# Role of Nucleoplasmin Family of Histone Chaperones, NPM1 and NPM2, in the Regulation of Transcription

A Thesis Submitted for the Degree of

Doctor of Philosophy

By

*Suchismita Dey*


To

Molecular Biology and Genetics Unit,
Jawaharlal Nehru Centre for Advanced Scientific Research,
(A Deemed University)
Jakkur, Bangalore – 560064, India.

December 2019

*Dedicated to my brother Sabyasachi, who has been my greatest support and source of inspiration*

# *Contents*

# *Declaration*

I hereby declare that this thesis entitled "Role of Nucleoplasmin Family of Histone Chaperones, NPM1 and NPM2, in the Regulation of Transcription" is an authentic record of research work carried out by me under the supervision of Prof. Tapas K. Kundu at the Molecular Biology and Genetics Unit, Jawaharlal Nehru Centre for Advanced Scientific Research (A Deemed University), Bangalore – 560064, India, and this work has not been submitted elsewhere for the award of any other degree.

In keeping with the standard norms of reporting scientific observations, due acknowledgments have been made wherever the work described was based on the findings of other investigators or assistance from other investigators was utilized towards the performance of experiments. Any omission, which might have occurred by oversight or misjudgment, is sincerely regretted.

Place: Bangalore Suchismita Dey

Date:

Prof. Tapas K. Kundu

Transcription and Disease Laboratory,

Molecular Biology and Genetics Unit,

Jawaharlal Nehru Centre for Advanced Scientific Research,

Bangalore – 560064, India.

# *Certificate*

This is to certify that the work described in this thesis entitled, "**Role of Nucleoplasmin Family of Histone Chaperones, NPM1 and NPM2, in the Regulation of Transcription**" is the result of the investigations carried out by Ms. Suchismita Dey in the Molecular Biology and Genetics Unit, Jawaharlal Nehru Centre for Advanced Scientific Research (A Deemed University), Bangalore – 560064, India, under my supervision, and the results presented in this thesis have not formed the basis for the award of any other previous diploma, degree or fellowship.

Place: Bangalore Prof. Tapas K. Kundu

Date:

# *Acknowledgments*

I take this opportunity to express my heartfelt gratitude to my mentor and research supervisor, Prof. Tapas K. Kundu. Although I have always felt that I was unable to live up to his expectations of performance, he was still very patient towards me, believed in my potential and capability to improve and perform better, and never gave up on me. Through all the criticisms, he always tried to bring out the best in me in terms of professional and personal development and I am sincerely grateful to him for that, that he always left room for me to improve further and grow better. I am thankful to him for involving me in the field of research which has been my interest since my student days in college, because of which I never felt too bored with my research work. I feel fortunate to have carried further, the legacy which my seniors and alumni – Dr. Venkatesh Swaminathan, Dr. Shrikanth Gadad, Dr. Jayasha Shandilya, and Dr. Parijat Senapati – have left behind. Their discoveries have been the foundations of my doctoral research and I have always admired and been inspired by them. Prof. Kundu gave me enough freedom to design my projects, explore new avenues in my research work independently, and think out of the box. I have learned a lot from his words of wisdom, and the philosophical and scientific discussions with him. His kindness, encouragement, support, and optimism have truly helped me in keeping a positive attitude towards work and life even during times of apparent failure and personal distress. He has the amazing perceptive quality of finding opportunities and making sense of scientific observations that appear inexplicable or meaningless to most others, and this has substantially helped me in taking my research work forward. I greatly admire his ambition, persistence and leadership skills which have fundamentally made him a successful scientist, team leader, and social worker. This has inspired me to contribute sincerely to the scientific community as well as to the general society by being part of social outreach programs. By working with him in a vibrant and enriching work environment, I have learned and inculcated several other qualities such as collaboration and coordination with team members, taking responsibility of workplace duties and for personal actions, soft skills and ethical behavior. It has been my honor and good fortune to be his Ph.D. student and I thank him for everything that he has done for me.

I thank Prof. M.R.S. Rao who has been a source of inspiration to me. I have admired his calm nature and immense knowledge and expertise in the field of chromatin biology. I have

learned important attributes for practicing good scientific research during my discussions with him. He has also offered valuable suggestions towards my research work which have been greatly helpful to me. I thank Prof. Ranga Uday Kumar for all his help, support and encouragement towards my work. He is a very methodical scientist and his attention to details and pursuit for root cause analysis, are a few things that I have found praiseworthy and inspiring. During his tenure as the chairman of the department, he has brought significant improvements in the central instrumentation facilities, general and biohazard safety regulations, the scientific and social temperament of the student community, all of which have collectively helped me in conducting research in a well-equipped, safe and good work environment. I also acknowledge the other faculties of the Molecular Biology and Genetics Unit and the Neuroscience Unit of JNCASR, Prof. Anuranjan Anand, Prof. Namita Surolia, Prof. Hemalatha Balaram, Prof. Maneesha Inamdar, Prof. Kaustuv Sanyal, Dr. Ravi Manjithaya, Dr. Sheeba Vasu, Dr. James P. Clement Chelliah, and Dr. Kushagra Bansal for the valuable suggestions, feedback and constructive criticisms which they have offered me regarding my research work during various research presentations and discussions, and also for the successful running and management of our departments and its facilities.

I am indebted to my senior Dr. Parijat Senapati for all her help, support, guidance, and encouragement towards me. She has taught me most of the techniques for carrying out my research and helped and guided me in designing and troubleshooting my experiments. Several leads which were obtained from her research work have contributed substantially to my thesis and she has been kind enough to share them with me. Specifically, I thank her for all the high-throughput genome-wide sequencing analyses which she has done for her manuscript and my thesis. I have learned a lot in several avenues from being a part of her manuscripts. Even in her busy schedule, she took time out to discuss with me about work and other matters which helped me in keeping my morale high and continue striving until I succeed. She always gave me valuable and honest advice about research presentations, career plans and otherwise which helped me make wiser decisions. I feel truly fortunate to have had her as a friend and my senior for my Ph.D. research.

I feel grateful for having been in the midst of friendly and helpful senior colleagues. I am thankful to Dr. Sadhan Das for his scientific contribution to the research work which I have been a part of. He has contributed greatly to the generation of reagents, designing, conducting and troubleshooting experiments, data analysis, and manuscript preparation. He

has always given me his best regards and offered to help me in whichever way I needed. I acknowledge the help of Deepthi Sudarshan, Dr. Surabhi Sudevan, and Dr. Manoj Kumar for some of my experiments. I also thank Dr. Stephanie Kaypee, Dr. Amit Behera, and Dr. Arnab Bose for sharing a few reagents generated by them, to conduct some of my experiments. All of them, as well as my other senior colleagues Dr. Dhanasekaran Karthigeyan, Dr. Amrutha Swaminathan, Dr. Mahadeva Swamy, and Dr. Sweta Sikder, have been truly supportive and encouraging, and have given valuable intellectual inputs in my work for which I am grateful to them. I also thank my junior colleagues – Aditya Bhattacharya for contributing towards and co-authoring our manuscript, Ila Joshi and Nikhitha Sreenivas for helping me with a few experiments, and Vinay J. Rao for doing some bioinformatics database analyses for me.

My sincere gratitude goes to Mune Gowda, Sunil Kumar, and Salauddin Sekh for their support and assistance in the lab and catering to our day-to-day needs. I thank Kruthi H.T. and all the previous secretaries of our laboratory for the efficient running and management of our laboratory. On the same note, I also acknowledge all the other members of our laboratory, past, and present, for their contributions towards the management and sustenance of the workplace.

The core facilities of our department have been greatly useful to me in conducting my research work. I thank Dr. R.G. Prakash at the Animal House facility for handling the animals (mice and rabbits) used for my studies, Ms. Suma Bangalore at the Confocal Imaging facility for performing the confocal imaging related to my experiments, Ms. Anitha at the Genomics facility for carrying out the sequencing of my constructs, and Dr. Narendra Nala at the Flow Cytometry facility for performing cell sorting techniques for my experiments. I acknowledge the faculty and student in-charges of the equipment at the Central Instrumentation Facility, Ms. Vidhya and all the previous secretaries of our department, safety officers of the institute Mr. Arun R and Ms. Lakshmi R, officers and staff at the institute's Academic, Purchase, Admin, Computer Laboratory, Library, Hostel, Dining Hall, Hostel mess, and Security sections, doctors and nurses at Dhanvantri, and everyone else at JNCASR for their services, help and cooperation whenever needed.

I acknowledge Eric Liao and the team at Quick Biology Inc. (Pasadena, CA, USA) (via Science Exchange, Palo Alto, CA, USA) for providing us the service of RNA-sequencing, Prof. Kazuhiko Igarashi and Dr. Hiroki Shima at the Tohoku University, Sendai, Japan, for

letting us use their Mass Spectrometry facility for an experiment, and the Proteomics Facility at the Molecular Biophysics Unit, Indian Institute of Science, Bangalore, India for their mass spectrometry (MALDI-TOF) service. I thank various scientists for providing us several reagents for our research work – Dr. Sagar Sengupta (National Institute of Immunology, New Delhi, India, for providing the mammalian expression construct of human c-jun and the c-jun antibody), Prof. Bert Vogelstein (Johns Hopkins University, Baltimore, USA, for providing the pCMV-wtp53 construct), Dr. K. Luger (University of Colorado, USA, for providing the bacterial expression constructs of *Xenopus* core histones), Prof. Purnima Dubey (Wake Forest School of Medicine, North Carolina, USA, for providing the mammalian expression construct of human YY1), Dr. Gautam Sethi (National University of Singapore, for providing the UM-SCC-1 cell line), Dr. Dirk Gründemann (University of Cologne, Germany, for providing the pEBTetD construct), Dr. Amit Dutt (Advanced Centre for Treatment, Research and Education in Cancer (ACTREC), Mumbai, India, for providing the AW13516, AW8507, OT9, NT8e and Het-1A cell lines), Prof. Susanne M. Gollin (University of Pittsburgh, USA, for providing the UPCI:SCC40, UPCI:SCC122 and UPCI:SCC-29B cell lines), Dr. Mitsuru Okuwaki (University of Tsukuba, Japan, for providing the human NPM2 expression construct), and Prof. Arturo Muga (Biofisika, Spain, for providing the *Xenopus* Nucleoplasmin expression construct).

I thank our collaborators Dr. Tessy T. Maleikal (Cancer Research Program, Rajiv Gandhi Centre for Biotechnology, Trivandrum, India) for conducting experiments related to the orthotopic animal model, Dr. K.S. Gopinath (HCG Cancer Centre and Bangalore Institute of Oncology, Bangalore, India) and Dr. Azeem Mohiuddin (Sri Devaraj Urs Academy of Higher Education and Research (SDUAHER), Kolar, India), for providing oral cancer patient samples for our research and helping with patient sample related data analysis. I am very grateful to Prof. Jerry Workman (Stowers Institute for Medical Research, USA), Prof. Kazuhiko Igarashi (Tohoku University, Japan), and Dr. Sagar Sengupta (National Institute of Immunology, New Delhi, India) for their valuable inputs in my research work which have directly helped me in troubleshooting several experiments, explore new angles in my projects and improve the quality of my data.

I acknowledge the University Grants Commission (New Delhi, India) for my research fellowship (JRF and SRF) and JNCASR (Department of Science and Technology, Government of India) for providing financial support during the last year of my Ph.D. and

partial financial assistance towards my travel for attending an international conference in the USA to present my research.

I thank the students and members of our department for extending their help and suggestions in times of need. My special thanks go to my senior from my past alma mater, Dr. Alok K. Singh and my friends Shazia Parveen, Nilesh Sharma, Amreesh Parvez, and Madhav K. Sarkar for being supportive, encouraging and caring towards me, especially during the hard and demoralizing moments. Dr. Alok has also taught me several research techniques, guided me in learning and developing professional and ethical scientific temperament, encouraged me to aim for higher in life and not be afraid to take up bigger challenges. He believed in my potential to perform better, appreciated my positive attributes and encouraged me to value and utilize such traits to gain success in my career.

Whatever little I have been able to accomplish so far, would not have been possible without the influence and contributions of my past teachers. I thank my teachers from Jawaharlal Nehru University, New Delhi, India, Prof. R. Madhubala who was my master's dissertation supervisor, Prof. K. Natarajan, Prof. S.K. Goswami, Prof. R. Muthuswami, Prof. P.C. Rath, and Prof. A.K. Nandi, for the wonderful lectures which they delivered. Their teaching led me to develop my interest in the fields of biochemistry, molecular biology, and transcription regulation. My teachers at St. Xavier's College, Kolkata, India, Prof. Arup K. Mitra, Prof. Kasturi Sarkar, Prof. Sudeshna Shyam Chowdhury, and Prof. Debjani Dutta, and Mrs. Menon, from whom I took biology lessons prior to joining college, also played a big role in developing my fundamental base in biology, which have had a compound effect on my career as a biological researcher.

I cannot emphasize enough on the love and support I have received from my family, my mother, my father, my uncle and my brother, which was probably the single most important factor that helped me cope with the stress of pursuing Ph.D. My parents, who came from very humble background having got fewer opportunities themselves to pursue higher studies, always encouraged me to study more and establish myself in the society. My brother has been the greatest blessing for me which I realized most during my Ph.D. tenure when I started getting negatively affected, mentally and physically, due to the continuous stretches of apparent failures in my research. His constant philosophical and practical guidance was most instrumental in bringing the faith back in me and making me see things and situations with a different and positive perspective. He made me realize that the journey

with the life experience one gets while pursuing a Ph.D., is what matters more than the short-term gains. My family has always emphasized on doing good and paying attention to good Karma to be happy and successful in career and life, and I have tried to follow that advice throughout, which has helped me achieve stability and peace of mind during my Ph.D. tenure.

Finally, I am deeply grateful to God Almighty and His Divine Intervention for bestowing me with the strength, courage, and everything else that I needed to survive this phase, for bringing me in contact with all the people and situations at the specific times which seemed to have a deeper significance in my life, and for all the opportunities for my future endeavors. It has by far been the most humbling experience of my life and I feel fortunate to have lived it.

# *Chapter 1: Introduction*

## *1.1. Organization of the eukaryotic genome into chromatin and its evolution*

The DNA, which is the bearer of the hereditary information of a cell, is massively larger in eukaryotic species compared to that in prokaryotes. The size of the genomic DNA generally tends to increase with increasing complexity in the species as we go up higher in the species evolutionary ladder. To contain such a huge length of genomic DNA into the limited confines of the nucleus, the DNA is packaged into a compact structure that is brought about by the action of specific proteins. This superstructure is called the chromatin which consists of a nucleic acid component (the DNA) and a protein component (the histone proteins) (Kornberg 1977). Although primitive forms of genome packaging mechanisms are present in prokaryotes (such as the presence of histone-like proteins in bacteria), basic mechanisms of eukaryotic genome packaging appear to share more commonalities with the archeal counterparts where we find proteins which are structurally similar to eukaryotic histones (Talbert et al. 2019). In eukaryotes, the histones are assembled into an octameric structure comprising of four types of histone proteins in a definite stoichiometry i.e. two dimers of histones H2A and H2B [2(H2A-H2B)] and a tetramer of two H3 and H4 histones each [$(H3\text{-}H4)_2$]. A single octamer of these 'core' histones is wrapped around by ~146 bp of DNA in roughly two superhelical turns (Luger et al. 1997) to form the nucleosome core particle (NCP) (Section 1.1.2) which is the fundamental unit of chromatin. Two core particles are connected by a linker DNA which is approximately 34 bp long and is associated with another type of histone called H1 or the linker histone that helps in nucleosome locking. The core particle along with a linker DNA constituting roughly 200 bp of DNA in total, is technically called the nucleosome (Oudet et al. 1975), while the nucleosome together with the linker histone is termed as the chromatosome (Simpson 1978) (Section 1.1.2).

The genomic DNA packaged into nucleosomes resembles a 'beads on a string' structure (Olins and Olins 1974; Olins and Olins 1979). This structure which is about 11 nm in diameter, is further packaged into higher-order structures such as the 30-nm fiber (Widom and Klug 1985; Williams et al. 1986; Olins and Olins 2003; Robinson et al. 2006),

chromatin loops with diameter of about 300 nm which is compressed to some extent forming the 250 nm fiber, condensed chromatin of diameter of about 700 nm, and finally the metaphase chromosome of diameter of about 1400 nm (Kireeva et al. 2004) (Figure 1.1).

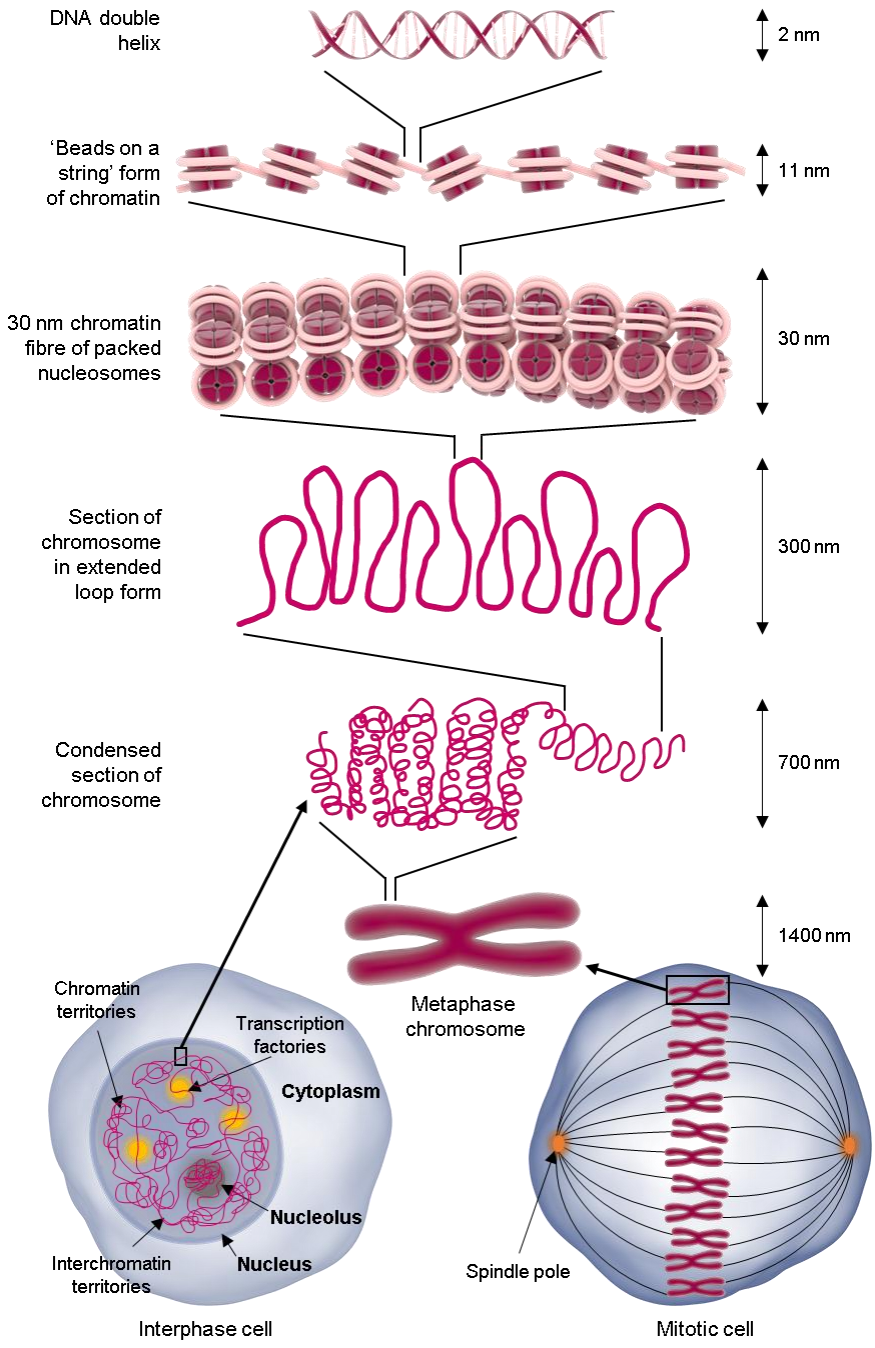



**Figure 1.1. Organization of the eukaryotic genome:** Schematic representation depicting the different levels of the packaging of eukaryotic genomic DNA (mentioned left), with the average diameter of the DNA and chromatin fibers indicated (mentioned right). Figure adapted from Senapati P, Ph.D. thesis, 2014.

The precise mechanisms of higher-order chromatin compaction are still largely obscure and the models which have been proposed in this field are majorly based on indirect evidence (Li and Reinberg 2011). However, with the advent of technologies such as cryogenic electron microscopy (cryo-EM) and higher-end X-ray crystallography, recent studies have been successful in gaining several insights into the structural aspects of low- and medium-order chromatin fibers which exist in different dynamic conformations. In a challenging study, Guohong Li and colleagues solved the cryo-EM structure of a 12-nucleosome array reconstituted in the presence of histone H1 at an 11-angstrom resolution, which has provided mechanistic insights into the compaction of nucleosomes to form the 30-nm fiber (Song et al. 2014). The structural analysis indicated that four nucleosomes are arranged back and forth in a zig-zag manner with a straight linker DNA to form a tetranucleosomal unit (Figure 1.2A – B). The presence of histone H1 in these units generates a left-handed twist (Figure 1.2A – B) such that the resulting chromatin fiber (reconstructed by the optimized docking of two tetranucleosome units (Figure 1.2C – D)) itself forms a double helical structure (Figure 1.2E) similar to that found in the DNA that it packages. This structure was a direct experimental validation of the proposed two-start zigzag configuration of the 30-nm fiber. However, other forms of chromatin fiber including the one-start solenoid structure may also exist under different conditions, such as in the presence of linker histone H5, different ionic concentrations, and different lengths of the nucleosome repeats as suggested previously (Robinson et al. 2006; Routh et al. 2008). For example, a recent study by Stefan Dimitrov and colleagues provides experimental evidence on the role of ionic conditions in determining the conformation of the chromatin (Garcia-Saez et al. 2018). In this study, the authors have solved the crystal structure of an H1-bound 6-nucleosome array at 9.7-angstrom resolution (Figure 1.3A – B). Using cryo-EM, biophysical and biochemical approaches, it was shown that this array adopted a less condensed ladder-like conformation that could transform into the twisted conformation in response to a shift in the ionic conditions (Figure 1.3C).

It should be noted that to date such models and mechanisms for higher-order chromatin structures and their interchangeable conformations have been elucidated based on evidence from *in vitro* reconstituted systems. The existence of such mechanisms *in vivo* is yet to be discovered. Multiple other factors such as histone modifications and histone variants in the nucleosomes are expected to play important roles in the regulation of these higher-order

chromatin structures through the modulation of the internucleosomal interactions between the multinucleosomal units.

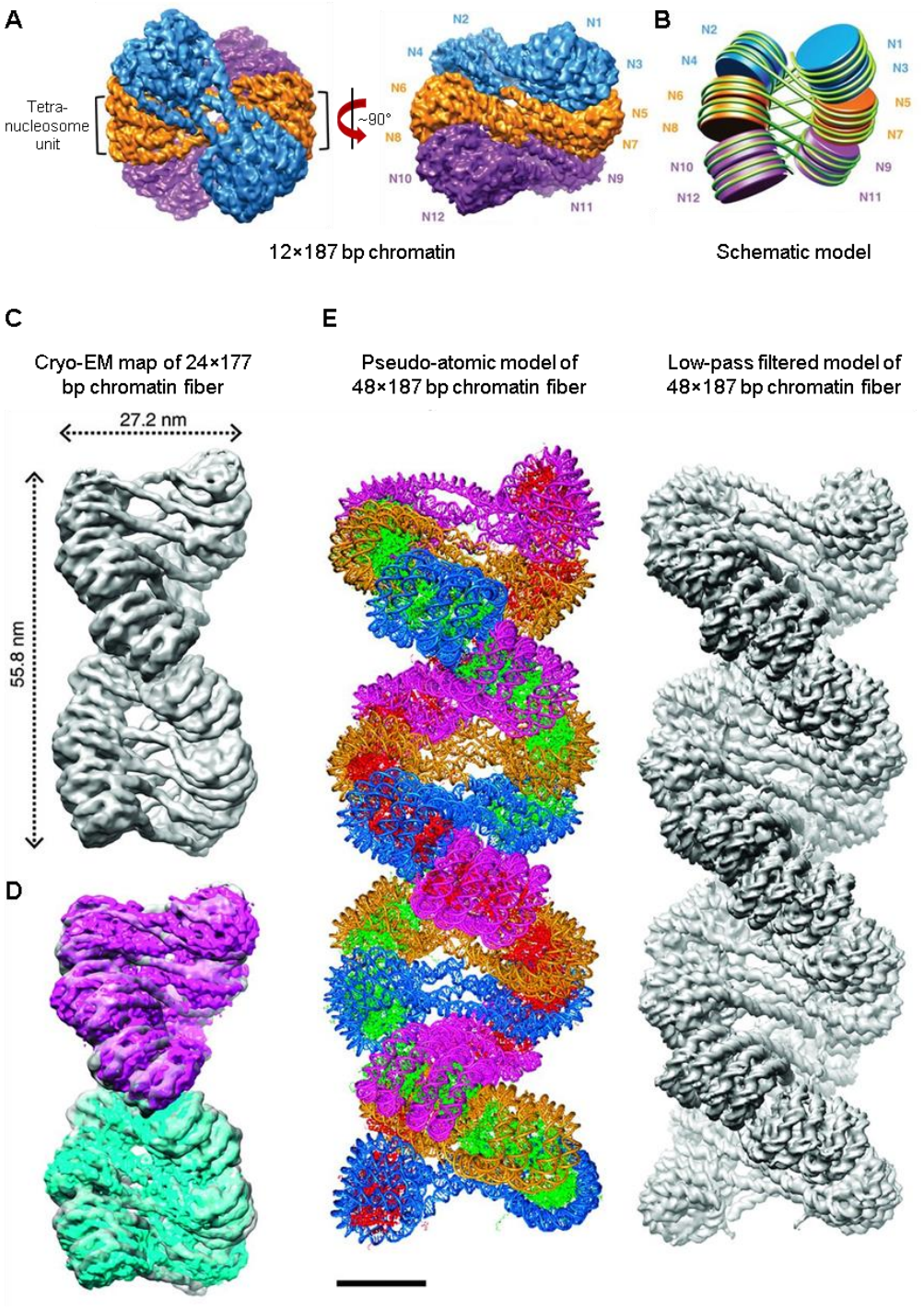



**Figure 1.2. Structure and model of the 30 nm chromatin fiber:** (A) 3D cryo-EM map of the 30-nm chromatin fiber reconstituted using a 12×187 bp DNA template showing the three tetranucleosomal structural units depicted in different colors and viewed from two different angles. (B) A schematic model of the cryo-EM structure of the 30-nm chromatin fiber. The nucleosomes are numbered. (C) The overall 3D cryo-EM map of the 30-nm chromatin fiber reconstituted using a 24×177 bp 601 DNA template. The length and diameter of the fiber are indicated. (D) The same structure of the 30-nm fiber as shown in (C) with two docked copies of the 12-nucleosome unit

structure shown in two different colors. (E) A reconstructed pseudo-atomic model (left, the structure of H1 is not included) and its corresponding density map low-pass filtered to 11 Å (right), created by directly stacking the cryo-EM structure of the dodecanucleosomal 30-nm fiber with 187-bp nucleosome repeat lengths on top of each other forming a continuous fiber. The scale bar is 11 nm. The figure has been obtained from (Song et al. 2014) and reused after obtaining copyright permission from the American Association for the Advancement of Science under license number 4718090760846.

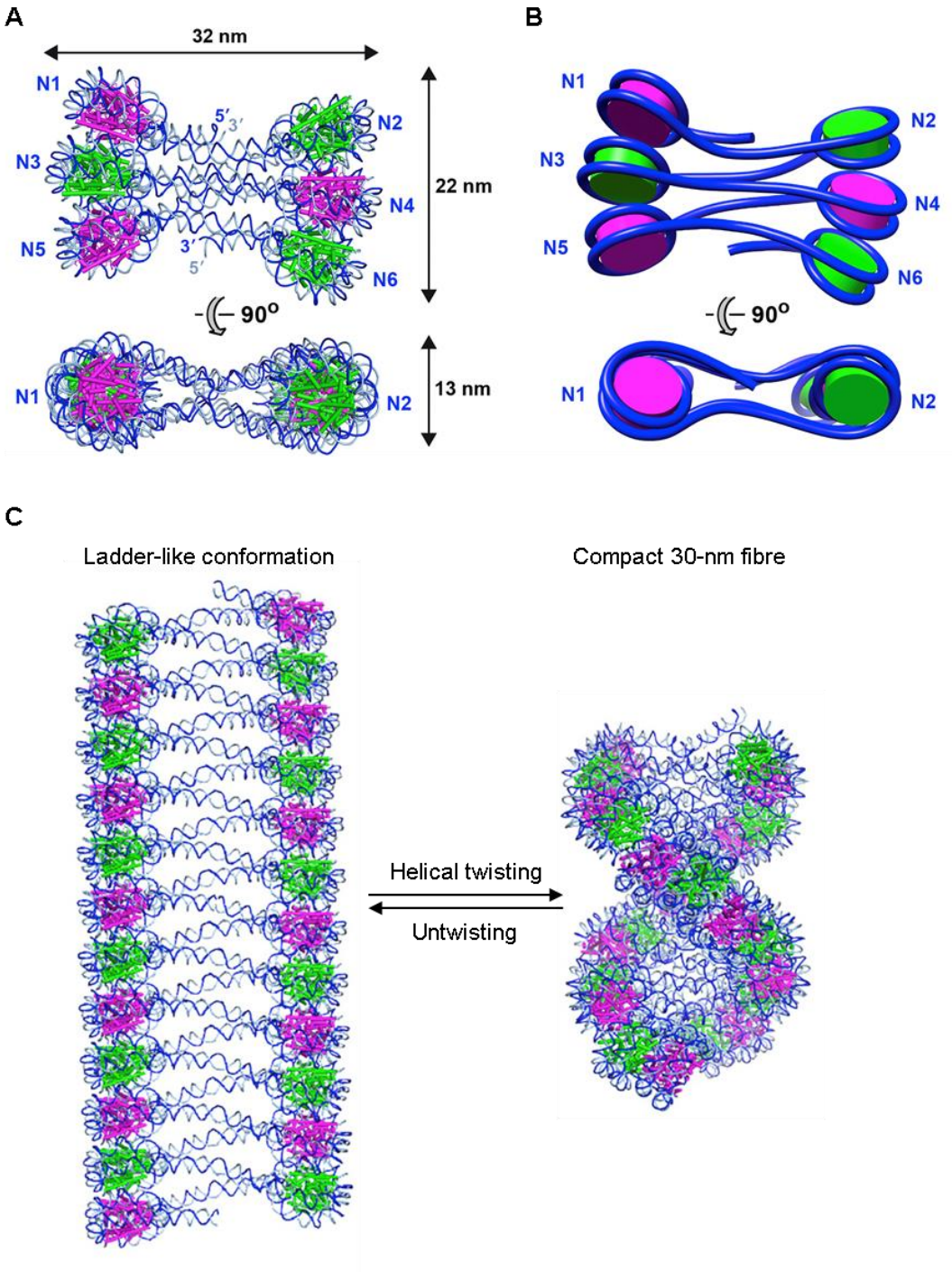



**Figure 1.3. Structural plasticity of the chromatin fiber:** (A) Crystal structure of the hexanucleosome. The core histones are depicted in magenta and green colors and the DNA in light and dark blue. The dimensions of the structure are indicated. (B) Cartoon model of the

hexanucleosome structure as shown in (A) highlighting the connectivity between nucleosomes. (A – B) The nucleosomes are numbered and the structures have been viewed in two different orientations. (C) Model depicting the reversible transformation between the open ladder-like conformation and the compact 30-nm fiber conformation of the chromatin that takes place with a shift in the ionic conditions of the environment. The figure has been obtained from (Garcia-Saez et al. 2018) and reused after obtaining copyright permission from Elsevier under license number 4718161227777.

### *1.1.1. Dynamic nature of the chromatin and its regulation*

In spite of the organized compaction of the eukaryotic chromatin, it is important that the chromatin be highly dynamic during the different DNA templated processes such as replication, repair, and transcription. This dynamicity is ensured and brought about by the action of various factors such as modulation of the chromatin structure by histone variants, covalent modifications of the histones, and methylation status of the DNA, as well as chromatin remodeling by the ATP-dependent remodelers and histone chaperones (Strahl and Allis 2000; Luger 2006; Luger et al. 2012; Becker and Workman 2013; Weber and Henikoff 2014; Soshnev et al. 2016). Various chromatin-associated proteins also regulate the state of chromatin compaction in different ways. Some of these above-mentioned factors follow certain common principles for their functioning. However, they also appear to have evolved to some extent in different eukaryotic species depending on the complexity of the genome and the requirement for the regulation of the chromatin-templated processes. The role and mechanisms of functioning of some of these factors are discussed below.

It was experimentally shown, using single-molecule fluorescence resonance energy transfer (FRET) for mononucleosomes, that DNA methylation induces the nucleosome to be more compact and rigid in nature (Choy et al. 2010) which explains their presence in heterochromatin or the tightly packed regions of the genome. Among histone modifications, histone H3 lysine 27 trimethylation (H3K27me3) and lysine 9 trimethylation (H3K9me3) correlate well with constitutive heterochromatin, while histone acetylation and histone H3 lysine 4 trimethylation (H3K4me3) are generally associated with euchromatin or the relatively accessible regions of the genome (Zhou et al. 2011). Likewise, histone variants such as H3.3 and H2A.Z usually mark the active regions of the chromatin. Two other variants namely macroH2A and H2A.Bbd, are also known to affect the accessibility of the chromatin for transcription in a context-dependent manner through their intrinsic and extrinsic effects on the stability and dynamics of the nucleosome (Li and Reinberg 2011; Zhou et al. 2011). In this context, the dynamics of the histone tails which

are the most common targets for the various post-translational modifications, play a significant role in modulating the higher-order organization of the chromatin (Fierz and Poirier 2019). Among other chromatin organizing factors are the linker histone H1 (which stabilizes the nucleosome and helps in the formation of stable higher-order chromatin structure (Thoma et al. 1979)), the high mobility group (HMG) proteins such as HMGB1 and HMGB2 (which loosen the chromatin through architectural alterations in the DNA structure (Thomas and Travers 2001; Thomas and Stott 2012)), heterochromatin protein 1 (HP1) (which recognizes the repressive histone modification, histone H3 lysine 9 di/trimethylation (H3K9me2/3) having consequent positive effects on heterochromatinization and gene silencing (James and Elgin 1986; Eissenberg and Elgin 2014; Sanulli et al. 2019)), positive coactivator 4 (PC4) (which induces chromatin condensation (Das et al. 2006; Das et al. 2010a)), MENT (also known as methylated in normal thymocytes, which can condense chromatin into unique secondary and tertiary structures (Springhetti et al. 2003)), the proto-oncogene DEK (which induces heterochromatinization through its interaction with HP1α and enhancement of its binding to the H3K9me3 mark on the chromatin (Waldmann et al. 2002; Waldmann et al. 2004; Kappes et al. 2011)), Polycomb group proteins (which mediate chromatin compaction (Francis et al. 2004)), methyl CpG binding protein 2 (MeCP2) (which brings about chromatin compaction through induction of secondary chromatin structures (Georgel et al. 2003)), the CCCTC-binding factor (CTCF) (which defines insulator regions of the genome and regulates the organization of the 3D genome in association with cohesin (Yusufzai et al. 2004; Parelho et al. 2008; Wendt et al. 2008; Rowley and Corces 2018; Braccioli and de Wit 2019)), poly(ADP-ribose) polymerase 1 (PARP1) (which can modulate the chromatin structure through various mechanisms such as its automodification (Kim et al. 2004b), histone PARylation (Gibbs-Seymour et al. 2016), inhibition of DNA methylation, association with CTCF and scaffold/matrix attachment regions (S/MARs) (Galande and Kohwi-Shigematsu 2000; Beneke 2012)), and non-coding RNA (which regulates chromatin organization through various mechanisms such as recruitment and regulation of chromatin modifying complexes, RNA-directed DNA methylation, association with and organization of nuclear matrix, insulator and boundary regions (Meller et al. 2015)), to name a few.

The collective contribution of the various factors discussed above facilitates the packaging and organization of the large genomic DNA into the chromatin, as well as the

compartmentalization of the chromatin into specific territories in three-dimensional space. The spatial organization of these chromatin territories also appears to be evolutionarily conserved (Tanabe et al. 2002). The euchromatin regions of the chromatin are usually found more towards the interior of the nuclear space while the heterochromatin regions are dispersed towards the nuclear periphery. The chromatin territories contain certain chromatin domains known as topologically associating domains (TADs). The size of these domains varies from several hundred kilobases to 1 – 2 megabases of nucleotides in the chromatin. These entities are characterized by a much higher frequency of interactions between distinct elements within the TAD (intra-TAD interactions) than those between different TADs (inter-TAD interactions) (Dekker and Heard 2015). With the development of techniques such as chromosome conformation capture (3C) and Hi-C, there has been extensive research regarding the functions of TADs in the regulation of gene expression and other nuclear processes.

### *1.1.2. Histone octamer and nucleosomes*

The nucleosome or more specifically, the nucleosome core particle, consists of the core histone octamer wrapped by ~146 bp of DNA in roughly two left-handed superhelical turns (Luger et al. 1997). The histone octamer comprises of two copies each of core histones H2A, H2B, H3, and H4 (Kornberg and Thomas 1974). These core histones proteins are also termed as 'canonical' which are expressed during the S-phase of the cell cycle and package the newly replicated DNA. The canonical histone coding genes are clustered in the chromosomes and are present in multiple copies. These proteins are highly conserved across eukaryotes and are abundant in their distribution. Occasionally, the core histones are substituted by similar proteins with certain distinctive features known as histone variants, which confer specific structural and functional properties to the nucleosome and consequently to the contextual chromatin. The histone variants are packaged with the DNA within the nucleosomes in a replication-independent manner and are synthesized throughout the cell cycle. The histone variant-encoding genes are usually single-copy genes and show a substantial degree of variation among species (Henikoff and Smith 2015).

Structurally, histone proteins are basic in nature and have a characteristic histone fold domain (HFD) within their protein sequences which is flanked by flexible N- and C-terminal tails. The average length of the HFD is about 70 amino acid residues and consists

of three alpha-helices connected by two short, unstructured loops (Arents and Moudrianakis 1995). In the core histone octamer, the H3-H4 tetramer is present in the middle whose each side is flanked by an H2A-H2B dimer (Luger et al. 1997). The linker histone H1 which is not as conserved across eukaryotic species as the core histones, is positioned on the dyad axis of the nucleosome and contacts the DNA entering and exiting the nucleosome core particle (NCP) (Noll and Kornberg 1977; Allan et al. 1980). Crystal structure of nucleosome with chicken linker histone H5 (Zhou et al. 2015) and cryo-EM structure of nucleosome with vertebrate linker histone H1 (Bednar et al. 2017) reveal that the globular domain of the linker histone interacts with the core DNA on the dyad axis and with both the DNA linkers, while its C-terminal domain is associated mainly with a single DNA linker. This arrangement collectively stabilizes and compacts the nucleosome.

### *1.1.3. Histone variants: Properties and functions*

As briefly mentioned in Section 1.1.2, histone variants are proteins that are structurally similar to their canonical counterparts, except for the presence of specific deviations in their amino acid sequences that confer distinct structural properties to the corresponding nucleosomes which have significant impacts on the downstream functions. Several variants have been identified for the core histones H3, H2A, and H2B, and the linker histone H1, as mentioned in the HIstome database (Khare et al. 2012). The functional aspects of some of the relevant and well-characterized histone variants are discussed briefly in the subsequent paragraphs.

There are numerous linker histone H1 variants across species and about eleven variants in mice and humans (Hergeth and Schneider 2015). The differences in these variants lie mainly in the non-globular N- and C- terminal regions (Kamakaka and Biggins 2005) which confer varying degrees of chromatin compaction ranging from weak (H1.1 and H1.2), intermediate (H1.3), and strong condensation (H1.0, H1.4, H1.5 and H1x) (Clausell et al. 2009). Histones H1.1 (H1a), H1.2 (H1c), H1.3 (H1d), H1.4 (H1e), and H1.5 (H1b) are expressed in most of the somatic cells in a replication-dependent manner (Marzluff 2005), while H1.0 ($H1^0$) and H1x (H1.10) are expressed in a replication-independent manner in terminally differentiated and/or tumor cells (Zlatanova and Doenecke 1994; Happel et al. 2005; Warneboldt et al. 2008). The other variants are testis-specific in their distribution such as H1oo (H1.8) which is expressed in the oocytes (Tanaka et al. 2001) and H1t (H1.6),

H1T2 (H1.7), and HILS1 (H1.9), which are expressed in the germ cells in testis (Drabent et al. 1991; Yan et al. 2003; Martianov et al. 2005; Tanaka et al. 2006; Hergeth and Schneider 2015).

H2A has been reported to have the largest number of variants among the core histones. These are H2A.Z, macroH2A, H2A.Bbd, and H2A.X (Kamakaka and Biggins 2005). The histone variant H2A.Z is evolutionarily conserved (Jackson et al. 1996). Its genomic localization varies depending on the context and hence is implicated both in transcriptional activation as well as repression (Giaimo et al. 2019). H2A.Z is often found to be enriched at the promoters of genes in the euchromatin and positioned in the two nucleosomes flanking the nucleosome-free region (NFR) that contains the transcription start site (TSS) (Raisner et al. 2005). It is also observed in the gene bodies as well as facultative heterochromatin in certain contexts such as transcriptional state, the activity of the histone chaperone FACT (SPT16 subunit) or SPT6, and so on (Hardy et al. 2009; Jeronimo et al. 2015; Lashgari et al. 2017). Such context-dependent genomic localization and functions related to transcriptional activation or repression could be attributed to the stability of the H2A.Z containing nucleosomes that is modulated by its different post-translational modifications, structural influence from other histones in the nucleosome such as H3, H4, and H3.3, and interactions with chromatin remodelers, transcriptional regulators, and histone chaperones (Fan et al. 2004; Draker and Cheung 2009; Thakar et al. 2009; Li and Reinberg 2011; Draker et al. 2012; Weber and Henikoff 2014; Dai et al. 2017; Zhang et al. 2017; Cakiroglu et al. 2019; Giaimo et al. 2019). Collectively, these properties of H2A.Z are implicated in transcriptional control, prevention of the ectopic spread of heterochromatin, and genome integrity.

The H2A variants macroH2A and H2A.Bbd, are found in vertebrates or mammals (Kamakaka and Biggins 2005). macroH2A localizes mainly in the inactive X-chromosome (Costanzi and Pehrson 1998), while H2A.Bbd is localized to the active X-chromosome and autosomes (Chadwick and Willard 2001). Structurally, macroH2A has more than 200 amino acid residues in addition to its sequence conserved with H2A, forming a C-terminal non-histone globular domain (Ladurner 2003). macroH2A is generally associated with transcription repression and heterochromatinization brought about through different mechanisms (Douet et al. 2017). However, a subset of genes such as serum-responsive genes, seem to be positively regulated by the presence of macroH2A in the autosomes (Gamble et al. 2010). The variant H2A.Bbd is considerably shorter in length than H2A and

lacks the C-terminal tail and part of the docking domain (Bönisch and Hake 2012). H2A.Bbd containing nucleosome wraps only about 118 bp of DNA (Bao et al. 2004) and is relatively less stable (Gautier et al. 2004). This explains its association with active transcription and enrichment at the TSS of genes (Doyen et al. 2006; Soboleva et al. 2011; Tolstorukov et al. 2012). Besides transcription, H2A.Bbd is also associated with cell cycle regulation, DNA damage repair, (Sansoni et al. 2014), as well as mRNA processing (Tolstorukov et al. 2012; Sansoni et al. 2014; Soboleva et al. 2017). H2A.Bbd is normally most highly expressed in the testis having roles in male germ cell development (Ishibashi et al. 2010; Talbert and Henikoff 2010), followed by brain, and sometimes aberrantly expressed in cancers such as lymphoma (Winkler et al. 2012).

The variant H2A.X is ubiquitously expressed and is distinguished by the presence of a C-terminal amino acid sequence motif, SQ(E/D)Ø, where Ø indicates a hydrophobic amino acid. The serine in this sequence motif is the site of phosphorylation mediated by members of the phosphoinositide 3-kinase (PI3K) family, producing a modified protein designated "γH2A.X" (Henikoff and Smith 2015). H2A.X phosphorylation is induced at the sites of DNA breaks upon DNA damage, which then helps in marking those sites for the recruitment and retainment of the DNA repair machinery, chromatin remodeling complexes such as INO80, chromatin-modifying enzymes such as Tip60, as well as stabilizers of chromatin structure such as cohesin (Lowndes and Toh 2005; Morrison and Shen 2005; Kuo and Yang 2008). Besides the universal function of DNA double-strand break (DSB) repair by γH2A.X, H2A.X phosphorylation in XY bivalent which is distinct from γH2A.X plays an important role in male germ cell development through its effect on chromatin remodeling and inactivation of sex chromosomes (Fernandez-Capetillo et al. 2003).

H2B has numerous variants across eukaryotic species and about 19 in humans. However, the H2B variants are considerably less studied and characterized. Few of them have been reported to have specialized functions related to chromatin compaction and transcription repression during gametogenesis (Kamakaka and Biggins 2005). For example, the testis-specific H2B variant TH2B along with the H2A variant that is TH2A, is associated with the chromatin dynamics in sperms and pluripotent cells, thus having important implications in male fertility (Shinagawa et al. 2015; Kutchy et al. 2017) as well as pluripotent stem cell generation (Huynh et al. 2016).

There are two major histone H3 variants which have been substantially studied. CENP-A is the H3 variant that is localized to the centromere of a chromosome, the locus of the DNA where the kinetochore protein complex is assembled that mediates chromosome segregation during eukaryotic cell division (Kamakaka and Biggins 2005). The ortholog of mammalian CENP-A which is found in other eukaryotes is generally referred to as CenH3 (Talbert and Henikoff 2010). CenH3s mark the identity of the centromeres and are essential for kinetochore assembly and chromosome segregation (Amor et al. 2004). In the field of the structure of CENP-A-nucleosome, different models such as octasome, hemisome, compact octasome, hexasome, and tetrasome, and their implications in the centromeric chromatin architecture, have been proposed in the course of time (Tachiwana and Kurumizaka 2011). However, human CENP-A-nucleosome was shown to be a homotypic, octameric structure (Nechemia-Arbely et al. 2017) and whose flexible ends modulate the interaction of the nucleosome with linker histone H1, thereby affecting the level of chromatin compaction as well as ensuring the fidelity of the mitotic process (Roulland et al. 2016). The other major H3 variant is H3.3 which is generally found at active chromatin including TSS of genes, enhancers, as well as gene bodies (Ahmad and Henikoff 2002; Jin et al. 2009; Goldberg et al. 2010). H3.3 and another variant, H3.4, have only four amino acid residues different compared to H3 or H3.1, while the isoforms H3.1 and H3.2 differ by a single amino acid which does not impart much difference in the functional properties of the two isoforms (Kamakaka and Biggins 2005; Henikoff and Smith 2015). While the deposition of H3.3 in the active chromatin regions is generally mediated by the histone chaperone HIRA, H3.3 is also found to be deposited at telomeres and pericentric heterochromatin by the action of the histone chaperones DAXX and ATRX (Goldberg et al. 2010; Santenard et al. 2010; Szenker et al. 2011). The regulation of chromatin dynamics by H3.3 has been implicated in the germline and cell fate transition (Henikoff and Smith 2015; Fang et al. 2018). The variant H3.4, also known as H3t, is a testis-specific histone that is essential for entry into spermatogenesis (Ueda et al. 2017). Two other primate-specific variants of H3 namely H3.X and H3.Y, are expressed in several normal and tumor cell types and are involved in the regulation of genes expressed in response to cellular stresses (Wiedemann et al. 2010). H3.Y is enriched at the TSS of genes and is involved in the regulation of transcription in human cells which is due, at least, to its less efficient interaction with linker histone H1 compared to that by H3.3. This could potentially result in a less compacted chromatin state (Kujirai et al. 2016).

H4 is one of the slowest evolving proteins and there is no known variant of this protein to date (Kamakaka and Biggins 2005). Recently, however, a novel Hominidae-specific H4 variant called H4G has been identified which is expressed in a variety of human cell lines and exhibits tumor-stage-dependent overexpression in tissues from breast cancer patients. This variant is primarily localized to the nucleolus where it can enhance rDNA transcription in breast cancer (Long et al. 2019).

In addition to the structural modulation of nucleosome brought about by the structures of the histone variants themselves, the various PTMs of the variants also affect the chromatin dynamics. Collectively, they serve as epigenetic markers for the recruitment of transcriptional coactivators and factors, which modulate gene expression. Understanding the complex roles of histone variant incorporation, their PTMs, and their turnover in the epigenetic control of chromatin organization, and their implications in disease and development is one of the active areas of research today.

## *1.2. Eukaryotic transcription: A global perspective of the basic mechanisms and their regulation*

For the survival and propagation of the living cell, the integrity of the genomic DNA is required to be maintained along with its proper replication during cell division and its transcription for gene expression. At the core of these basic processes is the need for the respective enzymatic machinery to access the genomic DNA. The packaging of the eukaryotic DNA into nucleosomes to form the chromatin poses a fundamental barrier to such machinery to access the DNA in the desired and specific regions. Hence the eukaryotes have developed a multitude of mechanisms and factors to mediate this process which has evolved in various species in the course of time but following some basic conserved principles. In the case of prokaryotes, we do not find such extensive packaging of the genomic DNA into chromatin. Hence there are some fundamental differences in the mechanisms of the DNA templated processes between prokaryotes and eukaryotes. In the context of transcription, these differences are well explained by the concept of the transcriptional 'ground state' which is defined as the inherent activity of the promoters (and hence the core transcription machinery) *in vivo* without the influence of other regulatory factors (Struhl 1999). In prokaryotes, there is no inherent restriction to access the DNA template by the RNA Polymerase for initiating transcription *in vivo* and hence ground state

of the DNA template is 'non-restrictive' (Figure 1.4). In the eukaryotes, due to the packaging of the DNA template into chromatin, the transcriptional ground state is 'restrictive' as a result of which the core transcriptional machinery cannot access the promoters by themselves. The prokaryotic DNA can be easily transcribed from the ground state by the influence of activators in case of weak promoters, or be repressed actively by the action of repressors (Figure 1.4). On the other hand, the eukaryotic chromatin can have multiple alternative states such as the 'silent state' where the chromatin is bound by repressors and further compacted by different mechanisms to reduce or inhibit gene expression, the transcriptional 'ground state' which is 'restrictive' and where there is an absence of a repressor or an activator, a 'poised state' where the chromatin-modifying enzymes have displaced the histones and loosened up the chromatin to make it accessible to the transcription machinery, and the 'active state' where the basal transcription machinery is recruited to the promoter for initiation of gene transcription (Struhl 1999).

The eukaryotes have developed and evolved multiple factors and mechanisms to overcome the nucleosomal barrier during the process of transcription, as well as control and fine-tune the level of transcription. Such factors include transcription factors, coactivators, chromatin remodelers, chromatin modifiers, and histone chaperones, apart from the multiple components of the transcription machinery (Roeder 1998; Roeder 2005). Many transcription factors which direct the expression of genes, are themselves regulated by environmental cues or signals such as physical, developmental, hormonal, pharmacological, stress, and so on (Brivanlou and Darnell 2002). Even the binding of the factors to their target DNA sequence could be near the gene or hundreds of kilobases away from the gene, which then work through direct or indirect long-range mechanisms such as recruitment of ancillary factors, to make the region in the chromatin permissive for transcription by the RNA Polymerase complex (Kleinjan and van Heyningen 2005). Further, several types of non-coding RNAs such as miRNAs, long non-coding RNAs have emerged as important players in the complex regulation of gene expression (Catalanotto et al. 2016; Fernandes et al. 2019). The physical properties of the proteins, the chromatin, the nuclear matrix and the fluid component of the cellular compartments also play a very important role in creating the environmental niche for transcription to occur by forming phase-separated condensates which can concentrate the different factors in space and act as hubs for the transcription factories (Boehning et al. 2018; Lu et al. 2018; Cramer 2019; Gibson et al. 2019).

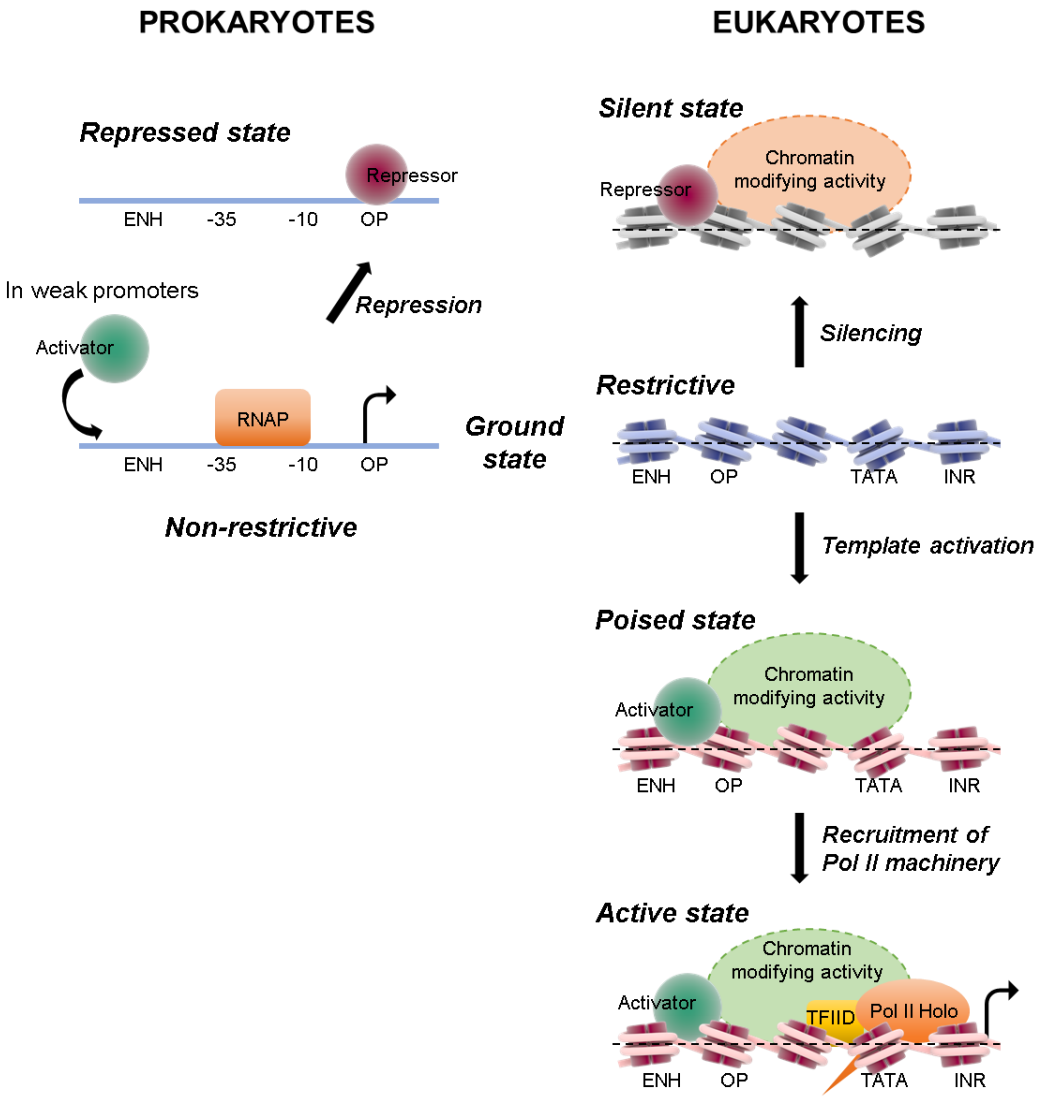



**Figure 1.4. Transcriptional states in prokaryotes and eukaryotes:** The transcriptional ground state in prokaryotes is non-restrictive which can be easily transcribed by the RNA Polymerase (RNAP), sometimes with the aid of activators in case of weak promoters, or be converted to a repressed state by the action of repressors. However, in eukaryotes the DNA is packaged into chromatin, causing an inherent restriction to the RNA Polymerase II machinery for accessing the DNA template which renders the ground state restrictive. The restrictive state of the chromatin (blue) is actively converted to a more silent state (grey) by the action of repressors and repressive chromatin-modifying activities (orange), or to an active state with an intermediate poised state (red) by the action of activators and activating chromatin-modifying activities (green). The latter state allows the recruitment of the RNA Pol II machinery (TFIID + Pol II holoenzyme) at the promoter. ENH: enhancer sequence, OP: operator sequence, TATA: TATA-box, INR: initiator sequence. Figure adapted from (Struhl 1999).

Despite the diversification in the mechanisms of eukaryotic transcription and its regulation, the core transcriptional machinery and the fundamental principles and mechanisms of transcription in prokaryotes and eukaryotes exhibit a marked degree of evolutionary

conservation (Ebright 2000; Cramer 2002). However, due to the evolution of the structures of various eukaryotic genes, eukaryotic RNA Polymerase has diversified and evolved to transcribe specific classes of these genes (Roeder and Rutter 1969; Roeder and Rutter 1970; Roeder 2003). These are RNA Polymerase I (which transcribes ribosomal DNA), RNA Polymerase II (which transcribes most of the genes including protein-coding as well as several non-coding genes), and RNA Polymerase III (which transcribes genes encoding the 5S rRNA, tRNA, and other small RNAs). Two other polymerases have also been identified in plants and are known as RNA Polymerase IV and V (which synthesizes silencing RNA (si-RNA) in plants) (Pikaard et al. 2008). The subsequent few sections would be devoted to brief discussions of RNA Polymerase II-mediated transcription. However, the other polymerases employ similar and analogous mechanisms of transcription, using functionally equivalent as well as some shared factors.

### *1.2.1. RNA Polymerase II: Basic structure and mechanisms of promoter recognition*

The RNA Polymerase II (RNA Pol II) is a multi-subunit enzyme complex that transcribes DNA to form protein-coding mRNA, small nuclear RNA (snRNA) and micro RNA (miRNA). The RNA Pol II holoenzyme complex is about 514 kDa in size and has 12 subunits in yeast and humans (Cramer et al. 2008). The largest subunits are RPB1 and RPB2 which are the orthologs of the β' and β subunits of the prokaryotic RNA Polymerase respectively. They form the core element with the central large cleft, the clamp element that moves to open and close the cleft, and the jaws that are thought to grab the incoming DNA template (Cramer et al. 2008). This catalytic site consists of highly conserved residues and magnesium ions buried deep within the cleft (Cramer 2002). The subunits RPB3 and RPB11 heterodimerize to form a part of the core complex which is functionally analogous to the prokaryotic RNA Polymerase alpha subunit homodimer that helps in the assembly of the RNA Polymerase complex. These four subunits namely RPB1, RPB2, RPB3, and RPB11 form the core of the eukaryotic RNA Pol II which is very similar to the prokaryotic RNA Polymerase (Zhang et al. 1999; Cramer et al. 2000). The subunit RPB6 is the sequence, structural and functional ortholog of the ω subunit of the prokaryotic RNA polymerase (Minakhin et al. 2001) which was found to have a structural role in the

maintenance of the conformation of the RNA Polymerase β' subunit, and the recruitment of β' to the enzyme assembly (Mathew and Chatterji 2006).

The eukaryotic RNA Pol II is capable of synthesizing RNA from the DNA template. However, it is not competent to initiate the process of transcription from the transcription start sites (TSS) of the genes. For this purpose, the eukaryotic system has developed ancillary factors known as general transcription factors (GTFs) which are functionally analogous to the sigma factors in the prokaryotic system. They assist in the proper orientation of the RNA Pol II on the cognate promoter to initiate transcription. These GTFs are TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH proteins (Orphanides et al. 1996; Roeder 1996; Hahn 2004; Thomas and Chiang 2006). Some of these GTFs are multi-subunit proteins such as TFIIA (with 2 subunits), TFIID (with 15 subunits including TBP and TAFs), TFIIE (with 2 subunits), TFIIF (with 3 subunits), and TFIIH (with 10 subunits) (Hahn 2004). By the concerted and sequential actions of the GTFs, the RNA Pol II is assembled at the promoter to form the pre-initiation complex (PIC) which is capable of initiating transcription (Thomas and Chiang 2006).

Promoter sequences possess certain physical properties, such as unusual DNA structures and low stability, which distinguishes them from the rest of the genome. Analysis of promoters from diverse organisms among bacteria, vertebrates, and plants have revealed that all these promoter sequences shared certain features, such as stability and bendability profiles, but had notable differences in their DNA curvature profiles and nucleotide composition (Kanhere and Bansal 2005). Besides this, the chromatin architecture including the epigenetic landscape around the promoters is also distinguishable from the rest of the genome (Haberle and Stark 2018). For example, the TSS of a gene is generally characterized by the presence of a nucleosome-free region (NFR), also known as nucleosome-depleted region (NDR), and the enrichment of histone modification marks H3K4me3 and H3K27ac, which affect the stability of the chromatin region at the promoter. The sequence structure of the eukaryotic promoter has also diversified in the course of evolution. However, certain conserved motifs are often found in the majority of the promoters, such as the TATA-box, Initiator element (INR), downstream promoter element (DPE), downstream core element (DCE), motif ten element (MTE), upstream TFIIB recognition element ($BRE_u$), downstream TFIIB recognition element ($BRE_d$), polypyrimidine initiator (TCT), X core promoter element 1 (XCPE1) (Tokusumi et al. 2007), X core promoter element 2 (XCPE2) (Anish et al. 2009), cell cycle-dependent

element (CDE)/cell cycle genes homology region (CHR) (Muller and Engeland 2010), dehydration-responsive element (DRE) (Maruyama et al. 2012), among others (Roy and Singer 2015; Haberle and Stark 2018). TATA, INR, DPE, DCE, MTE, and BRE are generally found in the promoters for most mRNA and miRNA genes, and hence known as canonical core promoter elements. There are also, TATA-less promoters, which generally give rise to some mRNA and other non-coding RNA such as Piwi-interacting RNA (piRNA), transcription initiation associated RNA (tiRNA), and transcription start site associated miRNA (TSSmiRNA). These promoters have elements such as CpG islands, ATG deserts, and transcription initiation platforms (TIPs), which are known as non-canonical promoter elements (Roy and Singer 2015).

The order of assembly of the GTFs to form the PIC has mostly been studied in the context of TATA containing promoter. However, several genes lack one or more of the core elements including the TATA element but may contain a subset of the other core elements, based on which there are different but analogous mechanisms of transcription initiation (Juven-Gershon et al. 2008). Although the promoters for RNA Pol I and Pol III are different from those of RNA Pol II promoter, all the RNA Polymerases employ the TATA-binding protein (TBP) or a TBP-like protein for the assembly of the PIC. The assembly and activity of the PIC control gene activity to a significant degree and are subject to a substantial level of regulation (Roeder 1998).

### *1.2.2. Transcription initiation*

According to the classical models of transcription based on *in vitro* biochemical studies, the initiation of transcription is marked by the ordered assembly of the pre-initiation complex (PIC) at the promoter of the gene (Buratowski 1994; Orphanides et al. 1996; Lemon and Tjian 2000). This begins with the binding of the TBP subunit of the TFIID complex to the TATA element through its interactions with the minor groove (Kim et al. 1993a; Burley and Roeder 1996). Structurally, the TATA-box recognition domain of TBP known as the TBP core, upon binding to the DNA, unwinds and sharply bends it to form a unique saddle-shaped structure over the bent DNA (Kim et al. 1993a; Kim et al. 1993b; Juo et al. 1996; Nikolov et al. 1996). This process is promoted by TFIIA which binds upstream to the TATA box, interacts with TBP and stabilizes the TBP-TATA DNA complex (Buratowski et al. 1989; Lee et al. 1992; Imbalzano et al. 1994; Geiger et al. 1996;

Tan et al. 1996). Besides this, TFIIA also counteracts several negative regulators of transcription which target TBP (Kim et al. 1996; Kokubo et al. 1998). The binding of TBP to TATA box was shown to take place in two steps – formation of an unstable complex between TBP and the unbent DNA, followed by a slow formation of a stable complex with the bent DNA, a process which is greatly accelerated by the TFIIB factor (Nikolov et al. 1995; Zhao and Herr 2002). TFIIB contacts the DNA both upstream and downstream of the TATA box at the TFIIB recognition element (BRE) (Lagrange et al. 1998). It contacts the TBP through its C-terminal domain (Nikolov et al. 1995) while its N-terminal domain interacts with and recruits the RNA Pol II (Kostrewa et al. 2009; Liu et al. 2010) bound with TFIIF to the core promoter (Flores et al. 1992). In the case of TATA-less promoters, the other subunits of the TFIID complex namely the TBP associated factors (TAFs) mediate the positioning of the TFIID through interactions with other sequence motifs or factors present in the vicinity of the core promoter (Martinez et al. 1995; Burke and Kadonaga 1997; Martinez et al. 1998). As an additional mechanism, there are also cell-type-specific TBP-related factors (TRFs), one of which was found to form a stable TRF-TFIIA-TFIIB-TATA DNA complex and substitute for TBP in directing RNA polymerase II transcription *in vitro* (Hansen et al. 1997).

Following the binding of the TFIIF-RNA Pol II to the promoter, the TATA DNA-TBP-TFIIB complex is further stabilized and TFIIE and TFIIH are recruited. The binding of TFIIE and TFIIH at the promoter along with the other assembled proteins results in the firing of the complex (Buratowski et al. 1989; Flores et al. 1992; Ohkuma and Roeder 1994). Mechanistically, TFIIE stimulates the TFIIH-dependent kinase activity that phosphorylates the CTD of the RPB1 subunit of RNA pol II. TFIIH also possesses a helicase activity. TFIIH can phosphorylate specific factors among the GTFs as well (Ohkuma and Roeder 1994; Lin and Gralla 2005). The function of the TFIIH helps in promoter melting that partially unwinds the TSS and converts the closed promoter complex of the PIC to an open promoter complex. This step is ATP-dependent and is controlled by the eleven subunits of the TFIIH complex. The helicase activity of the TFIIH is present in its ERCC3 (XPB) subunit that uses ATP hydrolysis to unwind the promoter DNA by a torsional activity (Tirode et al. 1999; Kim et al. 2000; Lin et al. 2005).

The mechanism of PIC assembly and its role in transcription has been modeled over the years through various biochemical studies and insights from the partial crystal and/or cryo-EM structures of the PIC (Cheung and Cramer 2012) such as RNA Pol II (Armache et al.

2005), RNA Pol II-TFIIB (Bushnell et al. 2004; Kostrewa et al. 2009; Liu et al. 2010; Sainsbury et al. 2013), RNA Pol II-TFIIS (Kettenberger et al. 2003), TFIIA-TBP-DNA (Tan et al. 1996), TFIIB-TBP (Tsai and Sigler 2000), TFIIE (Miwa et al. 2016), RNA Pol II-TFIIF (Chen et al. 2010), core TFIIH (Chang and Kornberg 2000; Greber et al. 2019), and so on. To obtain a more comprehensive understanding of the PIC, studies were conducted by some groups in the recent past to solve the structure of the complete PIC using cryo-EM approach combined with chemical crosslinking. One such study reported the 3D structure of yeast PIC comprising of roughly 32 proteins including the RNA Pol II subunits and the GTFs namely TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH (Murakami et al. 2013). In this model, the structure resembled that of the ribosome, having two distinct lobes – the P-lobe consisting of the RNA Pol II subunits, and the G-lobe formed by the GTFs. The template DNA is sandwiched between these two lobes but is mainly in contact with the G-lobe and not with Pol II. Previous studies of the partial structures of the components of the PIC and sequential assembly had not revealed this separation between the Pol II and the GTFs and indicated a direct DNA-Pol II interaction. However, this study proposed that the GTFs position the DNA above the Pol II cleft and its interaction with Pol II can only occur after DNA melting by TFIIH resulting in the bending of the DNA thus enabling its entry into the Pol II cleft. This mechanism could prevent premature Pol II-DNA interaction (Malik and Roeder 2013; Murakami et al. 2013). The structure was further refined with additional details in subsequent studies (Murakami et al. 2015). However, several other studies point towards the original classical model of core RNA Pol II PIC, where a significant extent of conservation has been found between mechanisms operating in yeast and human (He et al. 2013; Muhlbacher et al. 2014).

#### *1.2.2.1. Transcription initiation in the context of chromatin*

As discussed briefly in Section 1.2.1, chromatin at and near the gene promoters have special features since, with a regular packaging of the DNA, the promoter elements such as TATA-box may not be directly accessible to the TFIID complex for recognition and initiating the PIC assembly. As revealed from genome-wide nucleosome positioning studies, most gene promoters have a specific nucleosome positioning marked by the presence of nucleosome-depleted regions (NDRs) (Yuan et al. 2005; Lee et al. 2007b; Ozsolak et al. 2007; Mavrich et al. 2008; Schones et al. 2008). The two nucleosomes flanking an NDR at the TSS are

designated as –1 (present upstream of the TSS) and +1 (present downstream of the TSS) nucleosomes. These two well-positioned nucleosomes contain the histone variant H2A.Z which confers specific structural properties to the nucleosomes that influence transcriptional initiation at the TSS. The maintenance of the nucleosome positioning and the NDRs is achieved through several mechanisms such as presence of poly(dA:dT) tracts which inhibit nucleosome formation, CpG hypermethylation, activity of DNA-binding TFs which can recruit transcriptional coactivators, chromatin remodeling, chromatin modifications at specific nucleosomal histone residues at or near the TSS, specific histone variants such as H2A.Z and H3.3 which reduce or modulate the stability of the nucleosomes at or near the TSS (Jin et al. 2009), among others (Radman-Livaja and Rando 2010; Struhl and Segal 2013; Muller and Tora 2014). The chromatin modifications at the TSS are used by the transcription machinery to read and bind to the specifically modified nucleosomal histones. For example, the tandem bromodomains of the TBP-associated factor 1 (TAF1) bind to diacetylated histone H4 tails (Jacobson et al. 2000) which are often found to be associated with transcriptional activation (Dion et al. 2005). Acetylated H4 was also found to bind to Bromodomain Factor 1 (Bdf1), a protein that associates with TFIID (Matangkasombut and Buratowski 2003) and other transcriptional activators resulting in an enhancement of transcription (Vettese-Dadey et al. 1996). Several coactivator and chromatin-modifying complexes contain bromodomains which direct their recruitment to the acetylated histones at the promoters. These complexes can then further open up the promoter region by repositioning or removing the histones. Likewise, the PHD domain of TAF3, another subunit of the TFIID complex, was found to bind to H3K4me3 (Vermeulen et al. 2007) which is often present at the promoters (Zhang et al. 2009), and hence could recruit TFIID to the promoters to initiate PIC formation. Further, the H3K4me3-TAF3 mediated PIC formation could occur independently of or cooperatively with the TATA element in the promoter *in vitro*. Hence, this mechanism has important implications in transcription initiation from a TATA-less promoter *in vivo* (Lauberth et al. 2013). These findings indicate that specific histone modifications on the nucleosomes flanking the NDRs at the promoter regions are capable of recruiting TFIID through direct interactions, thereby leading to the formation and stabilization of the PIC.

### *1.2.3. Promoter clearance and escape, RNA Pol II pausing, and transcription elongation*

The process of transcription initiation following the assembly of the PIC constitutes several phases which were identified from the stalling of the RNA Pol II complex at defined positions of the DNA template in an *in vitro* reconstituted system (Holstege et al. 1997; Pal et al. 2001). In the first step, the open complex is formed due to the melting of the promoter in the –9/−2 region. In the second step, the transcription bubble is extended which allows the synthesis of a four-nucleotide RNA. Until the formation of the first three phosphodiester bonds, the formation of the open complex can be reversed which is associated with the generation of short abortive transcripts, and hence the bubble is required to be maintained by the ATP-dependent DNA helicase activity of TFIIH. When the synthesized RNA is at least 7 nucleotides long, a part of the transcription bubble abruptly closes which is known as bubble collapse, a phenomenon that marks the end of the requirement for the TFIIH helicase and defines RNA Pol II promoter clearance transition (Lin et al. 2005; Pal et al. 2005). The third step is accompanied by a switch from abortive to productive RNA synthesis, and this phase is referred to as promoter clearance (Holstege et al. 1997). The promoter clearance transition is also regulated by TFIIB whose N-terminal zinc ribbon contacts the 'dock' domain of the RNA Pol II near the path of RNA exit thereby interfering with the release of abortive transcripts (Bushnell et al. 2004). The subsequent release of TFIIB from the complex is associated with entry into the elongation phase (Pal et al. 2005; Kostrewa et al. 2009). Conformational and architectural changes to the structure of RNA Pol II occur during promoter clearance and transcription elongation (Gnatt et al. 2001; Westover et al. 2004; Barnes et al. 2015; Bernecky et al. 2016; Farnung et al. 2018). Upon reaching a length of 14 – 15 nucleotides by the nascent RNA, most of the GTFs are shed off. However, TFIIF remains associated with RNA Pol II. The other factors namely TFIIA, TFIID, TFIIE, TFIIH, and the Mediator complex remain attached as a scaffold complex to the promoter DNA that facilitates the efficient assembly of PIC for subsequent rounds of transcription of the gene, which would additionally require only TFIIB and TFIIF (Yudkovsky et al. 2000; Dvir et al. 2001; Dvir 2002). The intermediate structure of transcribing RNA Pol II formed during this process is sometimes termed as 'escape competent' which requires a DNA template of 40 – 50 bp length downstream of the TSS to synthesize at least 15 nucleotide long stretch of RNA (Dvir et al. 1997). If this 40 – 42 bp downstream DNA is absent, the RNA Pol II would be arrested at certain points

(depending on the length and not the sequence of the downstream DNA), before its escape (Wang et al. 2003). The terms "promoter clearance" and "promoter escape" are sometimes used to denote the initial and later stages of the transcription initiation process, where 'escape' could imply a complete disengagement of the RNA Pol II from the promoter. Some studies of the 3D organization of the actively transcribed chromatin have suggested that the RNA Pol II may retain promoter contacts during transcript elongation (Luse 2013).

The escaping RNA Pol II is characterized by a specific PTM in its CTD that is Ser5 phosphorylation of the heptapeptide sequence YSPTSPS which is repeated 52 times in the vertebrate RPB1 subunit of RNA Pol II. This phosphorylation is carried out by TFIIH-associated kinase CDK7 (Buratowski et al. 1989; Chapman et al. 2008; Egloff and Murphy 2008). The Ser5 phosphorylated CTD of the RNA Pol II interacts with the capping enzyme (Ho et al. 1998; Fabrega et al. 2003; Kim et al. 2004a) which adds the 7-methylguanosine cap to the 5'-end of the nascent RNA as it emerges out of the RNA exit channel (Fabrega et al. 2004). The capping enzyme also has regulatory effects on transcription elongation through its interactions with positive or negative regulators of RNA Pol II promoter escape such as DRB sensitivity-inducing factor (DSIF) and negative elongation factor (NELF) respectively (Mandal et al. 2004). The association of the transcribing RNA Pol II with DSIF and NELF causes a transient 'pausing' of the RNA Pol II (Yamaguchi et al. 2013). Subsequently, the action of the positive transcription elongation factor b (P-TEFb) comes into play which is recruited to the paused RNA Pol II as a part of the super elongation complex (SEC) (Luo et al. 2012). The CDK9 subunit of P-TEFb phosphorylates multiple factors in the paused elongation complex such as Ser2 of the CTD of RNA Pol II, DSIF and NELF which results in the dissociation of NELF from the complex and the conversion of DSIF into an elongation stimulating factor (Liu et al. 2015b). The phosphorylation of RNA Pol II CTD Ser2 is often facilitated through the 'priming' of the CTD by the prior phosphorylation of Ser7 mediated by the TFIIH-associated kinase CDK7 (Czudnochowski et al. 2012). Ser2 phosphorylation is also mediated by other kinases besides CDK9 such as CDK12 and CDK13 (Bartkowiak et al. 2010; Bowman and Kelly 2014; Core and Adelman 2019). In any case, RNA Pol II Ser2 phosphorylation mediates its interaction with additional elongation factors and RNA processing enzymes involved in polyadenylation and termination, thereby coupling transcription with post-transcriptional processes (Phatnani and Greenleaf 2006; Barrero and Malik 2013).

The transcription elongation factors facilitate the transcribing RNA Pol II complex to overcome the nucleosomal barrier in the chromatin template (Sims et al. 2004). These factors either mediate the remodeling of the chromatin during transcription elongation or alter the catalytic properties of RNA Pol II. For example the ATP-dependent chromatin remodeling complexes SWI/SNF, Acf1/ISWI, and CHD1 function to displace the nucleosomes during transcription elongation (Sims et al. 2004). The histone chaperone FACT, as the name suggests, has been shown to disrupt the nucleosomes ahead of the RNA Pol II elongating complex and act as acceptors of the displaced histones (Winkler and Luger 2011). Chromatin modifying activities also contribute to this process. For example, the lysine acetyltransferase and coactivator p300 synergizes with the SII (also known as TFIIS) component of a chromatin transcription-enabling activity (CTEA) to strongly enhance transcription elongation through several adjoining nucleosomes (Guermah et al. 2006). This transcription elongation factor TFIIS induces mRNA cleavage by enhancing the intrinsic nuclease activity of RNA Pol II. This helps in the removal of cryptic nascent RNA and misincorporated nucleotides that are generated due to the stalling and backtracking of the transcribing RNA Pol II upon encountering a nucleosomal block. Thus TFIIS assists the RNA Pol II to bypass blocks to transcription elongation (Wind and Reines 2000; Kettenberger et al. 2003).

The activation of RNA Pol II from its paused state appears to be the rate-limiting step of the transcription process and several factors have been found to regulate and contribute to the mediation of this step. However, in some cases for a certain number of genes, regardless of the state of transcription of the gene, genome-wide studies for paused RNA Pol II have revealed its presence at the +50 position (Muse et al. 2007; Zeitlinger et al. 2007; Margaritis and Holstege 2008). The paused RNA Pol II complex is also quite stable (Jonkers et al. 2014). In a kinetic analysis of the different RNA Pol II species, it was found by live-cell imaging that on average about 7% of RNA Pol II is freely diffusing, while 10% is chromatin-bound for 2.4 s during initiation, and 23% is promoter-paused for 42 s. However, initiating and promoter-proximal paused RNA Pol II species are dynamic having high turnover, in contrast to the 23 min that an elongating RNA Pol II resides on the chromatin, suggesting that the continuous release and reinitiation of the promoter-bound Pol II is an important aspect of transcriptional regulation (Price 2018; Steurer et al. 2018). The transcription elongation rate is also not uniform. The movement of RNA Pol II is slow at exons but significantly faster in the other regions of the genes which is influenced positively

by the presence of H3K79me2 and negatively by exon density and CG content within genes (Jonkers et al. 2014). Such findings suggest that the transcription of several genes probably initiate but is then temporarily held due to the RNA Pol II pausing. Upon receiving environmental, developmental or disease cues, transcription elongation is resumed and completed which potentially makes the process more efficient (Nechaev and Adelman 2008; Liu et al. 2015b; Core and Adelman 2019).

### *1.2.4. Role of coactivators and the Mediator complex in chromatin transcription*

The mechanism of transcription was deduced mostly in an *in vitro* reconstituted system where the RNA Pol II together with the GTFs constituting the general transcription machinery was sufficient to carry out basal transcription (Roeder 1996; Roeder 1998). However, in the cell, the process of transcription is more regulated by transcription factors (TFs) which are signal-dependent. This regulation could be brought about through direct interactions with the general transcription machinery or through indirect long-range interactions mediated by intermediary factors broadly known as coactivators (Roeder 1998). The coactivators can have other functions besides being a bridging platform between the general transcription machinery and the distantly located TFs, such as chromatin remodeling activity (for example, SWI/SNF, ISWI, CHD, INO80), chromatin-modifying activity (for example, Set1/MLL, CARM1, PRMT1, Set2, KDM5/Lid, SAGA, CBP/p300, Bre), and so on. Such functions are dictated by the presence of specific domains in their structure or subunits in their complex architecture (Krasnov et al. 2016). This brings about finer regulation in the transcriptional outcome which is often gene-specific and/or context-dependent (Naar et al. 2001). The Mediator which is a large multi-subunit coactivator complex with a modular organization is generally required for transcription by RNA Pol II and regulation of various steps of the transcription process. It functions mainly by transducing signals from the transcriptional activators bound to distant enhancer regions to the transcription machinery assembled at promoters in the form of the PIC, thereby modulating transcription initiation (Malik and Roeder 2010; Soutourina 2018). The Mediator comprises about 25 subunits in yeast and 30 subunits in humans that are organized in 3D in three main modules namely the head, middle, and tail, and a separable four-subunit kinase module (Verger et al. 2019). Accordingly, the molecular mass of this huge complex is about 0.8 – 0.9 MDa in yeast and about 1.4 MDa in humans. The subunits of this complex

have been assigned different numbers with the prefix 'MED'. Most of these subunits are conserved across eukaryotic species, with five subunits namely MED23, MED25, MED26, MED28, and MED30 apparently being metazoan-specific (Harper and Taatjes 2018). The head module of the Mediator is composed of seven highly conserved subunits namely MED6, MED8, MED11, MED17, MED18, MED20, and MED22, which together with the middle module (comprising up to nine subunits namely MED1, MED4, MED7, MED9, MED10, MED19, MED21, and MED31 as well as MED26 in mammals), plays a critical role during the PIC assembly by stabilizing the interactions of RNA Pol II with the GTFs. The head and the middle modules together constitute the core Mediator for their indispensable role in the Mediator function. The tail module subunits are the most evolutionarily divergent and include MED2, MED3, MED5, MED15, and MED16 in yeast. The subunit architecture and its conservation across various species have not been strictly elucidated. The structure of the tail module is highly dynamic which potentially serves different functions such as acting as an architectural backbone of the Mediator complex, modulating the global structure of the complex and promoting its stable association with the transcription machinery, as well as exerting an inhibitory effect on the head and/or middle modules in order to prevent promiscuous transcription in absence of activators. The four-subunit (CDK8, CycC, MED12, and MED13) dissociable CDK8 kinase module (CKM) has transcription regulatory functions which can be either repressive or activating, often determined by the CKM-induced structural rearrangements that can block the Mediator-Pol II interaction or vice versa (Verger et al. 2019). These functional propositions have been made through various biochemical and genetic analyses and insights from the structural deductions of partial constituents of the Mediator complex (Harper and Taatjes 2018) as well as the complete complex in association with RNA Pol II and PIC (Robinson et al. 2016).

Owing to the large and complex nature of the Mediator structure, it has obviously been found to interact, directly or indirectly, with a host of cellular proteins such as transcription factors, RNA Pol II subunits, coactivators, super elongation complex (SEC), structural proteins, nuclear receptors, as well as non-coding RNA, in various cell types, thereby regulating processes such as transcription initiation, re-initiation, promoter-proximal pause, elongation, termination, as well as mRNA processing, non-coding RNA activation, DNA looping, super-enhancer formation, higher-order chromatin organization, and genome stability (Fondell et al. 1996; Boyer et al. 1999; Blazek et al. 2005; Yin and Wang 2014;

Allen and Taatjes 2015; Jeronimo and Robert 2017; Sierecki 2018; Maji et al. 2019; Quevedo et al. 2019). As indicated previously, several of these interactions take place in response to specific signals and hence are implicated in specific cellular pathways in various developmental and disease contexts (Ito et al. 2000; Yin and Wang 2014; Allen and Taatjes 2015).

### *1.2.5. Termination of RNA Pol II-driven transcription*

Termination is the final step of the transcription process and has been the least studied among all the other steps of RNA Pol II-mediated transcription, hence its exact mechanism is still poorly defined. Nonetheless, this step is very important as it serves various purposes in the cell such as prevention of RNA Polymerase interference with adjoining DNA elements, recycling the RNA Polymerase, enabling RNA 3'-end processing, and contextual regulation of gene expression through premature termination of transcription referred to as attenuation (Kuehner et al. 2011). As has been realized through research over the years that eukaryotic transcription in the cell is pervasive with much of the genome having the potential of getting transcribed thereby generating a huge amount of premature or non-coding RNA, termination pathways have emerged as an important determining factor of whether the transcripts should be cleaned up by degradation (such as for cryptic unstable transcripts), directed towards maturation or limited processing (such as for snoRNAs), or engaged to yield spliced and polyadenylated mRNAs (Guenther et al. 2007; Loya and Reines 2016). RNA Pol II follows different pathways of termination for protein-coding mRNAs and non-coding RNAs which are the poly(A)-dependent pathway and Sen1-dependent pathway (Arndt and Reines 2015) respectively. The mechanisms and models of the transcription termination process are not further elaborated in this section as these are beyond the scope of the present study.

RNA Pol II transcription termination is coupled to 3'-end processing of the pre-mRNA (Birse et al. 1998; Hirose and Manley 2000; Yonaha and Proudfoot 2000; Buratowski 2005; Rosonina et al. 2006). Several components of the pre-mRNA processing machinery such as CPSF (cleavage and polyadenylation specificity factor) and CSTF (cleavage stimulation factor), are associated with the RNA Pol II elongation complex and come into action during termination (Richard and Manley 2009). This process is also suggested to be linked with initiating and paused RNA Pol II (Proudfoot 2004; Kazerouninia et al. 2010; Mapendano

et al. 2010; Fusby et al. 2016). The RNA Pol II CTD and its phosphorylation status play an important role in transcription termination and 3'-end processing by mediating interactions with different 3'-end processing factors (Komarnitsky et al. 2000; Licatalosi et al. 2002; Proudfoot et al. 2002; Gudipati et al. 2008; Richard and Manley 2009; Hsin and Manley 2012).

The processes of transcription elongation, termination and the coupled 3'-end processing are being further studied in different contexts of disease and development, one of them being the transcription of the variable regions of the immunoglobulin genes where pervasive transcription and premature termination have been proposed to have functional roles in the proper class switch recombination and somatic hypermutation in B lymphocytes.

### *1.2.6. Negative regulation of RNA Pol II transcription*

Uncontrolled or constitutive activation of transcription would be a highly energy-draining process that could significantly harm the cellular transcriptional homeostasis. Hence, to check and balance this effect, several modes of negative regulation of RNA Pol II transcription have developed in the cell. Several genes are known to be actively repressed by transcription factors and various aspects of the chromatin (such as the nucleosomes, chromatin histone modifications, repression through linker histone H1, non-histone chromatin-associated and chromatin-condensing proteins), which are then activated by a specific signal through the mechanism of anti-repression (Croston et al. 1992; Roeder 1998).

Some proteins have a general repressive effect on gene transcription. Such proteins include the histone deacetylases (HDACs) and co-repressors which when recruited to specific gene loci, modify the chromatin in a way that prevents the binding of GTFs, positive coactivators or Pol II itself. For example, the recruitment of HDAC2 and the corepressor mSin3A at the prolactin gene promoter in response to dopamine signal was found to cause rapid deacetylation of histones at the promoter resulting in its repression (Liu et al. 2005). Repression can also take place by directly targeting the components of the PIC. For example, the negative cofactor 2 (NC2) interacts with TBP and sterically inhibits TFIIA and TFIIB from entering the PIC (Kamada et al. 2001). Mot1, which is a global repressor of RNA Pol II transcription, displaces TBP from the promoter DNA through its ATPase

activity (Auble et al. 1994). Moreover, besides transcription activation, the Mediator has been shown to have negative regulatory functions. For example, the MED12 subunit of the Mediator complex was shown to link the RE1 silencing transcription factor (REST, also known as neuron restrictive silencer factor, NRSF) with G9a-mediated histone H3K9me2 resulting in a suppression of neuronal genes in non-neuronal cells (Ooi and Wood 2007; Ding et al. 2008). Thus, through the complexity of its composition, the Mediator can affect the phosphorylation of cyclin H subunit of TFIIH (Akoulitchev et al. 2000), perhaps occlude the binding of RNA Pol II (Naar et al. 2002), affect PIC formation, as well as integrate various signals originating from repressors and other coactivators (Croston et al. 1992; Malik and Roeder 2005).

### *1.2.7. RNA Pol II transcription and histone modifications*

In the context of the general notion that the packaging of the eukaryotic genome into nucleosomes presents a barrier to the RNA Pol II machinery for transcription, specific modifications of the nucleosomal histones have emerged as important factors for the enhancement of Pol II recruitment and thus transcription activation. For example, genome-wide studies have revealed that the histone modification marks H3K4me3, H3K9ac, and H3K14ac are found in the nucleosomes near TSS and strongly correlate with transcription initiation (Bernstein et al. 2002; Santos-Rosa et al. 2002; Schubeler et al. 2004; Guenther et al. 2007). H3K4me2 is found in the nucleosomes present throughout the coding regions (Bernstein et al. 2002) of both active and inactive euchromatic genes (Santos-Rosa et al. 2002). H3K36 methylations such as H3K36me2 and H3K36me3 are also found on nucleosomes present in the gene bodies and 3'-end regions of active genes and correlates with transcription elongation, termination, and/or early RNA processing (Bannister et al. 2005; Kharchenko et al. 2011). Further, H3K79 methylations are found within the bodies of active genes and implicated in transcription elongation by RNA Pol II (Wood et al. 2018). Histone acetylation is generally associated with transcription activation. However, some exceptions to this rule are also present such as H4K20ac which is associated with gene repression in human cells (Kaimori et al. 2016). H3K9ac is distributed in the promoter regions of genes and has been recently shown to mediate the transition from RNA Pol II initiation to elongation by recruiting the SEC and promoting Pol II pause release (Gates et al. 2017). H3K27ac is also a well-known modification that is enriched at the promoter and

enhancer regions of the chromatin and correlates very well with active transcription (Creyghton et al. 2010). Among repressive marks, H3K9me3 and H3K27me3 are the most well-characterized which generally mark heterochromatin regions of the genome having very low transcriptional activity (Richards and Elgin 2002; Wiles and Selker 2017).

It is not well understood if the histone modifications are the cause or the consequence of transcription since there are reasons to believe that the answer to this question might not be straight-forward and would probably be contextual. For example, the Set1 enzyme-containing complex COMPASS which methylates H3K4 in yeast was shown to be recruited by the Paf1 complex associated with the elongating RNA Pol II, which is followed by the methylation of histone H3 (Krogan et al. 2003). The MLL complex which is the human homolog of Set1 exhibited indirect interactions with Ser5 phosphorylated CTD of RNA Pol II implying post-recruitment methylations of H3K4 in the promoter chromatin regions (Hughes et al. 2004). However, another study has shown that the TFIID complex is recruited at some promoters through the binding of its component TAF3 with H3K4me3, an interaction that was found to be selectively inhibited by asymmetric dimethylation of H3R2 but promoted by H3K9ac and H3K14ac (Vermeulen et al. 2007). The H3K4me3-TAF3 interaction could facilitate global TFIID recruitment at active genes and stimulate PIC formation independently or in cooperation with the TATA element, to regulate selective p53 target gene expression in response to genotoxic stress (Lauberth et al. 2013). This suggests that H3K4 methylation could be a consequence as well as a cause of Pol II recruitment and transcription under specific circumstances. The Set2-mediated methylation of H3K36 in the coding regions of actively transcribed genes appears to occur post-recruitment of elongating RNA Pol II which is then recognized by the Eaf3 subunit of the histone deacetylase Rpd3S. Rpd3S deacetylates the acetylated histones which were deposited in the gene bodies by co-transcriptional histone exchange (Venkatesh et al. 2012) thereby preventing cryptic transcription from within genes (Lee and Shilatifard 2007).

The specific histone modifications discussed above are just a few well-studied examples among the plethora of modifications whose number and types keep increasing as active research in this field continues (Bannister and Kouzarides 2011; Zhao and Garcia 2015; Lawrence et al. 2016). Several histone acylations have been discovered whose roles are being studied in different cellular processes including transcription under various developmental and disease contexts (Barnes et al. 2019).

## 1.3. Histone Chaperones

Histone chaperones were originally defined by their property of preventing the non-specific charge-based histone-DNA interactions (that form insoluble aggregates at physiological ionic strength (Stein 1979)), and in the process facilitating ordered nucleosome assembly (Laskey et al. 1978). However, accumulating evidence in this field suggest multiple other roles of histone chaperones beyond just nucleosome assembly which are implicated in different cellular processes including transcription. Fundamentally, these chaperone proteins function by binding to the histones, partially through the acidic stretches in their structures, and systematically depositing them onto the DNA to form the nucleosome, without being a part of the final product (De Koning et al. 2007). They help in shielding the positive charge of the histones from non-specifically interacting with the negatively charged DNA. Histone chaperones also mediate the reverse process that is the removal or eviction of the histones from the nucleosome known as nucleosome disassembly (Akey and Luger 2003) (Figure 1.5). Further, histone chaperones carry out other processes such as transfer of the histones from one chaperone to another, and transfer of histones to enzymes that use them as substrates (De Koning et al. 2007). Nucleosome assembly is initiated with the deposition of the H3-H4 tetramer that has a higher affinity for DNA, followed by the sequential binding of two dimers of H2A and H2B, which have a high affinity for H3-H4 bound to DNA (Annunziato 2013; Elsasser and D'Arcy 2013). During chromatin disassembly, histone chaperones act as ‘acceptors’ for histones and shield them until they have been assembled (Das et al. 2010b) (Figure 1.5).

### 1.3.1. Types and Functions of Histone Chaperones

Different histone chaperones have been discovered having a strict or preferential binding affinity towards one or more histones and/or histone variants (Table 1.1). Accordingly, they have been implicated in different cellular processes as a consequence of this preference. For example, the histone chaperones Nap1 and Nucleoplasmin have a binding preference towards H2A-H2B, while others such as Asf1, NASP (N1/N2), CAF-1, HIRA, Vps75, SET, RbAp46, and RbAp48 preferentially bind to H3-H4 (Eitoku et al. 2008). Some histone chaperones, however, can bind to both H2A-H2B and H3-H4 such as FACT, and Nucleophosmin (NPM1) (Swaminathan et al. 2005; Bowman et al. 2011; Formosa 2012).

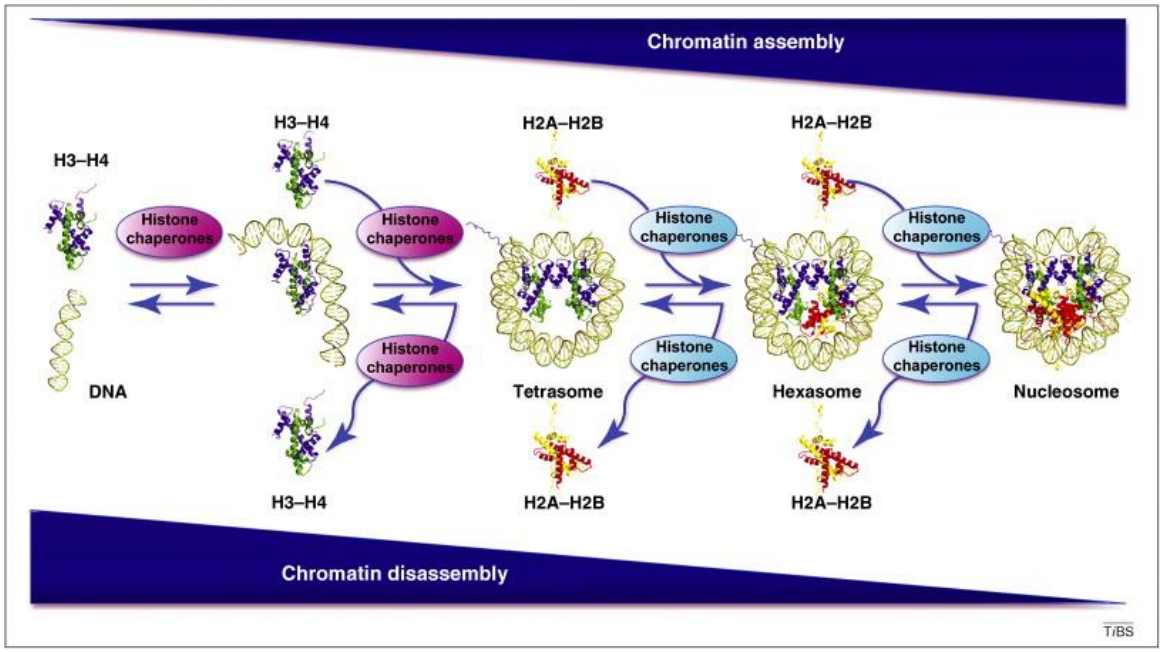



**Figure 1.5. Role of histone chaperones in the stepwise assembly and disassembly of nucleosomes:** DNA is wrapped around two H3-H4 dimers forming the more stable [$(H3\text{-}H4)_2$] tetramer, and two H2A-H2B dimers to form the nucleosome core particle. This ordered and stepwise assembly occurs through different possible intermediates such as the tetrasome and hexasome. Each step of the assembly/disassembly process is mediated by histone chaperones. Generally, histone chaperones have a binding preference towards H3-H4 or H2A-H2B. Some chaperones can, however, bind to both (not depicted here). Histone H2A is shown in yellow, H2B in red, H3 in blue, and H4 in green. The figure has been obtained from (Das et al. 2010b) and reused after obtaining copyright permission from Elsevier under license number 4718600857602.

Apart from the classical function of nucleosome assembly and disassembly, histone chaperones play roles in other processes as well by virtue of their histone-binding property. They may help in the nuclear import of histones after their synthesis in the cytoplasm. They may also participate in histone storage, supply, recycling, and exchange. Some histone chaperones perform other specialized functions such as modulation of histone PTMs (Avvakumov et al. 2011; Hondele and Ladurner 2011). Based on the time or the cell cycle phase where specific roles of histone chaperones come into play, these functions of the histone chaperones can be broadly categorized as replication-coupled and replication uncoupled functions (Burgess and Zhang 2013). In the context of chromatin, the actions of histone chaperones play important roles in modulating the chromatin dynamics in different regions such as genic regions, centromeres, telomeres, which impact specific processes such as replication, transcription, repair. This is eventually manifested at the organismal level where proper functioning of histone chaperones in regulating the histone traffic results in normal development and sustenance of the organism while abnormal functioning of these

proteins could lead to different pathological conditions including cancer (Gurard-Levin et al. 2014) (Figure 1.6). The general functions of histone chaperones are broadly depicted in Figure 1.7 while Table 1.1 lists the histone specificity and functions of well-studied histone chaperones. A discussion of the roles played by specific histone chaperones in eukaryotic DNA replication and repair is beyond the scope of the present study and hence is not elaborated further. The following section deals with the roles of histone chaperones in transcription regulation.

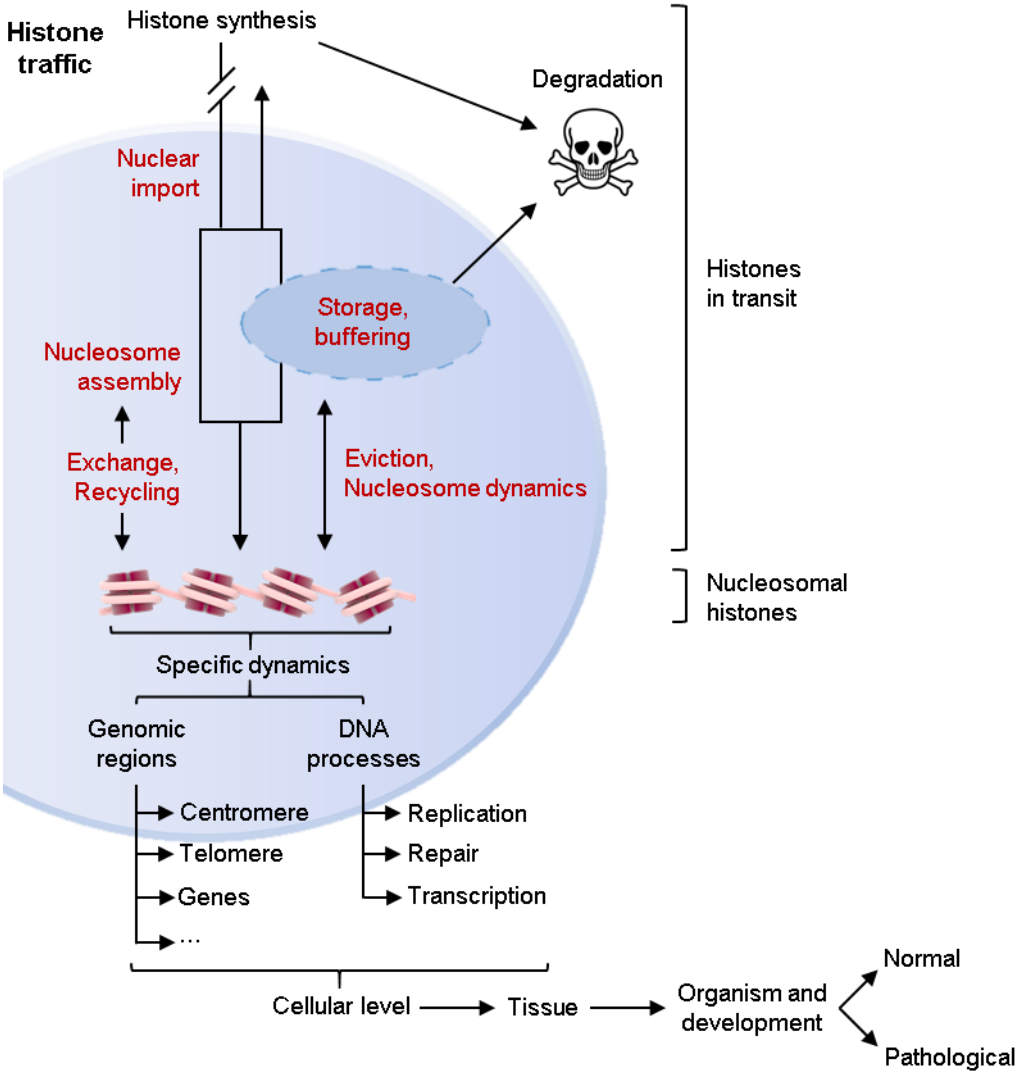



**Figure 1.6. Role of histone chaperones in various processes of histone traffic:** The histone chaperones carry the histones during their cellular life and shuttle them into various pathways for their nuclear import, storage, degradation, and dynamics at the chromatin such as assembly, exchange, and eviction. Together, these processes contribute to different DNA templated processes such as transcription, replication, and repair, whose effects are manifested in normal physiological or pathological conditions at the organismal level. Figure adapted from (Gurard-Levin et al. 2014).

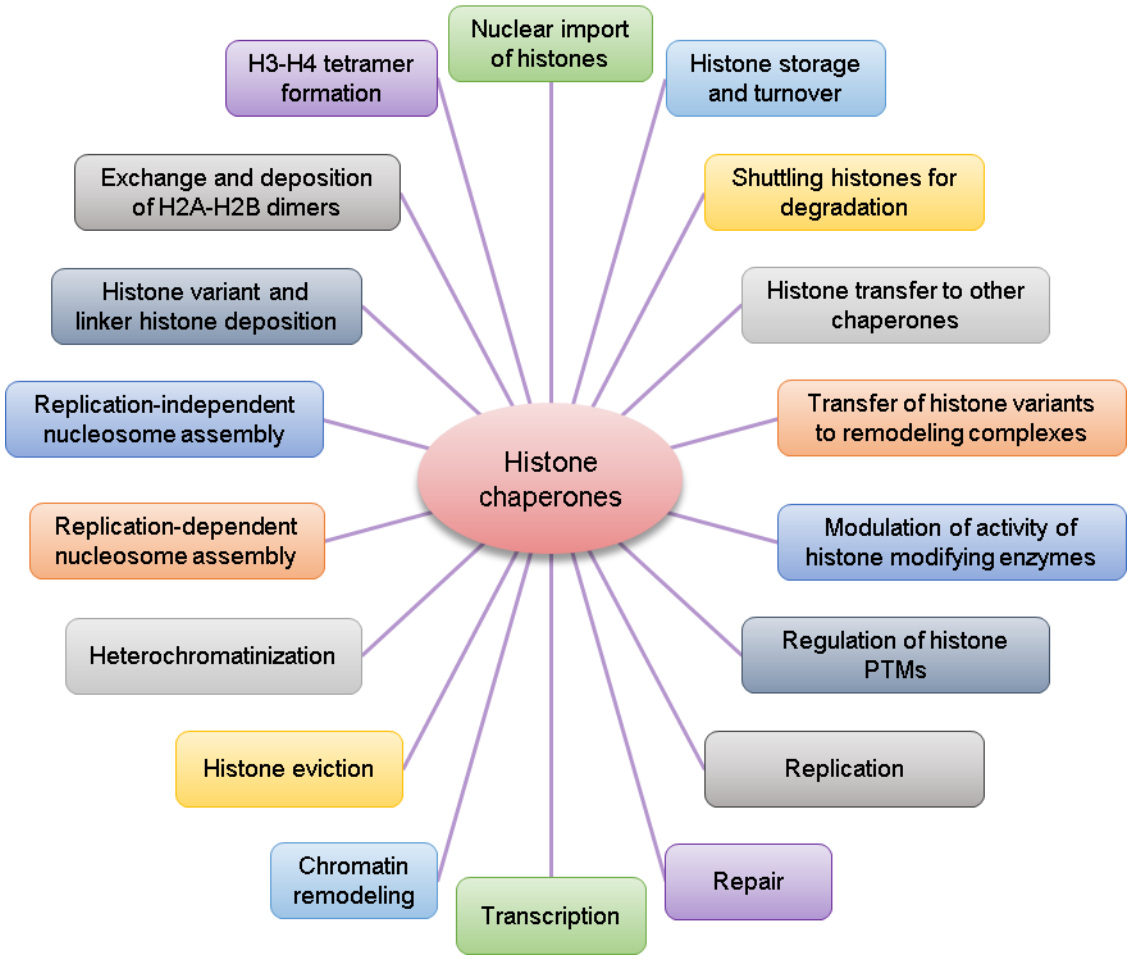



**Figure 1.7. Various functions of histone chaperones.** Figure adapted from Senapati P, Ph.D. thesis, 2014.

| Histone chaperone | Histone cargo | Function |
|---|---|---|
| **Asf1** | H3.1-H4, H3.2-H4, H3.3-H4 | Histone import; histone transfer to CAF-1 and HIRA; regulation of H3K56ac, H3K9ac, H4K5ac, H4K12ac, H3K36me3, H3K4me3, parental histone PTMs during replication; histone removal during transcription initiation and histone removal as well as deposition during transcription elongation; histone sink during replicational stress; repair. |
| **CAF-1 complex (having p150, p60, RbAp48)** | H3-H4, H3.1-H4 | H3.1-H4 deposition coupled to replication and repair; $(H3–H4)_2$ formation; regulation of H3K56ac, H4K5/12ac in import. |
| **Daxx (with ATRX)** | H3.3-H4 | Replication-independent H3.3-H4 deposition at telomeric heterochromatin; maintenance of ribosomal DNA and pericentric heterochromatin. |
| **DEK** | H3.3-H4 | Regulation of H3.3-H4 incorporation; maintenance of heterochromatin; transcriptional coactivator. |

| | | |
|---|---|---|
| **HIRA complex (having HIRA, Cabin1, UBN1)** | H3.3-H4 | Replication-independent deposition of H3.3-H4 in genic regions; regulation of H3ac. |
| **NASP or N1/N2** | H3-H4, linker histones | Storage of H3-H4 in *Xenopus laevis* oocytes; histone supply and turnover; linker histone chaperone; protects H3-H4 from degradation in human cells. |
| **Rtt106** | H3-H4 | Formation and deposition of $(H3\text{-}H4)_2$ tetramer; heterochromatin silencing; regulation of H3K56ac. |
| **HJURP** | CENPA-H4 | Regulation of incorporation of the H3 variant CENPA; centromere maintenance. |
| **FACT complex (having Spt16 and SSRP1)** | H3-H4, H2A-H2B, H2A.X-H2B | Deposition and exchange of H3-H4, H2A-H2B, H2A.X-H2B; Transcription elongation; assists in chromatin remodeling; replication; repair; regulation of H2BK123ub; regulation of H2A.XS139(129)ph; contribution to the formation of cohesin-dependent TADs, in the context of nuclear organization during interphase and mitotic chromosome folding. |
| **Nap1** | H3-H4, H2A-H2B, H2A.Z-H2B, H1 | H2A-H2B nuclear import and deposition during replication and transcription; nuclear import of H2A.Z-H2B heterodimers; linker histone chaperone; regulation of H3ac, H3K9ac, H3K4me3. |
| **Chz1** | H2A.Z-H2B | H2A.Z-H2B deposition; transfer of H2A.Z-H2B to SWR1 remodeling complex for assembly into chromatin; transcription; regulation of H2BK123ub. |
| **APLF** | Core histones, macroH2A.1-H2B | Regulation of macroH2A.1 incorporation during DNA damage. |
| **Hif1** | H3-H4 | Component of Hat1 complex; assists HAT enzyme. |
| **Rsf1** | H3-H4 | Assists remodeling complex. |
| **Spt6** | H3-H4 | Transcription initiation and elongation; regulation of H3K36me2/3. |
| **FKBP** | H3-H4 | Ribosomal DNA silencing. |
| **Nucleolin** | H2A-H2B, macroH2A-H2B | Transcription elongation; assists in chromatin remodeling. |
| **Nucleoplasmin** | H2A-H2B | Storage of H2A-H2B in *Xenopus laevis* oocytes; sperm chromatin remodeling; cytoplasmic-nuclear transport. |
| **Nucleophosmin (NPM1)** | H3-H4, H2A-H2B, H1 | Transcriptional regulation; other functions separately covered in Section 1.4.1.3. |
| **JDP2** | H2A-H2B, H3-H4 | Regulates transcription via inhibition of p300 mediated histone acetylation and chromatin assembly. |
| **ANP32E** | H2A.Z-H2B | H2A.Z eviction; transcription. |
| **Swc2** | H2A.Z-H2B | H2A.Z incorporation. |

| | | |
|---|---|---|
| **Arp4** | Not determined | Assists remodeling complex. |
| **Arp7, Arp9** | Not determined | Assists remodeling complex. |
| **Arp8** | H3-H4 | Assists remodeling complex. |
| **Acf1** | H3-H4, H2A-H2B | Assists remodeling complex. |
| **Spt2** | H3-H4 | Transcription elongation. |
| **RbAp46** | H3-H4 | Histone transport; assists HAT enzyme. |
| **Vps75/SET** | H3-H4 | Chromatin assembly and disassembly at both active and inactive genes; regulation of H3K9ac, H3K23/27ac. |
| **MCM2** | CENPA-H4, H3.1-H4, H3.2-H4, H3.3-H4 | Replication; nucleosome position memory. |
| **TONSL** | H3-H4 | Replication; recognizes unmethylated H4K20 mark of newly replicated DNA. |
| **HSP90A/B** | H3-H4, H2A-H2B, H1 | Collaborates with NASP to assemble H3-H4 dimer; mediates histone degradation. |
| **HSC70** | H3-H4, H2A-H2B, H1 | Collaborates with NASP to assemble H3-H4 dimer; mediates histone degradation. |
| **IPO4** | H3.1-H4, H3.2-H4, H3.3-H4 | Nuclear import of monomeric histones from the cytosol and transferring them to NASP in the nucleus before their heterodimerization. |

**Table 1.1. Histone chaperones and their functions:** Information in the table has been compiled from (Avvakumov et al. 2011; Burgess and Zhang 2013; Gurard-Levin et al. 2014; Venkatesh and Workman 2015; Hammond et al. 2017; Reddy et al. 2017) with references therein, and (Jin et al. 2006; Campos et al. 2010; Straube et al. 2010; Xue et al. 2013; Huang et al. 2015; Saredi et al. 2016; Apta-Smith et al. 2018; Garcia-Luis et al. 2019; Schlissel and Rine 2019).

### *1.3.2. Histone chaperones and the regulation of RNA Pol II transcription*

By virtue of the ability to bind to histones in a regulated manner, histone chaperones have been found to have significant involvement in assisting the transcription machinery during chromatin transcription. They contribute to this process in several ways which fundamentally can be narrowed down to histone removal ahead of transcription, histone reassembly following transcription, exchange of histone variants and regulation of histone modifications that modulate the chromatin structure making it amenable to transcription (Figure 1.8). Besides these, some histone chaperones can collaborate with other transcription regulatory proteins such as chromatin remodelers, histone modifiers, and

coactivators (Venkatesh and Workman 2015). The following paragraphs give an overview of the well-studied transcription regulatory mechanisms of some histone chaperones.

Several histone chaperones have been found to play a role in chromatin remodeling during transcription initiation and elongation to increase the rate of transcription (De Koning et al. 2007; Kulaeva et al. 2007). The histone chaperone Asf1, which was originally identified as a transcriptional derepressor in a yeast genetic screen (Le et al. 1997), has been observed to associate with gene promoters and evict histones or enhance nucleosome disassembly during transcription initiation (Adkins et al. 2004; Korber et al. 2006). Asf1 has also been demonstrated to mediate histone H3 eviction as well as deposition during transcription elongation by RNA Pol II where its actions could inhibit transcription initiation from cryptic promoters within gene bodies (Schwabish and Struhl 2006).

Some histone chaperones deposit histone variants that have been associated with transcription activation, at the promoter regions. For example, the H2A variant H2A.Z has been observed in the −1 and +1 nucleosomes flanking the nucleosome-depleted region (NDR) near the TSS. The enhanced acetylation status of H2A.Z prevents the formation of a repressive chromatin structure (Draker and Cheung 2009). The H3 variant H3.3 is found to be enriched at actively transcribed regions of the chromatin and hence associated with transcription activation (Ahmad and Henikoff 2002; McKittrick et al. 2004). In this context of transcriptional activation due to the presence of H2A.Z and H3.3 at the promoters (Jin et al. 2009; Thakar et al. 2009), the actions of their specific chaperones become relevant. The H3.3-specific chaperone HIRA deposits H3.3 at the transcriptionally active regions (Szenker et al. 2011). HIRA also interacts with RNA Pol II complex itself as shown through various studies (Osborn and Greer 2015), possibly helping in its recruitment to enhance transcription since depletion of HIRA in yeast resulted in impaired Pol II recruitment and nucleosome eviction at the gene promoters (Chujo et al. 2012). The H2A.Z-specific histone chaperone and remodeling complex Swr1 deposits the H2A.Z-H2B heterodimer in the chromatin through the ATP-driven exchange of the variant H2A.Z (Mizuguchi et al. 2004). The chaperones Nap1 and Chz1 help in the transfer of H2A.Z-H2B heterodimers to Swr1 in exchange for canonical H2A-H2B. Nap1 is involved in the nuclear import of H2A.Z-H2B while Chz1 is involved in transferring the heterodimer to Swr1 in the nucleus (Straube et al. 2010). Another H2A.Z-specific chaperone, ANP32E, is involved in the resolution of the non-nucleosomal H2A.Z aggregates possibly facilitating the removal of H2A.Z at the +1 nucleosomes, which could help RNA Pol II to overcome the first nucleosomal barrier

(Mao et al. 2014). Thus, the exchange of these histone variants such as H2A.Z at the specific loci in the chromatin is highly regulated and brought about by specific histone chaperones and chromatin remodeling complexes which facilitates their easy eviction during transcription initiation (Billon and Cote 2013).

The histone chaperone FACT has been best described for its role in transcription (Orphanides et al. 1998). This complex consists of two subunits Spt16 which binds to H2A-H2B, and SSRP1 that can bind to H3-H4. FACT has been shown to evict H2A-H2B dimers thereby promoting nucleosome disassembly during transcription (Orphanides et al. 1999). The FACT-mediated removal of H2A-H2B dimers resulting in nucleosome disruption is facilitated by a conserved HBR (H2B repression) domain in the H2B tail (Zheng et al. 2014). Regarding the mechanism of FACT-mediated nucleosome disassembly, two models have been proposed. According to the first model known as the 'dimer displacement model', the Spt16 subunit of FACT which can bind to both H2A-H2B dimers and nucleosomes, first removes an H2A-H2B dimer. This stimulates the interaction of the SSRP1 subunit of FACT with the nucleosome with an altered structure (or a hexasome). Thus FACT promotes the destabilization of the dimer-tetramer interactions during transcription (Belotserkovskaya et al. 2003). The second model known as the 'global accessibility' model proposes that the actions of FACT do not displace the H2A-H2B dimer of the nucleosome first, but instead reorganize it to form a looser, more dynamic structure with the original composition (Xin et al. 2009). The evidence presented to support this model shows that FACT increases the sensitivity of the reorganized nucleosome to hydroxyl radicals and endonucleases globally and not just in regions contacted by H2A-H2B. The looser nucleosome thus formed would be more prone to the displacement of H2A-H2B dimers which is in accordance with the 'dimer displacement model' but with a different mechanistic cause (Formosa 2012). FACT also mediates the reassembly of the nucleosomes transversed by the elongating RNA Pol II (Formosa 2012). This property helps prevent cryptic transcription from within the gene bodies and maintain the fidelity of transcription (Mason and Struhl 2003). Transcription elongation is also facilitated by FACT through another mechanism where FACT promotes RNF20/40 and UbcH6-mediated histone H2B monoubiquitination which is a mark associated with transcriptional activity. H2B monoubiquitination further enhances FACT activity on the eviction of H2A-H2B dimers from the nucleosomes, thereby stimulating transcript elongation with the generation of longer transcripts (Pavri et al. 2006).

Some histone chaperones mediate transcription activation indirectly by enhancing the activities of other chromatin remodelers implicated in transcription. For example, the histone chaperone Nucleolin has been shown to destabilize the histone octamer as well as stimulate the chromatin remodeling activities of the SWI/SNF and ACF complexes. Nucleolin can thus facilitate transcription through the nucleosome similar to the FACT complex (Angelov et al. 2006). Nucleolin can also assist in gene transcription such as that of the rDNA genes where it helps in the removal of macroH2A from the nucleosomes present at the promoters of methylated rDNA genes (Cong et al. 2014).

Some histone chaperones can regulate activator-dependent transcription by interacting with specific activators. For example, the histone chaperone NAP1 can interact with the HIV-1 Tat protein and enhance Tat-mediated activation of viral gene expression. This effect was further increased in the presence of p300, which is a known coactivator for both Tat and NAP1 (Vardabasso et al. 2008). Nap1 can also mediate the reassembly of nucleosomes during transcription elongation (Del Rosario and Pemberton 2008).

Sometimes the collaborative or opposing roles of multiple histone chaperones determine their effect on transcription activation. For example, both of the H3-H4 chaperones Asf1 and HIRA can deposit histones but they have different effects on histone exchange. Asf1 could mediate the incorporation of external H3-H4 and renewal of the pre-existing histones, which was opposed by HIRA. Hence, a balance of these two opposing activities might be the determining factor for the state of the chromatin in the context of transcription (Kim et al. 2007).

Finally, as mentioned previously, several histone chaperones help in the regulation of histone PTMs associated with transcriptional activation (Avvakumov et al. 2011). For example, Rtt109-mediated H3K56ac on soluble histones is stimulated by the histone chaperone Asf1 (Adkins et al. 2007; Tsubota et al. 2007) and is implicated in different DNA templated processes including transcription (Williams et al. 2008). The FKBP family protein Fpr4 is a histone chaperone that negatively regulates Set2-mediated H3K36 methylation through isomerization of Pro38 of histone H3, thereby affecting transcription induction kinetics (Nelson et al. 2006). Further, a positive feedback loop takes place between FACT and monoubiquitination of H2BK123, which facilitates the timely assembly and disassembly of nucleosomes during transcription as well as prevent cryptic transcription (Pavri et al. 2006; Fleming et al. 2008). Histone chaperones like Asf1 and

Nap1 can also promote the removal of H3K4me3 and H3ac at promoters and enhancers of certain genes to silence their expression (Moshkin et al. 2009). Spt6 has been shown to enhance the removal of PRC2-mediated H3K27me3 mark by recruiting the demethylase KDM6A (UTX) to the chromatin, thereby leading to activation of gene expression during myogenesis (Wang et al. 2013). On the other hand, HIRA was found to be important for the establishment of PRC2-mediated deposition of H3K27me3 at promoters of developmental genes in mouse embryonic stem cells through the recruitment of PRC2 in coordination with histone H3.3 (Banaszynski et al. 2013).

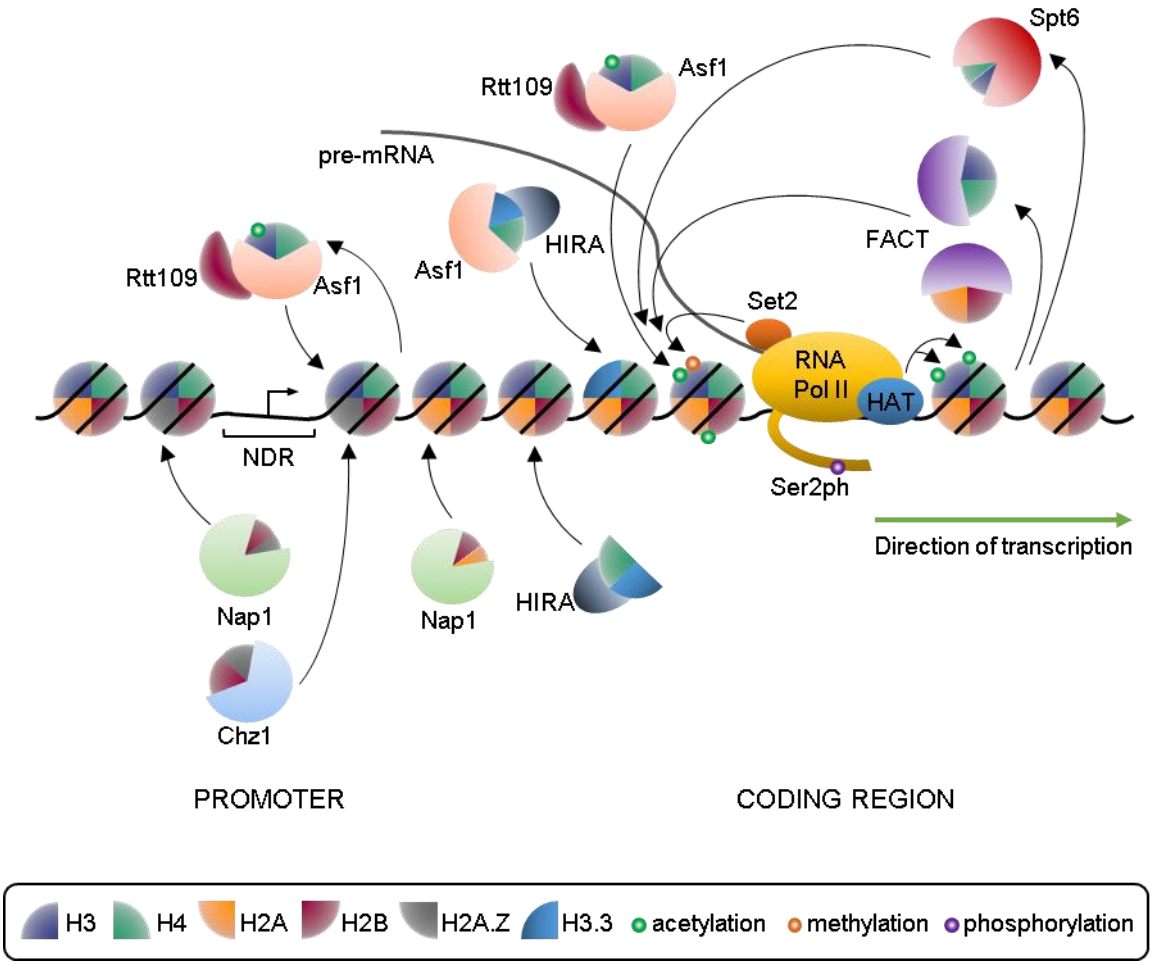



**Figure 1.8. Multiple roles of histone chaperones during RNA Pol II-mediated transcription:** Some of the well-characterized mechanisms of histone chaperone-mediated regulation of RNA Pol II-driven transcription are illustrated. Note that the organization of the core histones in the nucleosome in this illustration is not as per the true structural organization of the nucleosome core particle. In the promoter regions, the transcription start site (TSS, denoted by the arrow) is generally devoid of nucleosomes and hence is known as the nucleosome-depleted region (NDR). The two nucleosomes flanking the NDR generally contains the H2A variant H2A.Z. Histone chaperones Nap1 and Chz1 help in the deposition of H2A.Z-H2B heterodimers in the promoter chromatin through Swr1 chromatin remodeler (not shown here). H2A.Z containing nucleosomes being labile are more prone to removal when transcription initiation takes place. Asf1 is involved in the removal

of histones from the promoters. It also facilitates the acetylation of the H3-H4 dimers displaced from the promoter by its associated histone acetyltransferase (HAT), Rtt109 at the H3K56 site which helps in their removal during the subsequent cycles of transcription initiation. In the coding region, histones displaced by the advancing RNA Pol II are accepted by the histone chaperones Spt6 and FACT, which are then used for reassembly of the nucleosomes with the transcribed DNA in its wake. Asf1, Nap1, and the H3.3-specific chaperone HIRA mediate the deposition of newly synthesized (not shown here) histones. Nucleosome disassembly ahead of the elongating RNA Pol II is facilitated by acetylation of histones by different candidate HAT complexes coming into play during transcription elongation. The Spt6 chaperone also promotes the recruitment of the Set2 methyltransferase that methylates reassembled nucleosomes at the H3K36 site, which is then used by the Rpd3S HDAC complex to remove histone acetylation (not shown here), thereby re-establishing a repressive chromatin structure that prevents cryptic transcription from within the coding region. Asf1 also has the capability to evict H3-H4 tetramer during elongation and deposit them back (not shown here). These steps are often collaborated by chromatin remodelers which are not shown here. The figure has been adapted from (Avvakumov et al. 2011).

Thus, the interplay of different histones chaperones along with other factors regulates the distinct steps of transcription. While histone chaperone-mediated eviction or exchange of histones from the nucleosomes at the promoters make the chromatin template permissive for transcription initiation, disassembly and subsequent reassembly of the nucleosomes in the gene bodies coordinated with the movement of the RNA Pol II machinery facilitates faithful transcription elongation.

### *1.3.3. Histone chaperones in development and disease: An overview*

As described previously, histone chaperones by virtue of their regulated and specific interactions with histones, play crucial roles in various cellular processes of histone metabolism and traffic. Besides the very important process of chromatin organization during DNA replication, repair, and transcription, the mere storage, buffering and nuclear import of the histones itself is critically regulated and coordinated by these chaperones that prevent the development of a cellular stress condition due to the presence of such highly charged proteins free in solution. These molecular functions have important implications in the general health of the cell which is manifested during various developmental stages of the organism while their deregulations are implicated in different diseases including cancer. The role of histone chaperones in development has been further validated from the evidence from various knockout and mutagenesis studies of specific chaperones (Table 1.2). The knockout phenotypes of some of these proteins are embryonic lethal which proves the critical roles played by these chaperones in development. The deposition of histone

variants in specific genomic locations is regulated in a spatial and temporal manner during development by the respective chaperones (Banaszynski et al. 2010; Filipescu et al. 2013). For instance, the histone chaperone HIRA mediates the assembly of H3.3 containing nucleosomes in the decondensing male pronucleus after fertilization in *Drosophila* (Loppin et al. 2005; Bonnefoy et al. 2007). HIRA is also required for the proper deposition of H3.3 and regulating the chromatin dynamics during gastrulation in *Xenopus laevis* (Szenker et al. 2012) as well as artificial reprogramming of *Xenopus* eggs and oocytes (Jullien et al. 2012). Similarly, CAF-1-mediated chromatin assembly is critical for regulating the nuclear organization and cell cycle progression during the rapid early cleaving stages of *Xenopus* development (Quivy et al. 2001). Chromatin dynamics during fertilization are also regulated by the classical histone chaperone Nucleoplasmin which was identified as an important factor mediating sperm chromatin remodeling after fertilization (Philpott et al. 1991). Knockout of mouse Nucleoplasmin or NPM2 phenotypically resulted in sub-fertility in females while the males had no apparent defects. In this case, although sperm chromatin decondensation proceeded without NPM2, there were clear abnormalities related to the nuclear and nucleolar organization in oocytes and early embryonic nuclei, which indicated a critical role of NPM2 in normal embryonic development (Burns et al. 2003). The paralog of NPM2 that is NPM1 is however essential for embryonic development and genomic stability, which can be attributed to other histone chaperone-unrelated functions of this protein as well (elaborated in Section 1.4.1.3). Thus, the absence of specific histone chaperones might result in the improper deposition of histone variants or flaws in chromatin organization which can interfere with the lineage-specific expression profiles that regulate various developmental programs.

| Histone chaperone | Knockout phenotype | Reference |
|---|---|---|
| **HIRA** | Embryonic lethal at E11; abnormal gastrulation; death probably resulting from abnormal placentation and failure of cardiac morphogenesis. | (Roberts et al. 2002) |
| **DAXX** | Embryonic lethal at E9.5; occurrence of extensive apoptosis. | (Michaelson et al. 1999) |
| **ATRX** | Embryonic lethal at E9.5 pc; abnormal trophoblast development. | (Garrick et al. 2006) |
| **Asf1a** | Embryonic lethal at midgestation. | (Hartford et al. 2011) |

| | | |
|---|---|---|
| **CAF-1 p150** | Developmental arrest by 16-cell stage; abnormal heterochromatin (pericentric) organization. | (Houlard et al. 2006) |
| **NPM2 (Nucleoplasmin)** | Sub-fertility in females; improper nuclear and nucleolar organization in oocyte and early embryonic nuclei; null embryos did not cross the 2-cell stage. | (Burns et al. 2003) |
| **NPM1 (Nucleophosmin)** | Embryonic lethal between E11.5 and E16.5; severe anemia resulting from defects in primitive hematopoiesis. | (Grisendi et al. 2005) |

**Table 1.2. Knockout phenotypes of various histone chaperones.** The table has been adapted from (Gurard-Levin et al. 2014).

The importance of histone chaperones in physiological processes is also manifested when mutations in the histone chaperone encoding genes lead to altered or loss of functions in the respective proteins that often result in abnormal growth of the cell or cancer. Table 1.3 enlists the disease phenotypes for mutations of in of the histone chaperones. Such revelations clearly suggest that genome instability and altered gene expression network are the consequences of mutations in factors such as histone chaperones, involved in nucleosome assembly which ultimately promote the development of various diseases (Burgess and Zhang 2013). For example, the mutations in the genes belonging to the Daxx-ATRX-H3.3 deposition pathway have been denoted as 'driver' mutations that promote cancer pathogenesis and manifested in the form of altered telomeres and gene expression profiles (Heaphy et al. 2011; Jiao et al. 2011; Schwartzentruber et al. 2012). A fusion protein of the histone chaperone DEK and nucleoporin 214, DEK-NUP214 (formerly known as DEK-CAN) is found in about 1% of AML patients where it significantly reduces the formation of the functional DEK histone chaperone complex, thereby inhibiting the normal functioning of the protein (Sawatsubashi et al. 2010) as well as promoting cancer through various other mechanisms (Mendes and Fahrenkrog 2019). On the other hand, overexpression of certain histone chaperones such as Asf1b, CAF-1 p60, HJURP Nucleolin, and Nucleophosmin (Table 1.3) can also cause an imbalance in the cellular homeostasis and drive them towards cancer through different mechanisms requiring altered or hyperactivity of these nucleosome assembly factors. However, it can be the other way around as well. Since proteins like Asf1b, CAF-1, NPM1, and Nucleolin themselves play important roles in cell proliferation, increased protein abundance of these factors in cancer cells could be a result of the enhanced proliferation status of cancer cells. Nevertheless, an

elevated level of these chaperones can alter nucleosome assembly, resulting in genome instability and the promotion of tumorigenesis. Hence, while it is difficult to assess the cause or the consequence of cancer, it can be surmised that the role of histone chaperones in the regulation of telomeres, gene expression and heterochromatin organization are important for cell survival and abnormalities in these functions are associated with disease pathogenesis.

| Histone chaperone gene | Disease phenotype due to mutation/overexpression | Reference |
|---|---|---|
| ***HIRA*** | Deletion or haploinsufficiency associated with DiGeorge Syndrome (a congenital developmental disorder characterized by heart defects and poor immune system function). | (Lorain et al. 1996) |
| ***ATRX*** ***DAXX*** | Mutations observed in pancreatic neuroendocrine tumors (PanNETs); mutations observed in pediatric glioblastoma. | (Jiao et al. 2011; Schwartzentruber et al. 2012) |
| ***DEK*** | Mutation (chromosomal translocation forming hDEK-CAN fusion protein) observed in a subset of acute myeloid leukemia (AML); overexpression in glioblastoma, melanoma, bladder carcinoma, cervical cancer. | (Soekarman et al. 1992; von Lindern et al. 1992; Wise-Draper et al. 2009) |
| ***ASF1B*** | Overexpression observed in breast cancer. | (Corpet et al. 2011) |
| ***CHAF1B*** **(CAF-1 p60)** | Overexpression observed in renal, endometrial, breast, cervical cancer. | (Polo et al. 2010) |
| ***HJURP*** | Overexpression observed in breast, lung cancer. | (Kato et al. 2007; Hu et al. 2010) |
| ***NCL*** **(Nucleolin)** | Overexpression observed in multiple cancers. | (Grinstein et al. 2002; Storck et al. 2007) |
| ***NPM1*** **(Nucleophosmin)** | Mutation (chromosomal translocation, deletion) and overexpression observed in blood and other cancers respectively. | (Grisendi et al. 2006) |

**Table 1.3. Disease phenotypes due to mutation/overexpression of histone chaperones.**

## *1.4. Nucleoplasmin Family of Histone Chaperones*

Nucleoplasmin was the first histone chaperone to be described as the factor that prevented histone-DNA aggregations during nucleosome assembly in *Xenopus laevis* oocytes (Laskey et al. 1978). The *Xenopus* Nucleoplasmin (NP) was subsequently found to have orthologs in other vertebrates. The mammalian orthologs of NP, known as NPM2, share about 40 – 50% identity in protein sequence with NP. Two other proteins were found to share similar structural properties with NPM2 and were assigned to this family (Schmidt-Zachmann et al. 1987; Zirwes et al. 1997). These proteins are NPM1, also known as Nucleophosmin/Numatrin/B23, and NPM3. The members of the NPM family differ in their expression patterns and localizations and have been implicated in different functions based on both *in vitro* and *in vivo* evidence which would be discussed further in the subsequent sections. However, all these proteins have a characteristic domain structure wherein the N-terminal core domain (also known as the oligomerization domain) is quite conserved in sequence and helps in homo- and/or hetero-oligomerization of the protein as well as interact with core histones, while the C-terminal region is disordered that stabilizes the interactions of NPM with the histones mainly through its conserved acidic stretches (Figure 1.9 and Figure 1.11A – B). The crystal structures of the core domains of five of the Nucleoplasmin family proteins reveal that this domain forms a beta-barrel like structure with a jelly-roll topology which associate to form pentamers (Dutta et al. 2001; Namboodiri et al. 2003; Namboodiri et al. 2004a; Namboodiri et al. 2004b; Lee et al. 2007a; Platonova et al. 2011) (Figure 1.10). Two pentamers associate in a head-to-head fashion to form a decamer which is believed to be the functional form of these chaperones that bind to the histone proteins (Figure 1.11B). The oligomer form of NPM is resistant to reducing agents as they do not associate through disulfide bonds but instead are assembled mainly by hydrophobic interactions within the core domain.

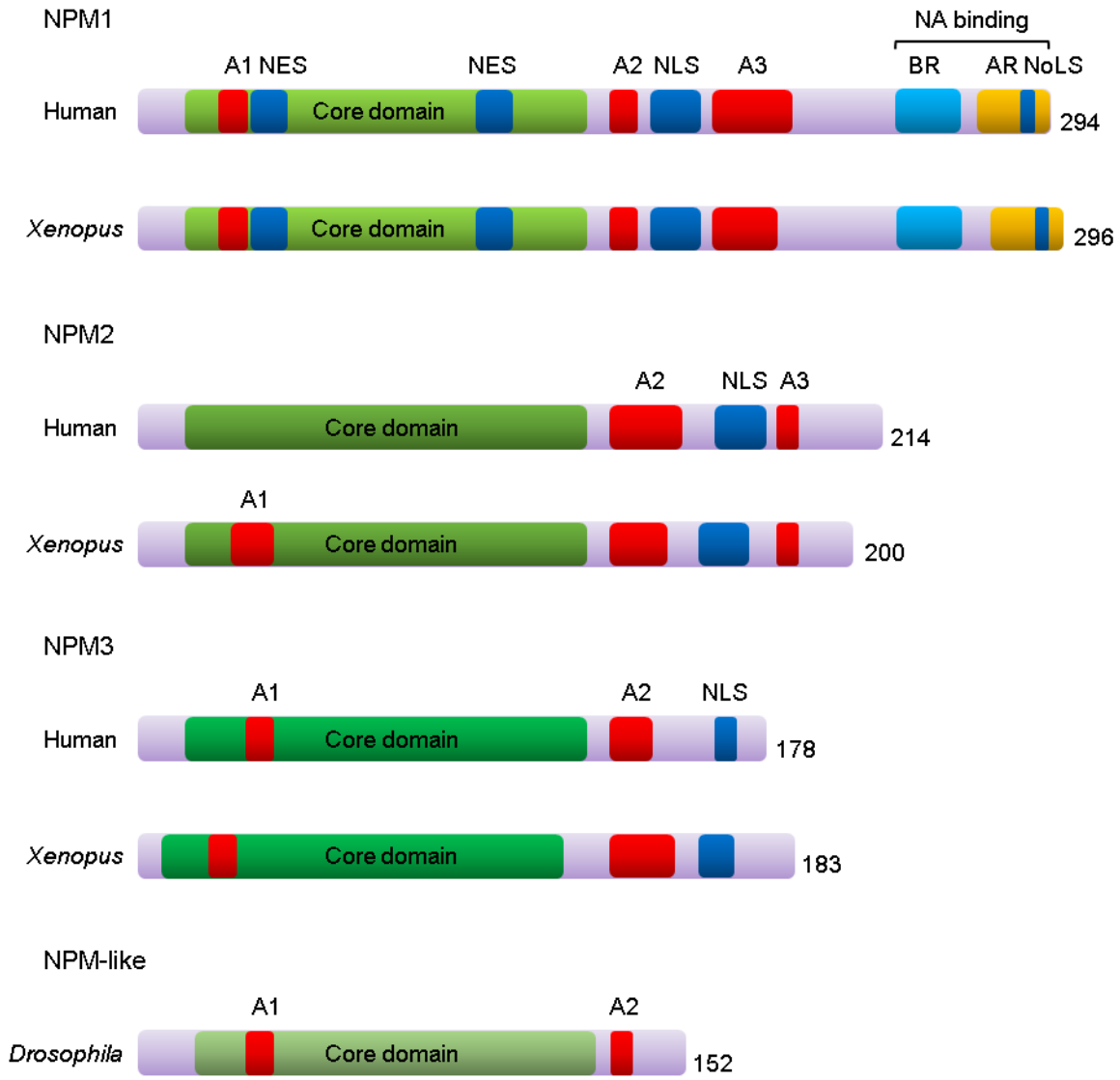



**Figure 1.9. Domain architecture of the Nucleoplasmin family of histone chaperones:** Domain structures of the human, *Xenopus,* and *Drosophila* homologs of Nucleoplasmin (NPM) family members namely NPM1, NPM2, NPM3, and NPM-like protein (NLP) shown for comparison. The N-terminal conserved core or oligomerization domain are shown in different shades of green, acidic stretches (A) are numbered (A1, A2, and A3) and shown in red, nuclear export signal (NES), bipartite nuclear localization signal (NLS) and nucleolar localization signal (NoLS) are shown in dark blue, the basic region (BR) is shown in light blue, and the aromatic region (AR) is shown in yellow. The nucleic acid (NA)-binding region is indicated. Numbers at the right indicate the length of the polypeptide. The figure has been adapted from (Frehlick et al. 2007).

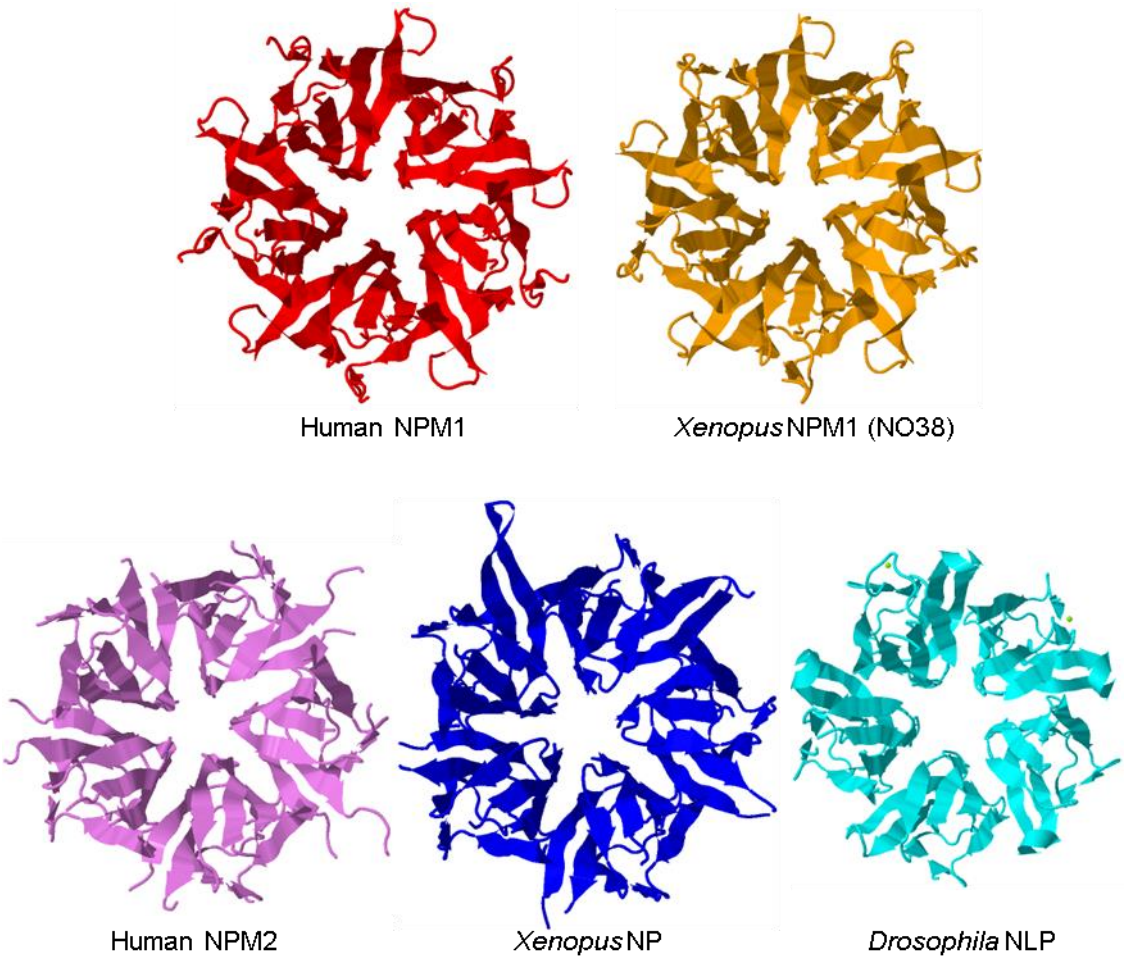



**Figure 1.10. Structural homology of the members of Nucleoplasmin family:** Pentameric core domain structure of human NPM1 (red), *Xenopus* NPM1 known as NO38 (yellow), human NPM2 (violet), *Xenopus* Nucleoplasmin (NP) (blue), and *Drosophila* Nucleoplasmin like protein (NLP) (cyan) are shown. Cartoon structures of the core domains of different NPM proteins, namely human NPM1 (PDB ID: 2P1B), *Xenopus* NO38 (PDB ID: 1XB9), human NPM2 (PDB ID: 3T30), *Xenopus* NP (PDB ID: 1K5J), and *Drosophila* NLP (PDB ID: 1NLQ) were obtained from RCSB PDB archive (https://www.rcsb.org).

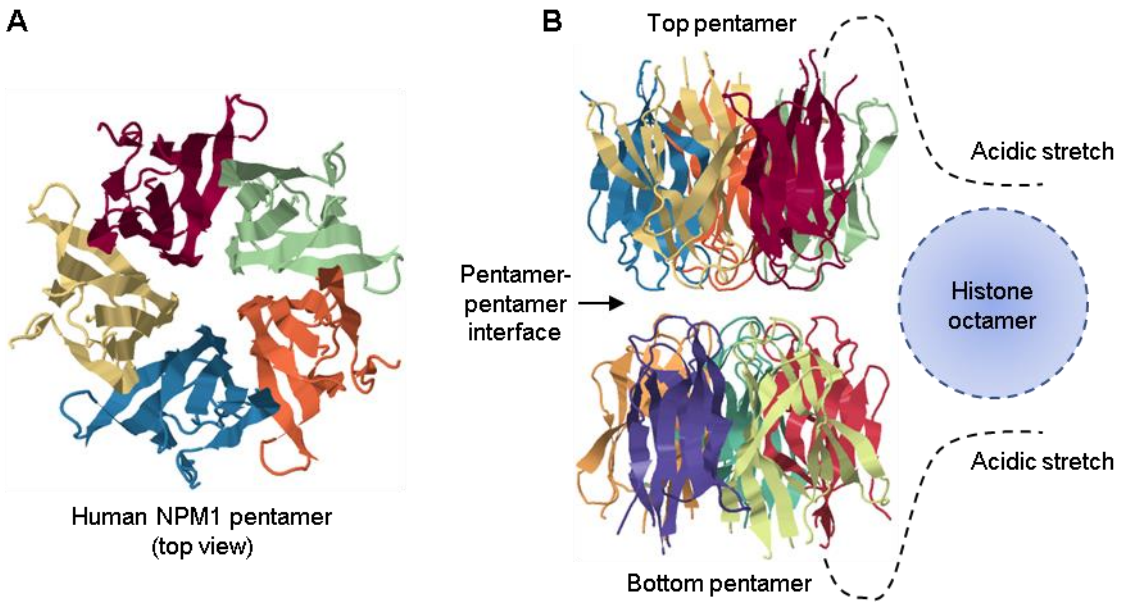



**Figure 1.11: Mode of histone binding by Nucleoplasmin homologs:** (A) Pentameric structure of the core domain of human NPM1 as seen from the top. The individual monomers are shown in different colors. (B) Structure of the NPM1 core domain in the decameric assembly of two pentamers oriented in a head-to-head manner. The lateral surface of the NPM1 decamer is proposed to contact the histone-octamer, an interaction that is further stabilized by the flexible and disordered acidic stretches indicated in black dashed lines. Cartoon structures of the core domains of human NPM1 (PDB ID: 2P1B) were obtained from the RCSB PDB archive (https://www.rcsb.org).

As mentioned previously, the three members of this family have orthologs in other vertebrates, which have been well characterized only in the case of *Xenopus laevis*. In invertebrates like *Drosophila melanogaster*, there is the presence of a Nucleoplasmin like protein (NLP) (Kawasaki et al. 1994; Ito et al. 1996) (Figure 1.9) and a homolog of Nucleophosmin (NPH) (Eirin-Lopez et al. 2006; Emelyanov et al. 2014). Recently, other proteins sharing close homology to the Nucleoplasmin family and having the characteristic pentameric Nucleoplasmin fold in their structures, have also been reported in *Drosophila*, *Arabidopsis thaliana*, *Saccharomyces cerevisiae*, and *Tetrahymena thermophila*, which can associate with chromatin, indicating that even if Nucleoplasmin family is not found in invertebrates, structurally and functionally similar proteins are present in other organisms as well (Edlich-Muth et al. 2015; Kozlowska et al. 2017; Leung et al. 2017; Ashraf et al. 2019).

### *1.4.1. Nucleophosmin (NPM1)*

NPM1 or Nucleophosmin is one of the most extensively studied proteins in this family because of its diverse functions and involvement in diseases like cancer. Also known as B23 and Numatrin in mammals, and NO38 in amphibians, NPM1 was originally identified as a nucleolar phosphoprotein, isolated from the nucleoli of rat Novikoff hepatoma ascites cells by the group of Harris Busch, at around 1973 – 1974 (Orrick et al. 1973; Kang et al. 1974; Olson et al. 1974). It was named B23 as it was the 23$^{rd}$ spot in the B region of the 2D gel where spots from the resolved proteins were numbered in the order of decreasing mobility. Independently, another group found this protein to be tightly associated with the nuclear matrix and named it Numatrin (Feuerstein and Mond 1987; Feuerstein et al. 1988a). Parallelly the *Xenopus* ortholog of B23 that is NO38 was identified and found to be homologous to Nucleoplasmin (Schmidt-Zachmann et al. 1987). Over time, 'Nucleophosmin' emerged as the most popular name for this protein.

The human Nucleophosmin (NPM1) gene is present in chromosome 5 in the 5q35.1 region. The gene is about 23.7 kb in length with 13 exons as per the current human genome assembly (GRCh38.p13). According to this assembly, there are 8 transcript variants of NPM1 and 6 isoforms. Variant 1 and 7 both encode the longest, major, and most well-studied isoform (294 amino acids long). Transcript variant 3 utilizes an alternate 3'-terminal exon (exon 10), resulting in a shorter protein (isoform 3 of 259 amino acids length) with a distinct C-terminus. The isoforms 1 and 3 of humans are B23.1 and B23.2 respectively in the rat (Chang and Olson 1989). Transcript variants 2, 4, 5, and 6 have alterations in their sequences due to the skipping of specific exons or differences in their 5'- or 3'-UTR sequences, resulting in shorter isoforms consisting of 265, 230, 230, and 167 amino acids respectively. Transcript variant 8 uses an alternate splice junction and lacks an alternate internal exon compared to variant 1. This variant is represented as non-coding since the predicted protein does not meet the quality criteria of the RefSeq database. The predominant isoform 1 of NPM1 (or B23.1 in rat) is localized to the nucleolus (Michalik et al. 1981; Spector et al. 1984; Okuwaki et al. 2001a; Okuwaki et al. 2002) whereas isoform 3 or B23.2 is expressed at much lower levels in cells and localized in the nucleoplasm or cytoplasm (Wang et al. 1993; Dalenc et al. 2002; Okuwaki et al. 2002). The isoforms other than 1 and 3 are not well-characterized. Recently it has been found that in acute myeloid leukemia (AML) where the *NPM1* gene is commonly mutated, deregulation in its splicing

pattern results in the formation of circular RNAs of *NPM1* (Hirsch et al. 2017). The precise functions of these *NPM1* circRNAs have not been understood yet.

#### *1.4.1.1. Domain architecture, structure, and, cellular localization of NPM1*

The NPM1 protein has several domains and sequence motifs, some of which are conserved and present in other members of the Nucleoplasmin family, while some are unique to NPM1 (Figure 1.9). As a result of these domains, NPM1 can perform multiple functions, some being exclusive to itself and not shared by the other paralogs (Frehlick et al. 2007).

The N-terminal region (residues 1 – 119 of human NPM1) contains the most conserved and characteristic domain of the Nucleoplasmin family members, known as the core domain (Section 1.4 and Figure 1.9). This domain is responsible for the oligomerization property and histone chaperone activity of NPM1 (Swaminathan et al. 2005). It attains a very conserved and stable secondary structure which is a beta-barrel with jelly-roll topology (Section 1.4). The core domain of one monomer interacts with the core domain of another monomer primarily through hydrophobic interactions to attain a pentameric form (Figure 1.11A) which then aligns with another pentamer in a head-to-head fashion to attain a decameric structure (Figure 1.11B) (Lee et al. 2007a). The pentamer-pentamer association is contributed through hydrogen bonding interactions by water molecules at the interface of the two pentamers (Figure 1.11B), as well as through electrostatic interactions by the Asp (of a highly conserved AKDE loop) and Lys residues. The crystal structures of the core domains of different NPM proteins (Figure 1.10) indicate that they, including NPM1, bind to the histone octamer mainly through the lateral surface of the decamer and the flexible acidic stretches then stabilize the complex (Figure 1.11B). The first acidic stretch (A1) is present within the core domain whose function is not clearly understood in the case of human NPM1 but is required for sperm chromatin decondensation in the case of *Xenopus* Nucleoplasmin. The core domain contains another conserved motif known as the GSGP loop which is important for the stability and oligomerization of NPM1, its localization to the nucleolus, as well as mediating its interaction with tumor suppressor ARF (Enomoto et al. 2006). The core domain also contains two stretches (residues 42 – 49 and 94 – 102) of the nuclear export signal (NES) (Figure 1.9). The first NES helps in the nuclear export of ribosomal subunit proteins and rRNA in a CRM1-dependent manner (Yu et al. 2006; Maggi et al. 2008). The second NES helps in the export of NPM1 to the cytoplasm through the

Ran-Crm1 pathway and is required for the localization of NPM1 to the centrosome to prevent premature centrosome reduplication (Wang et al. 2005). Following the core domain, the remaining structure of NPM1 is majorly disordered. In this region, there are two more acidic stretches A2 (residues 120 – 132) and A3 (161 – 188) which are separated by a characteristic bipartite nuclear localization signal (NLS; residues 152 – 157 and 191 – 197) (Figure 1.9). The acidic stretches mediate the binding of NPM1 to basic proteins such as histones and sperm basic proteins presumably through electrostatic interactions to promote nucleosome assembly and chromatin remodeling (Swaminathan et al. 2005; Gadad et al. 2011b) as well as help in stimulating Adenoviral DNA replication *in vitro* (Okuwaki et al. 2001a) and interactions with viral proteins and peptides (Adachi et al. 1993; Szebeni et al. 1995; Szebeni et al. 1997; Huang et al. 2001; Samad et al. 2007). Following the acidic stretches, there is a basic region (residues 219 – 242) which has a predominance of basic residues lysine and arginine (Figure 1.9). This is followed by an aromatic region (residues 258 – 294) which is enriched in aromatic residues. The basic and aromatic regions constitute the nucleic acid-binding region (residues 219 – 294) which mediates the binding of NPM1 to DNA/RNA molecules (including G-quadruplex DNA structures) that also has an effect on nucleosome assembly. This region enhances a ribonuclease activity located in the central region of NPM1 (Hingorani et al. 2000) as well as binds to tumor suppressor p53 leading to its enhanced stability and transcriptional activity (Colombo et al. 2002). The C-terminal aromatic region which is unique to NPM1 isoform 1, contains the nucleolar localization signal (NoLS) (Figure 1.9) where two tryptophan residues namely Trp288 and Trp290, are critical for the nucleolar localization of NPM1 (Nishimura et al. 2002; Falini et al. 2006a; Falini et al. 2006b). In a particular type of acute myeloid leukemia (AML) known as NPM1c+ AML, a frame-shift mutation in the 12$^{th}$ exon of the *NPM1* gene results in the generation of an altered C-terminal sequence with mutations of the Trp288 and Trp290 residues and disruption of the NoLS as well as the generation of a new NES at the C-terminus. This synergistically causes the delocalization of the mutated NPM1 from the nucleolus to the cytoplasm (Falini et al. 2006a; Falini et al. 2006b). Structural studies of the C-terminal end of NPM1 has revealed that specific conserved residues such as Trp288, Trp290, Phe268, Phe276 along with certain surface-exposed lysine residues contribute towards the structural integrity of the NoLS at the C-terminus (Grummitt et al. 2008). The C-terminal region (residues 242 – 294) is also capable of binding to ATP (Chang et al. 1998) with Lys263 specifically being responsible for this activity as well as enhancing NPM1 stability (Choi et al. 2008).

As discussed previously, the various sequence motifs in NPM1 protein structure such as the NES, NLS, and NoLS allow NPM1 to be a highly mobile nucleo-cytoplasmic shuttling protein (Borer et al. 1989), although the majority of NPM1 is found to be localized to the nucleolus, in its granular region (Spector et al. 1984). Due to this property, NPM1 can mediate the shuttling of proteins between the nucleolar, nuclear, and cytoplasmic compartments, such as that seen in case of the ribosomal proteins (Yu et al. 2006; Maggi et al. 2008), small basic viral proteins (Fankhauser et al. 1991; Adachi et al. 1993; Szebeni et al. 1995; Li 1997; Szebeni et al. 1997), among others. A minor pool of NPM1 is present in the nucleoplasm often harboring specific PTMs such as acetylation and phosphorylation. Acetylated NPM1 present in the nucleoplasm was found to colocalize with RNA Pol II foci indicating its role in transcription (Shandilya et al. 2009). CKII-mediated phosphorylated NPM1 at Ser125 was also found to be localized to the nucleoplasm during interphase of the cell cycle (Negi and Olson 2006; Shandilya et al. 2014a). During mitosis, NPM1 has been observed to be localized at various regions of the cell such as in certain structures known as the nucleolus-derived foci (NDF) (Dundr and Olson 1998), the mitotic spindle poles (Zatsepina et al. 1999), the perichromosomal layer, the prenucleolar bodies (PNBs) (Zatsepina et al. 1997), as well as the midbody regions during cytokinesis (Shandilya et al. 2014a). NPM1 is also found to be localized to the centrosome for which its NES is required (Wang et al. 2005). The nucleolar localization of NPM1 is not only determined by the presence and integrity of its NoLS as discussed earlier but also on its functional oligomeric state (Enomoto et al. 2006) and stability imparted by the residues Lys263 and Lys267 (Choi et al. 2008; Grummitt et al. 2008). The cytoplasmic localization of NPM1 often occurs abnormally in AML due to mutation-induced disruption of the C-terminal NoLS and gain of an extra NES (Falini et al. 2006a; Falini et al. 2006b).

#### *1.4.1.2. Post-translational modifications of NPM1*

NPM1 undergoes various post-translational modifications (PTMs) in residues present throughout its protein structure. Some of these PTMs such as phosphorylation, acetylation, SUMOylation, among others, have been well studied and found to impart several important functions to the protein as discussed briefly in the following sections.

### *1.4.1.2.1. Phosphorylation*

Phosphorylation is the most predominant PTM of NPM1. NPM1 was first identified as a major phospho-protein in the nucleolus (Orrick et al. 1973; Kang et al. 1974; Olson et al. 1974). Since then, several kinases have been found to phosphorylate NPM1 such as CKII (Szebeni et al. 2003), Plk1 (Zhang et al. 2004), Plk2 (Krause and Hoffmann 2010), cdc2 (Peter et al. 1990), cyclin E/CDK2 (Okuda et al. 2000), IKKα (Xia et al. 2013), AURKA (Reboutier et al. 2012), AURKB (Shandilya et al. 2014a), GRK5 (So et al. 2012), and CDK6 (Lin et al. 2019a). Phosphorylation of NPM1 has been implicated in various cellular processes and contexts. For example, cyclin E/CDK2-mediated phosphorylation of NPM1 at Thr199 resulted in the dissociation of NPM1 from the centrosomes and allowed centrosome duplication (Okuda et al. 2000; Tokuyama et al. 2001) as well as targeted NPM1 to the nuclear speckles and repressed pre-mRNA processing (Tarapore et al. 2006). Thr199-phosphorylated NPM1 was also found to enhance the inhibition of GCN5-mediated histone acetylation and transactivation (Zou et al. 2008). cdc2-mediated phosphorylation of NPM1 is important for its RNA-binding and rRNA chromatin-binding activity that regulates the organization of the nucleolar chromatin (Okuwaki et al. 2002; Hisaoka et al. 2010). Plk-mediated phosphorylation of NPM1 at Ser4 triggers centriole duplication and ensures proper mitotic progression (Zhang et al. 2004; Krause and Hoffmann 2010). Aurora Kinase B-mediated phosphorylation of NPM1 at Ser125 was also found to be important for mitotic progression (Shandilya et al. 2014a). CKII-mediated phosphorylation of NPM1 promotes its substrate release capacity thereby affecting its molecular chaperone activity (Szebeni et al. 1997; Szebeni et al. 2003). Thus, phosphorylation can affect various aspects of NPM1 function, structure, and cellular localization (Negi and Olson 2006; Mitrea et al. 2014). NPM1 phosphorylation has been found to be upregulated in several cancers indicating its role in cancer manifestation and progression (Shandilya et al. 2014a; Destouches et al. 2016; Lin et al. 2019a). A schematic representation of some of the phosphorylation sites in NPM1 protein sequence is shown in Figure 1.12A.

Dephosphorylation of NPM1 has been reported to be mediated by PP1β in response to DNA damage besides other genotoxic stress and growth conditions, that facilitate the DNA repair process (Lin et al. 2010).

### *1.4.1.2.2. Acetylation*

NPM1 has been reported to be acetylated by the lysine acetyltransferase p300 at multiple residues (Swaminathan et al. 2005; Shandilya et al. 2009) (Figure 1.12B). Acetylation of NPM1 enhances the histone chaperone activity and increases its binding affinity towards core histones, thereby potentially resulting in enhanced nucleosome disassembly. In agreement with this observation, acetylated NPM1 was found to be a better activator of acetylation-dependent chromatin transcription *in vitro* (Swaminathan et al. 2005). Acetylated NPM1 also helps in activator-dependent transcription which is seen in the case of the transactivator protein HIV-Tat (Gadad et al. 2011a). *In vivo*, acetylated NPM1 was found to be localized in the nucleoplasm, colocalizing with RNA Pol II foci indicating its role in cellular transcription. Its levels were observed to be elevated in oral cancer and it could associate with RNA Pol II at the promoters of genes implicated in cancer manifestation, bringing about the upregulation of their expression (Shandilya et al. 2009). Acetylated NPM1 was also found to be higher in senescent cells indicating its role in the process of aging (Lee et al. 2014).

Deacetylation of NPM1 has been reported to be brought about by SIRT1 (Shandilya et al. 2009), SIRT6 and SIRT7 (Lee et al. 2014; Kiran et al. 2015).

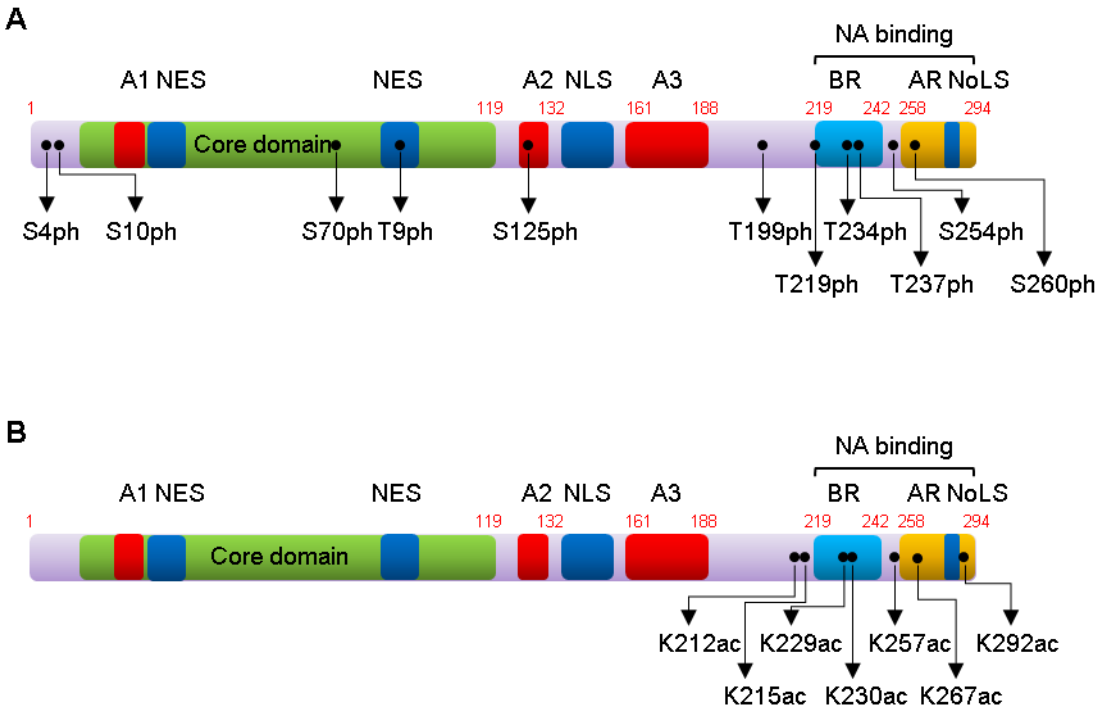



**Figure 1.12. Post-translational modifications of NPM1:** (A – B) Schematic representation of the NPM1 domain structure with the (A) phosphorylated and (B) acetylated sites indicated in black

circles. The residues are mentioned. Numbers in red indicate the residues at the boundaries of the various domains. A1, A2, and A3: acidic stretches 1, 2, and 3, NES: nuclear export signal, NLS: nuclear localization signal, BR: basic region, AR: aromatic region, NoLS: nucleolar localization signal, NA: nucleic acid, ph: phosphorylation, ac: acetylation. The figure has been adapted from Senapati P, Ph.D. thesis, 2014.

#### *1.4.1.2.3. SUMOylation*

NPM1 is reported to be SUMOylated which was induced by the tumor suppressor ARF (Tago et al. 2005). SUMOylation at Lys263 of NPM1 was found to be important for its centrosomal and nucleolar localization, its interaction with the tumor suppressor protein Rb and ErbB3-binding protein 1 (EBP1), regulation of E2F1-mediated transcriptional activity, as well as its resistance to apoptosis (Liu et al. 2007b; Okada et al. 2007). TRIM28 is the E3 ligase for ARF-mediated NPM1 SUMOylation (Neo et al. 2015) while SENP3 is the SUMO deconjugating enzyme for NPM1 (Haindl et al. 2008; Nishida and Yamada 2008). The induction of NPM1 SUMOylation by ARF occurs by promoting the turnover of SENP3 (Kuo et al. 2008). SUMOylation also has an effect on the stability of NPM1 (Lee et al. 2008; Vishwamitra et al. 2015). In the context of DNA repair, NPM1 SUMOylation was found to be important for the recruitment of DNA repair proteins at the initial stages of DNA-damage response (DDR), and SUMOylated NPM1 affected the assembly of the BRCA1 complex in AML cells (Xu et al. 2019).

#### *1.4.1.2.4. Other PTMs of NPM1*

NPM1 has been reported to get polyubiquitinated which was promoted by ARF and linked to its degradation (Itahana et al. 2003; Enomoto et al. 2006). However, the polyubiquitination of NPM1 mediated by the BRCA1-BARD1 complex causes its stabilization rather than degradation since this particular ligase catalyzes the conjugation of untraditional polyubiquitin chains. This modification has been suggested to regulate NPM1 function during mitosis (Sato et al. 2004). USP36 is the deubiquitinating enzyme for NPM1 that causes NPM1 protein stabilization and improved nucleolar function (Endo et al. 2009).

NPM1 undergoes poly(ADP-ribosyl)ation after exposure to ionizing radiation (Ramsamooj et al. 1995). It is associated with the poly(ADP-ribose) polymerases PARP1 and PARP2

accumulated in transcriptionally active nucleoli (Meder et al. 2005). The functional significance of this association is unclear.

NPM1 was found to be citrullinated by PADI4 at Arg197 residue which resulted in the translocation of NPM1 from the nucleoli to the nucleoplasm (Tanikawa et al. 2009). S-nitrosylation of NPM1 by GAPDH at Cys275 has been shown to confer a neuro-protective role of NPM1 by preventing SIAH1-GAPDH death signaling under stress-induced conditions in the brain (Lee et al. 2012b). Further, NPM1 has been reported to get S-glutathionylated at Cys275, which promotes the dissociation of NPM1 from the nucleolar nucleic acids thereby mediating the nucleoplasmic translocation of NPM1 for stress-induced activation of p53 (Yang et al. 2016a).

#### *1.4.1.3. Functions of NPM1*

NPM1 performs diverse functions that are attributed to the presence of various domains and sequence motifs in NPM1 structure as well as its partial disordered nature that enables NPM1 to interact with a multitude of proteins and thereby participating in several cellular processes. These functions are summarized in Figure 1.13 and described briefly in the following sections.

##### *1.4.1.3.1. Protein chaperoning*

NPM1 acts as a molecular chaperone for several proteins (Szebeni and Olson 1999). This activity is mainly present in its N-terminal predominantly hydrophobic core domain, with contributions from its acidic stretches. Hence oligomerization of NPM1 is necessary for the manifestation of its molecular chaperone activity (Hingorani et al. 2000). Through this property, NPM1 can prevent temperature-dependent and independent aggregation of proteins, preserve the activities of several enzymes during thermal denaturation, promote renaturation of pre-denatured proteins, and preferentially bind to denatured proteins and expose their hydrophobic regions during interaction with other proteins (Szebeni and Olson 1999). Although NPM1 can bind to ATP (Chang et al. 1998), ATP is not required for its molecular chaperone activity (Szebeni and Olson 1999). The molecular chaperone activity of NPM1 is utilized especially in the process of ribosome biogenesis where it prevents the aggregation of ribosomal proteins and promotes the assembly of ribosomal subunits, as

well as for the assembly of viral particles (Szebeni and Olson 1999; Bevington et al. 2007; Finsterbusch et al. 2009; Ugai et al. 2012; Day et al. 2015). This specific function in one way or the other has also been implicated in different neuronal pathways and deregulation of NPM1 chaperone activity and/or expression could contribute to the development of neurological disorders such as schizophrenia (Mladinov et al. 2016), Parkinson's disease (Xie et al. 2016), Huntington's disease (Pfister and D'Mello 2016), spinal cord injury (Guo et al. 2014), and spinocerebellar ataxia type 17 (Lee et al. 2009).

NPM1 can also bind to histones and act as a histone chaperone mediating the ordered assembly and disassembly of nucleosomes (Okuwaki et al. 2001b; Swaminathan et al. 2005). For this function also, the N-terminal oligomerization domain of NPM1 is necessary, while the C-terminal region contributes further to the activity (Swaminathan et al. 2005; Lee et al. 2007a). The histone chaperone activity of NPM1 is also manifested towards linker histone H1 where it could facilitate the deposition or removal of H1 to or from specific *in vitro* assembled nucleosomal templates (Gadad et al. 2011b).

#### *1.4.1.3.2. Ribosome Biogenesis*

NPM1 was first isolated from the nucleolus which is the site for ribosomal RNA synthesis and processing. Initial studies with this protein suggested that NPM1 could be a ribosome assembly factor or a ribosome chaperone. NPM1 was found to be associated with the pre-ribosomal particle such as the 60S subunit, as well as the assembled 80S ribosome and polysomes (Yu et al. 2006). The properties of NPM1 that determine its nucleolar localization, nucleo-cytoplasmic shuttling, nucleic acid binding, and interaction with and transport of the pre-ribosomal particles are implicated in its role in ribosome biogenesis (Olson et al. 1986; Borer et al. 1989; Dumbar et al. 1989; Yun et al. 2003). The intrinsic ribonuclease activity of NPM1 (Herrera et al. 1995) helps in the processing of pre-ribosomal RNA by cleaving at a specific site in the ITS2 (internal transcribed spacer 2) of the pre-5.8S rRNA (Savkur and Olson 1998). NPM1 also helps in the transcription of rRNA genes and its histone chaperone activity plays an important role in this process (Murano et al. 2008). NPM1 interacts with a number of ribosomal proteins such as RPL5 (Yu et al. 2006), RPS9 (Lindstrom and Zhang 2008) and RPL23 (Wanzel et al. 2008). Given these specific roles played by NPM1 in ribosome biogenesis, it has been observed as expected that reducing the levels of functional NPM1 through knockdown, degradation or inhibition

of activity results in changes in the profile of ribosomes including defects in the pre-ribosomal RNA processing and export of pre-ribosomes (Itahana et al. 2003; Grisendi et al. 2005; Maggi et al. 2008). However, knockout studies in mice have suggested that NPM1 is not essential for ribosome biogenesis since there seems to be a compensation of NPM1 function for ribosome biogenesis by other factors having overlapping roles (Colombo et al. 2005; Grisendi et al. 2005; Grisendi et al. 2006).

#### *1.4.1.3.3. DNA replication and cell cycle regulation*

NPM1 was found to associate with and enhance the activity of DNA Polymerase α, which is responsible for the initiation of DNA replication and synthesis of the Okazaki fragments (Feuerstein et al. 1990; Takemura et al. 1994; Takemura et al. 1999). This activity is localized to the C-terminal 12 residues of NPM1, which was responsible for the increase in the amount and length of the product DNA but not the processivity of the DNA Polymerase enzyme. NPM1 also conferred stability to the enzyme and protected it from heat inactivation (Umekawa et al. 2001). NPM1 could stimulate Adenoviral DNA replication *in vitro* (Okuwaki et al. 2001a; Ugai et al. 2012) as well as bovine immunodeficiency virus (BIV) replication (Passos-Castilho et al. 2018). However, in other instances, NPM1 was found to have a negative effect on the replication of the Chikungunya virus (Abraham et al. 2017) and Adeno-associated viruses (Satkunanathan et al. 2017). NPM1 was also found to regulate translesion DNA synthesis (TLS) by interacting with the catalytic core of DNA polymerase-η (polη). NPM1 deficiency or presence of NPM1c+ mutant was found to result in defects in TLS due to proteasomal degradation of polη (Ziv et al. 2014). In addition, NPM1 was found to be important for telomerase activation and maintenance of telomere length (Ho et al. 2019).

NPM1 plays an important role in the process of cell cycle and centrosome duplication. It has been observed that there is a redistribution of NPM1 from the nucleolus to various cellular locations during the different stages of mitosis (Hernandez-Verdun and Gautier 1994; Zatsepina et al. 1997). A fraction of NPM1 was found to be localized to the mitotic spindle poles along with nuclear mitotic apparatus protein (NuMA) during the metaphase stage of the cell cycle (Zatsepina et al. 1999). Phosphorylation of NPM1 has been shown to have major implications in the mitotic process (Bergstralh et al. 2007). In the context of normal development, NPM1 helps in the maintenance of normal centrosome since reduced

levels or absence of NPM1 results in unrestricted centrosome duplication and genomic instability (Grisendi et al. 2005). At the molecular level, NPM1 is associated with unduplicated centrosomes, which then dissociates upon cyclin E/CDK2-mediated phosphorylation at Thr199 (which can also be induced by a loss of the protein 14-3-3γ), thereby allowing centrosome duplication (Okuda et al. 2000; Tokuyama et al. 2001; Mukhopadhyay et al. 2016). The phosphorylation of NPM1 by various other kinases at different stages of the cell cycle (Jiang et al. 2000) such as cdc2, AURKB, CDK6, has been shown to have implications in centrosome duplication and mitotic progression (Peter et al. 1990; Cuomo et al. 2008; Shandilya et al. 2014a). In addition, BRCA1-BARD1-mediated ubiquitination of NPM1 has been linked to its localization to the centrosomes (Sato et al. 2004) and suggested to be important for maintaining the integrity of spindle poles and the ploidy of cells (Grisendi et al. 2006).

#### *1.4.1.3.4. Cell survival and apoptosis*

NPM1 has been shown in general to promote cell survival and proliferation and inhibit apoptosis which gives significant advantages to cancerous cells (Grisendi et al. 2006). It acts as a nuclear receptor for the second messenger phosphatidylinositol (3,4,5)-triphosphate [PI(3,4,5)P3] that inhibits caspase-activated DNase (CAD) in nerve growth factor (NGF)-treated PC12 cells and prevents DNA fragmentation (Ahn et al. 2005). Hypoxia-induced overexpression of NPM1 was found to inhibit p53-mediated apoptosis (Li et al. 2004). NPM1 also had a protective role from UV-induced cell death in NIH3T3 cells by enhancing nucleotide excision repair and expression of PCNA (Wu and Yung 2002). Further, NPM1 inhibits apoptosis by negatively regulating pro-apoptotic proteins such as PKR (Pang et al. 2003) or being associated with anti-apoptotic factors such as the long non-coding RNA Lnc_bc060912 (Luo et al. 2015) under specific contexts. In agreement with the strong anti-apoptotic and pro-survival effect of NPM1 on cancer cells, it was found that NPM1 was transcriptionally downregulated upon chemical-induced apoptosis such as by sodium butyrate and vanadate in human leukemia HL-60 cells (Liu and Yung 1998) and phycoerythrin in human SW480 tumor cells (Li et al. 2016). Likewise, targeting NPM1 through factors such as miRNA-181a in laryngeal cancer cells (Wang et al. 2019) or inducing cytosolic translocation and degradation of NPM1 by treating human lung non-small cell carcinoma cell line H460 with aloe-emodin (1,8-dihydroxy-3-

(hydroxymethyl)-anthraquinone) (Lee et al. 2005a) thereby reducing its levels, was found to enhance apoptosis.

However, a recent study has suggested a pro-apoptotic role of NPM1 as well, where NPM1 was found to play a direct and conserved role in DNA damage-induced assembly of the PIDDosome complex, the activating platform for caspase-2, in the nucleolus, and inhibition of NPM1 was found to impair caspase-2 processing, apoptosis, and caspase-2-dependent inhibition of cell growth (Ando et al. 2017; Sidi and Bouchier-Hayes 2017).

#### *1.4.1.3.5. Stress response and DNA repair*

NPM1 is highly responsive to cellular stresses and several studies have indicated NPM1 to play a role in the repair of DNA damage which is often induced by different kinds of stress. DNA damage often causes the translocation of NPM1 from the nucleolus to the nucleoplasm where it associates with the chromatin (Lee et al. 2005b). Inhibiting the ribosome biogenesis pathway at any step without causing much DNA damage can also cause nucleoplasmic translocation of NPM1 from the nucleolus (Yung et al. 1985; Chan 1992) where it can interact with and regulate several pathways, thus acting as a nucleolar stress sensor (Scott and Oeffinger 2016; Yang et al. 2016a). Under oxidative stress, NPM1 levels have been found to decrease, and its rRNA binding properties are affected by its interaction with CacyBP/SIP under these conditions (Rosinska and Filipek 2018). NPM1 is also recruited at specific DNA damage foci in its Thr199 phosphorylated form through RNF8-dependent ubiquitin conjugates, upon IR-induced DNA damage (Koike et al. 2010). UV-induced DNA damage results in an upregulation of NPM1 levels (Wu et al. 2002a; Wu and Yung 2002) as well as an increase in its RNA-binding activity (Yang et al. 2002). An enhanced nucleotide excision repair (NER) pathway and PCNA expression resulting from the increased levels of NPM1, renders the cells more resistant to UV-induced cell death (Wu et al. 2002b). In response to gamma radiation, NPM1 was found to get dephosphorylated at Thr199 and Thr234/237, redistributed throughout the cell (nucleolus, nucleoplasm, cytoplasm), and promote tumor cell survival post-irradiation (Wiesmann et al. 2019). NPM1 was also found to be critical for DNA double-strand break (DSB) repair pathway where loss of NPM1 increases the radiation sensitivity of MEFs (Sekhar et al. 2014), as well as base excision repair (BER) pathway where it controls BER protein levels, regulates total BER capacity, and modulates the nucleolar localization of several BER

enzymes (Vascotto et al. 2009; Poletto et al. 2014; Vascotto et al. 2014). NPM1 could potentially function as a histone chaperone during or after DNA damage to shield the histones after their removal and reassembling them into the chromatin following DNA damage.

#### *1.4.1.3.6. Transcription*

NPM1 has been shown to regulate both RNA Pol I- and RNA Pol II-mediated transcription. It regulates the synthesis of rRNA in the nucleolus for which its histone chaperone activity was found to be important (Murano et al. 2008). It is required for the upregulation of the rDNA transcription factor TAF1A (also known as TAF(I)48) (Bergstralh et al. 2007). The binding of NPM1 to the rDNA or the nucleolar chromatin (Hisaoka et al. 2010) and its ability to facilitate the nucleolar translocation of the RNA Pol I transcription termination factor TTF1 (Lessard et al. 2010), are also implicated in the process of regulation of RNA Pol I transcription. In the context of epigenetic regulation of transcription, it has been recently shown that NPM1 can bind to the H3K4me2 mark but not H3K4me1 and H3K4me3, through its acidic tracts, having potential implications in the transcriptional activation of rDNA genes (Wu et al. 2017). The organization of the nucleolar chromatin by the association of NPM1 with HP1γ, core, and linker histones, also has implications in the regulation of rDNA transcription (Holmberg Olausson et al. 2014). Recently it has been reported that NPM1 can interact with a novel histone H4 variant H4G and localize it to the nucleolar chromatin where the latter can enhance rDNA transcription in breast cancer (Long et al. 2019). RNA Pol I-mediated rDNA transcription is commonly promoted in cancer cells, aided by the inactivation of tumor suppressors such as p53 and hyperactivation of oncogenes such as c-myc (White 2008; Drygin et al. 2010). In this context, NPM1 can also indirectly affect RNA Pol I transcription by modulating the properties and functions of tumor suppressors and oncogenes (Section 1.4.1.3.7).

Several studies have indicated the involvement of NPM1 in RNA Pol II-driven transcription. It can mediate various activator-dependent transcription in cells such as by Gal4-VP16, Tat (Swaminathan et al. 2005; Gadad et al. 2011a), and so on. It has been found to regulate the expressions of specific sets of genes through its interactions with specific transcription factors such as p53 (Colombo et al. 2002), NF-κB (Dhar et al. 2004; Lin et al. 2017), YY1 (Inouye and Seto 1994), c-myc (Li et al. 2008), IRF1 (Kondo et al. 1997; Abe

et al. 2018), STAT1 (Abe et al. 2018), Oct4, Sox2, Nanog (Johansson and Simonsson 2010), AP-2γ (Lin et al. 2016), or viral proteins such as EBNA2 (Liu et al. 2012a). The mechanisms of NPM1-meditated transcriptional regulation could be diverse. For example, NPM1 can increase the stability of a transcription factor such as in the case of p53, thereby enhancing p53 transcriptional activity under stress conditions (Colombo et al. 2002). NPM1 can act both as a coactivator such as in the case of NF-κB in regulating the expression of the human SOD2 gene (Dhar et al. 2004), or a co-repressor such as in the case of AP-2 alpha and AP-2 gamma where its recruitment at the AP-2 alpha or gamma target gene promoters, resulted in their decreased expression during the specific contexts of RA-mediated cellular differentiation (Liu et al. 2007a) or endometrial cancer (Lin et al. 2016). Its interaction with YY1 was found to relieve the YY1-mediated transcriptional repression (Inouye and Seto 1994). NPM1 is also suggested to co-occupy gene promoters along with other TFs such as c-myc and IRF1 and enhance their cognate gene expression (Li et al. 2008; Abe et al. 2018). It can modulate the ability of certain TFs and chromatin-associated proteins such as IRF1, NF-κB, E2F1, Rb, and HMG1 to bind to specific gene promoters, thereby having either a positive or negative effect on their target gene expression (Kondo et al. 1997; Lin et al. 2006; Guery et al. 2011; Arnoldo et al. 2015; Lin et al. 2017). In a specific instance, wild-type NPM1 was found to inhibit BRD4-mediated core transcriptional activation program in AML, which was relieved upon NPM1c+ mutation and its cytoplasmic dislocation (Dawson et al. 2014). The association of NPM1 with chromatin remodeling complexes such as ISWI, NuRD, and P/BAF, can also serve as a mechanism to recruit them at the gene promoters thereby regulating downstream gene expression (Darracq et al. 2019). NPM1 can induce the reduction of the functional forms of transcription inhibiting proteins such as HEXIM1 through degradation or cytoplasmic translocation, thereby resulting in an increase in the P-TEFb-mediated RNA Pol II transcription in this particular scenario (Gurumurthy et al. 2008). NPM1 can also associate and/or modulate the activities of chromatin-modifying enzymes that impact transcription. For example, NPM1 was found to be present in the complex of the methyltransferase DOT1L, which catalyzes the methylation of H3K79 that has been implicated in active transcription (Park et al. 2010). NPM1 could also bind to the lysine acetyltransferase GCN5 and inhibit the GCN5-mediated acetylation of free and mononucleosomes and GCN5-mediated transactivation. This effect was more pronounced in the case of Thr199-phosphorylated NPM1 during mitosis suggesting that NPM1 plays a role in the mitotic inhibition of GCN5-mediated histone acetylation and transactivation (Zou et al. 2008). In

another instance, NPM1 was found to induce the autoacetylation of p300 through its oligomerization property or molecular chaperone activity, thereby enhancing the catalytic activity of p300 towards its substrates including nucleosomal histones implicated in transcriptional activation (Arif et al. 2010; Kaypee et al. 2018b). NPM1 itself gets acetylated which was found to enhance its ability to activate acetylation dependent chromatin transcription *in vitro* (Swaminathan et al. 2005). In cells, the acetylated NPM1 was found to colocalize with RNA Pol II in the nucleoplasm as well as co-occupy promoters of genes implicated in oral cancer (Shandilya et al. 2009).

#### *1.4.1.3.7. Modulation of protein activities*

Being a molecular chaperone, NPM1 has been shown to protect and/or modulate the catalytic activities of enzymes including some of the chromatin-modifying enzymes (discussed in Sections 1.4.1.3.1 and 1.4.1.3.6) (Szebeni and Olson 1999; Kaypee et al. 2018b). Besides enzymes, NPM1 has also been found to modulate the functional properties of tumor suppressors and oncogenes through their interactions which have important implications in cancer. In support of this fact, NPM1 has been found to be associated or colocalized with other tumor suppressors and oncoproteins in cells under different contexts (Li et al. 2010). Substantial investigations have been performed over the years regarding the modulation of activity and stability of two well-known tumor suppressors ARF and p53 by NPM1. ARF can inhibit cell proliferation through both p53-dependent (Honda and Yasuda 1999; Weber et al. 1999; Llanos et al. 2001) and independent (Weber et al. 2000) pathways. The well-studied interaction of NPM1 with ARF has been shown to result in various functional consequences in relation to ribosome biogenesis, nucleo-cytoplasmic shuttling, protein degradation, protein-protein associations, and regulation of the nucleolar homeostasis (Itahana et al. 2003; Bertwistle et al. 2004; Brady et al. 2004; Korgaonkar et al. 2005; Luchinat et al. 2018; Mitrea and Kriwacki 2018). The delineation of the universal biological consequence of NPM1–ARF interaction has not been clearly established (Colombo et al. 2011). NPM1 interacts with ARF and increases its stability by retarding its ubiquitination-mediated degradation (Kuo et al. 2004; Kuo et al. 2008) while ARF induces the degradation of NPM1 (Itahana et al. 2003). However, in another study, it was found that the NPM1–ARF association did not correlate with any growth suppression. Here, NPM1 controlled the nucleolar localization of ARF and competed with Mdm2 to associate

with ARF. The downregulation of NPM1 was found to result in an enhanced ARF–Mdm2 association and reduced ARF nucleolar localization. p53 is thus relieved from Mdm2-mediated inhibition leading to its activation and subsequent p53-mediated cell cycle arrest and apoptosis. Thus, in this case, NPM1 inhibited ARF's p53-dependent activity by targeting it to nucleoli and impairing ARF–Mdm2 association (Korgaonkar et al. 2005).

NPM1 has also been implicated in the p53 pathway both directly and indirectly. p53 gets activated and stabilized in response to various cellular stresses such as DNA-damaging drugs, UV radiation or metabolic stress, which are sensed by the nucleolus and often result in the disruption of the nucleolus (Rubbi and Milner 2003). Nucleolar disruption can result in the redistribution of ARF as well as NPM1 from the nucleolus to the nucleoplasm where they can stabilize p53 by associating with Mdm2, the E3 ubiquitin ligase of p53 (Kurki et al. 2004a; Kurki et al. 2004b). In a similar mechanism, downregulation of another protein DDX31, in renal cell carcinoma, resulted in the translocation of NPM1 from the nucleoli to the nucleoplasm where it bound to HDM2, thereby reducing HDM2–p53 association causing p53 stabilization (Fukawa et al. 2012). In another pathway, NPM1 was found to interact with the lncRNA SAMD12-AS1 in HBV-induced hepatocellular carcinoma, which resulted in an enhanced association of HDM2 with p53. This caused the ubiquitination-mediated p53 degradation with a concomitant increase in cell growth and inhibition of apoptosis (Liu et al. 2019). Similar downstream effect, that is the reduction in p53 levels with a concomitant enhancement of cell survival, is seen after the downregulation of NPM1 expression by phosphorylated STAT5 in cancer cells (Ren et al. 2016). Further, the NPM1-mediated stabilization of p53 also enhances its transcriptional activity (Colombo et al. 2002).

NPM1 has been shown to regulate the turnover of oncoproteins such as c-myc by mediating the nucleolar localization and stability of the F-box protein Fbw7γ which is a component of the E3 ubiquitin ligase complex for these oncoproteins including c-myc. In the absence of wild-type NPM1 or the presence of the NPM1 mutant NPM1c+ which was found to delocalize Fbw7γ to the cytoplasm, the oncoproteins could be stabilized whereby they could manifest their oncogenic properties (Bonetti et al. 2008). The presence of NPM1 promoted the stemness-like properties of lung cancer cells mediated by the oncogene CUG2 and downstream TGF-β signaling (Kaowinn et al. 2019). Further, the chimeric protein NPM-ALK generated due to *NPM1* gene translocation in anaplastic large cell lymphoma (ALCL), was found to enhance the DNA binding and transcriptional activity of

the oncoprotein FOXM1 (Haque et al. 2019). Interestingly, FOXM1 can regulate the stability of NPM1 and there appears to be a mutual regulation of each other by both these two proteins (Pandit and Gartel 2015).

Thus, NPM1 plays critical roles in modulating the properties and functions of tumor suppressors and oncogenic proteins in the complicated networks operating in cancer.

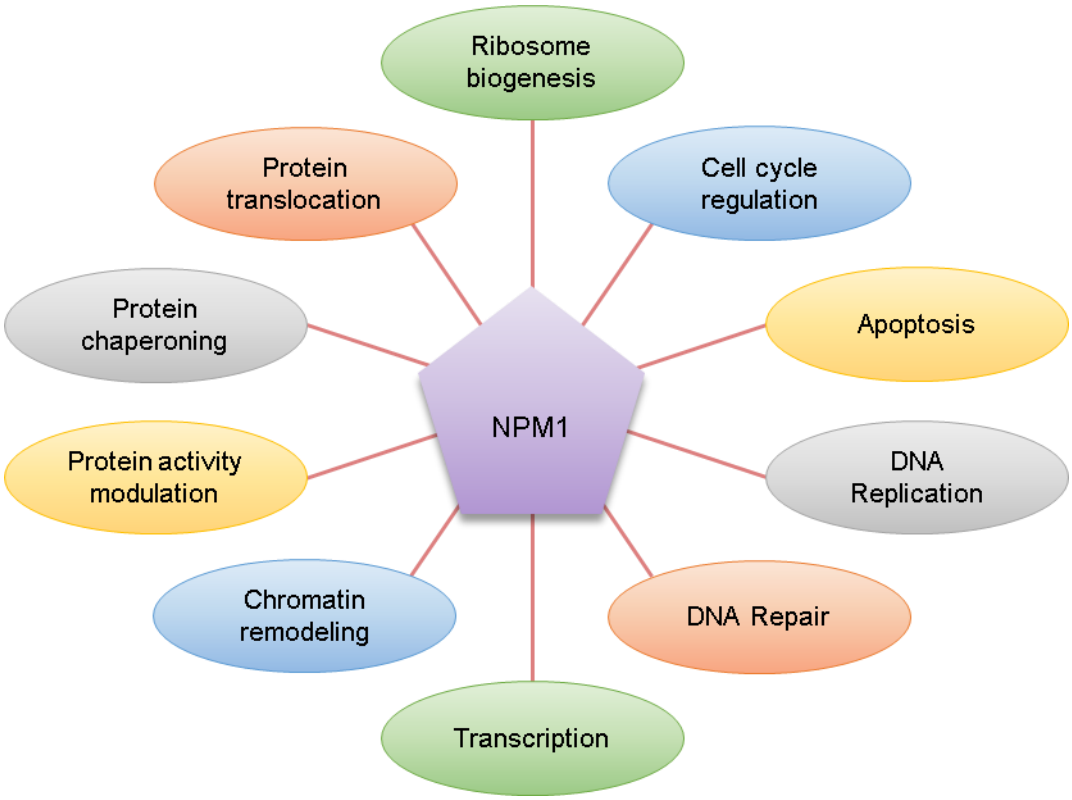



**Figure 1.13. Functional diversity of NPM1:** The figure shows the major molecular and cellular processes where the role of NPM1 has been implicated. Protein chaperoning includes both its molecular as well as histone chaperone functions. See the text (Section 1.4.1.3) for details.

### *1.4.1.4. NPM1 and Cancer*

NPM1 is essential for cellular survival and normal organismal development as it plays critical roles in several important cellular processes. However, the physiological level of NPM1 protein needs to be critically maintained since both absence as well as overexpression can result in detrimental consequences. The functions of NPM1, hence, can be both tumor suppressive as well as oncogenic, the latter being more predominant. As evident from the discussion in the previous sections, an overexpression of NPM1 can promote cell division, growth, and survival through multiple ways such as an increase in

ribosome biogenesis, stimulation of DNA replication and transcription, facilitating DNA repair, suppression of apoptosis, as well as modulation of the functional properties of tumor suppressors and oncogenes, directing the cell towards cancer (Grisendi et al. 2006; Box et al. 2016). In agreement with these functions of NPM1, it is indeed found to be overexpressed in multiple types of cancers (elaborated in Chapter 3, Section 3.1) and mutated to gain more oncogenic properties, in certain blood cancers. In non-cancerous cells such as fibroblasts, overexpression of NPM1 was found to result in the malignant transformation of mouse fibroblast NIH3T3 cells (Kondo et al. 1997) as well as induce p53-dependent premature senescence in human fibroblasts cells (Colombo et al. 2002). In such instances, NPM1 behaves like oncogenes such as Ras. Although it is not clearly understood whether the upregulation of NPM1 in tumor cells is a cause or a consequence of cancer which appears to vary in different contexts, there is no doubt that overexpression of NPM1 presents several advantages to the sustenance of the tumor cells by promoting higher growth rate and meeting an increased demand for ribosome biogenesis (Ruggero and Pandolfi 2003). It can also protect cancer cells from cytotoxicity induced through drugs such as platinum-containing compounds (Malfatti et al. 2019) and adriamycin (Yang et al. 2007). NPM1 overexpression and its occasional delocalization in cancer are often found to correlate well with their specific clinical prognostic features as seen in liver cancer (Yun et al. 2007; Yang et al. 2016b), oral cancer (Shandilya et al. 2009; Peng et al. 2019), bladder cancer (Tsui et al. 2004), breast cancer (Skaar et al. 1998), ovarian cancer (Fan et al. 2017), lung cancer (He et al. 2016), and thyroid cancer (Pianta et al. 2011). Sometimes autoantibodies to NPM1 have also been detected in the serum of patients suffering from cancer such as liver cancer (Liu et al. 2015a; Wang et al. 2017), lung cancer (Dai et al. 2016b), and prostate cancer (Dai et al. 2014; Dai et al. 2016a). In addition, the association of NPM1 with metastasis and radiation-resistance has been observed in some instances (Chen et al. 2013; Qu et al. 2016; Peng et al. 2017). In all such cases, NPM1 overexpression can serve as a diagnostic marker for cancer. Its upregulation can be caused by various pathways, one of them being its induction upon mitogenic signals (Feuerstein and Mond 1987; Feuerstein et al. 1988a; Feuerstein et al. 1988b).

Genetic alterations of NPM1 have not been reported in solid tumors including lung, hepatocellular, breast, colorectal, and gastric carcinomas (Jeong et al. 2007). However, the *NPM1* gene is frequently mutated in different ways such as frameshift mutations, translocations, and deletions, in blood cancers (lymphomas and leukemias) (Falini et al.

2006b; Naoe et al. 2006; Falini et al. 2007; Rau and Brown 2009; Kunchala et al. 2018) namely acute promyelocytic leukemia (APL), anaplastic large cell lymphoma (ALCL) and in premalignant myelodysplastic syndrome (MDS) (Grisendi et al. 2006). Among the translocations that have been reported, the N-terminal region of NPM1 is found to fuse with different partner genes such as anaplastic lymphoma kinase (ALK), retinoic acid receptor alpha (RAR-α), and myeloid leukemia factor 1 (MLF1) in ALCL, APL, and MDS respectively (Falini et al. 2007). Such fusions provide additional and/or enhanced functional properties to the chimeric proteins which contribute to oncogenesis. For example, the chimeric NPM1-ALK protein is able to oligomerize and activate the kinase signaling cascade thus exhibiting its oncogenic potential (Morris et al. 1994; Bischof et al. 1997). NPM1-ALK was also found to manifest its oncogenic properties by epigenetically silencing the tumor suppressor gene *STAT5A* (Zhang et al. 2007). Loss of an allele of *NPM1* has been reported in de novo and therapy-related MDS (Grisendi et al. 2006). Further, the C-terminal region of NPM1 (exon 12) is often prone to mutations in at least one-third of the adult AML cases, which result in the disruption of its NoLS and/or gain of an additional NES motif in the C-terminus causing an aberrant cytoplasmic localization of NPM1 (denoted as NPM1c+) (Falini et al. 2005; Falini et al. 2009). The NPM1c+ mutant has several altered properties compared to wild-type NPM1. It can aberrantly delocalize proteins from the nuclear and nucleolar compartments to the cytoplasm, thereby inhibiting their normal functions in the nucleus as seen in the case of Fbw7γ, the E3 ubiquitin ligase of various oncoproteins, resulting in increased levels of the oncoproteins such as c-myc (Li and Hann 2009). Recently, germline mutations of *NPM1* have been identified in patients with dyskeratosis congenita which is a rare and progressive bone marrow failure syndrome. In this context, wild-type NPM1 was found to regulate the modification 2'-O-methylation on rRNA by directly binding to small nucleolar RNA and decreasing their interaction with the rRNA methyltransferase FBL, thereby modulating translation via epitranscriptomic regulation. These mutations inhibit this function of NPM1 that is important for cellular growth, differentiation and hematopoietic stem cell maintenance (Nachmani et al. 2019; Zhou and Muller-Tidow 2019).

Since the overexpression of NPM1 in various cancers is associated with the proliferation of tumor cells, targeting the functional form of NPM1 in cancer cells through RNAi or small molecule inhibitors is emerging as a promising therapeutic approach to treat cancer. NPM1-targeting molecules may work by suppressing specific functions of NPM1, interfere

with its subcellular localization, with its oligomerization properties or drive its degradation (Di Matteo et al. 2016). The compound NSC34884 disrupts NPM1 oligomerization by binding to a specific hydrophobic pocket required for oligomerization, which could induce p53-mediated apoptosis in cancer cells (Qi et al. 2008). Other known molecules target NPM1 in the following ways: the alkaloid Avrainvillamide alkylates Cys275 of NPM1 and induces p53 activity, the porphyrin TmPyP4 induces nucleolar disruption by binding to C-terminal region of NPM1 and G-quadruplexes, ATRA/ATO promotes degradation of NPM1c+ by oxidation of Cys288 of NPM1, Deguelin and ECGT promote degradation of NPM1c+ by an unknown mechanism, N6L binds to NPM1 and interferes with some of its protein-protein associations, and YTR107 binds to N-terminal region of NPM1 and promotes its monomerization (Di Matteo et al. 2016). High-affinity binding of NPM1 with peptides such as Rev has been shown to disrupt its transactivation ability with a concomitant reduction in the tumorigenic properties of the cells (Chan et al. 2005). The NPM1-binding peptide CIGB-300 binds to Ser125 and prevents the phosphorylation of NPM1 by CK2 leading to nucleolar disruption and apoptosis induction in cancer cells (Perera et al. 2009). RNA aptamers are another kind of molecule that can bind to the central region of NPM1 and disrupt its oligomerization causing delocalization of NPM1 from the nucleolus and sensitizing the cells to apoptosis (Jian et al. 2009). Besides targeting NPM1 itself, some molecules have also been developed to target the interaction of NPM1 with other proteins such as APE1, which were found to display anti-tumor properties (Poletto et al. 2016).

As mentioned previously, in some instances NPM1 has been implicated as a tumor suppressor. For example, in blood cells, NPM1 was suggested to act as a haploinsufficient tumor suppressor of myeloid and lymphoid malignancies (Sportoletti et al. 2008). Loss of NPM1 in mice or presence of NPM1c+ mutant in AML were found to affect the nucleolar localization of the tumor suppressor ARF and destabilize it which prevents the activation of p53 upon oncogenic stimuli and results in oncogenic transformation (Colombo et al. 2005; den Besten et al. 2005; Sharpless 2005; Colombo et al. 2006). NPM1 heterozygosity in mice has been found to result in genomic instability due to aneuploidy, increased centrosome numbers and DNA damage checkpoint activations which at a later stage enhance oncogenic transformation. In concordance with this observation, *Npm1*$^{+/-}$ mice often develop hematological malignancies that resemble myelodysplasia in humans (Grisendi et al. 2005). Besides mutations, downregulation of NPM1 expression has also

been reported in specific cohorts of certain cancers such as gastric cancer (Leal et al. 2014) and breast cancer (Karhemo et al. 2011), where NPM1 was less associated with metastasis. NPM1 can associate and cooperate with factors that have tumor-suppressive functions such as HLJ1 (Chang et al. 2010). For instance, its interaction with HDAC2 and lncRNA EPB41L4A-AS1 in cancer cells have a repressive effect on the Warburg effect (Liao et al. 2019). However, the upregulation of NPM1 per se was found to promote aerobic glycolysis and repress fructose-1,6-bisphosphatase 1 (FBP1) in pancreatic cancer cells (Zhu et al. 2015). From these studies, we get the idea that NPM1 is not a classical tumor suppressor gene, but behaves as a tumor suppressor only under certain cellular contexts since even with the loss of an NPM1 allele which promotes tumorigenesis, NPM1 itself does not repress cell cycle or induce growth arrest in response to DNA damage. The proper regulation of NPM1 levels and cellular localization often determines its role as a tumor suppressor.

### *1.4.2. Nucleoplasmin (NPM2)*

NPM2 is the mammalian ortholog of *Xenopus* Nucleoplasmin which is the founding member of the Nucleoplasmin family of histone chaperones. It was discovered as an acidic, thermostable, multi-subunit complex in *Xenopus laevis* oocytes capable of binding to histones and transferring them to DNA. This property was demonstrated *in vitro* by the generation of supercoils in a relaxed DNA template, a reaction that does not require cofactors like Mg2+ and ATP and is inhibited at higher NaCl concentrations (Laskey et al. 1978). This protein in *Xenopus* was termed as Nucleoplasmin indicating its occurrence in the soluble fraction of the nucleus in different vertebrate cell types (Earnshaw et al. 1980). Much of its molecular functions have been elucidated in *Xenopus*. However, the mouse and zebrafish orthologs have also been studied and characterized. A few studies have been done using the orthologs in cow, some species of cyprinid fish, and humans.

The human *NPM2* gene is located in the short arm of chromosome 8 in the region 8p21.3 according to the latest GRCh38.p13 assembly of the human genome. It is about 13.6 kb long and has 11 exons. It gives rise to 3 transcript variants through alternative splicing. Two transcript variants code for the major protein isoform whereas the third variant codes for a truncated form of the protein having the oligomerization domain and only a part of

the nuclear localization sequence. The mouse *Npm2* gene is of similar size and is reported to give rise to a single transcript variant.

Human NPM2 shares about 63% sequence identity with mouse Npm2 and about 48% identity with the orthologs in *Xenopus* sp. Figure 1.14 shows an alignment of the protein sequences of human, mouse and *Xenopus* Nucleoplasmin proteins. In comparison to *Xenopus* Nucleoplasmin, the human and mouse orthologs lack one of the three acidic stretches, which is the first acidic tract present within the N-terminal core domain (denoted as A1) (Figures 1.9 and 1.14). This structural feature has been correlated with the reduced efficiency of histone chaperone activity of human NPM2 in comparison to *Xenopus* Nucleoplasmin since the acidic tracts are believed to play a significant role in the binding of Nucleoplasmin to histones (Platonova et al. 2011). Mechanistically, it has been proposed that two pentamers of Nucleoplasmin can form decamers under physiological conditions. The decameric form is stabilized by its interaction with H2A-H2B dimers through its lateral surface. This complex can reversibly bind to free or chaperone-bound H3-H4 tetramer, thereby forming Nucleoplasmin-histone octamer assemblies which can potentially mediate nucleosome assembly (Dutta et al. 2001; Platonova et al. 2011). The interaction of Nucleoplasmin with histones is modulated by dynamic intramolecular competitions among the largest acidic stretch (A2), the C-terminal basic stretch, and the basic nuclear localization signal present in the Nucleoplasmin tail domain (Warren et al. 2017).


**Figure 1.14. Comparison of protein sequences of Nucleoplasmin orthologs:** Multiple sequence alignment of Nucleoplasmin orthologs in human (NPM2), mouse (Npm2), and *Xenopus* (NP) using ClustalW2 sequence alignment tool. The region enclosed in the green box denotes the oligomerization or core domain characteristic of the members of Nucleoplasmin family. The residues highlighted in yellow and enclosed in red boxes represent the three acidic tracts (denoted

as A1, A2, and A3), those highlighted in turquoise and bright green represent the conserved AKD/EE and GSGP loops respectively. Note the highly reduced length of acidic tract A1 in human and mouse Nucleoplasmin compared to that in *Xenopus.* Residues highlighted in grey are basic in nature. The residues enclosed in the purple box constitute the characteristic bipartite nuclear localization signal while those enclosed in the blue box constitute the C-terminal basic region.

NPM2 is mostly localized in the nucleus of the cell. While some studies have reported the exclusion of NPM2 from the nucleolus (Burns et al. 2003), other studies have suggested that it is an important component of the nucleolus (Ogushi et al. 2017) or nucleolus-like bodies (Inoue and Aoki 2010) in oocytes, playing a role in nucleolar organization, which would be discussed in the subsequent sections. The *Drosophila* NLP has been found to be localized to the centromere through its binding with the protein HMR, in a domain distinct from the one having dCENP-A (Anselm et al. 2018).

### *1.4.2.1. Post-translational modifications of Nucleoplasmin/NPM2*

The histone binding activity of Nucleoplasmin is modulated by its PTMs. Nucleoplasmin was found to be majorly phosphorylated in *Xenopus* eggs from where it was first isolated. Phosphorylation is higher in the laid eggs compared to the oocytes. It occurs in multiple sites throughout the protein and has been found to increase the activity of Nucleoplasmin (Cotten et al. 1986; Leno et al. 1996; Banuelos et al. 2007). One of the kinases phosphorylating Nucleoplasmin is Casein Kinase II (Taylor et al. 1987; Vancurova et al. 1995). Human NPM2 gets phosphorylated more efficiently by mitotic kinases (Okuwaki et al. 2012). However, the specific kinases phosphorylating NPM2 have not been identified yet. Nonetheless, the protein harbors consensus sites for some cell cycle-dependent kinases (Finn et al. 2012).

Among other PTMs that have been reported, Nucleoplasmin has been shown to get symmetrically di-methylated at Arg187 by PRMT5-Mep50 complex, which modulates its activity towards histones, indicative of a regulatory role of NPM2 *in vivo*. This methylated NPM2 is physiologically detected late in oogenesis and is believed to contribute to the development of pluripotency (Wilczek et al. 2011). Nucleoplasmin has also been shown to be glutamylated in addition to the PTMs phosphorylation and arginine methylation. These PTMs are developmentally regulated and they greatly impact the conformation of Nucleoplasmin causing it to either deposit the histones onto DNA (mediated by its C-

terminal tail phosphorylation, PRMT-5-mediated arginine methylation, and TTLL-4-mediated glutamylation present in oocyte Nucleoplasmin), or sequester them (mediated by hyper-phosphorylated egg Nucleoplasmin) (Vitale et al. 2007; Onikubo et al. 2015).

### *1.4.2.2. Functions of Nucleoplasmin/NPM2*

Nucleoplasmin and its orthologs have been proposed to play a role in regulating the histone dynamics in different cellular processes by virtue of their histone-binding properties (Figure 1.15). Some of its functions have been implicated in organismal development as discussed further.

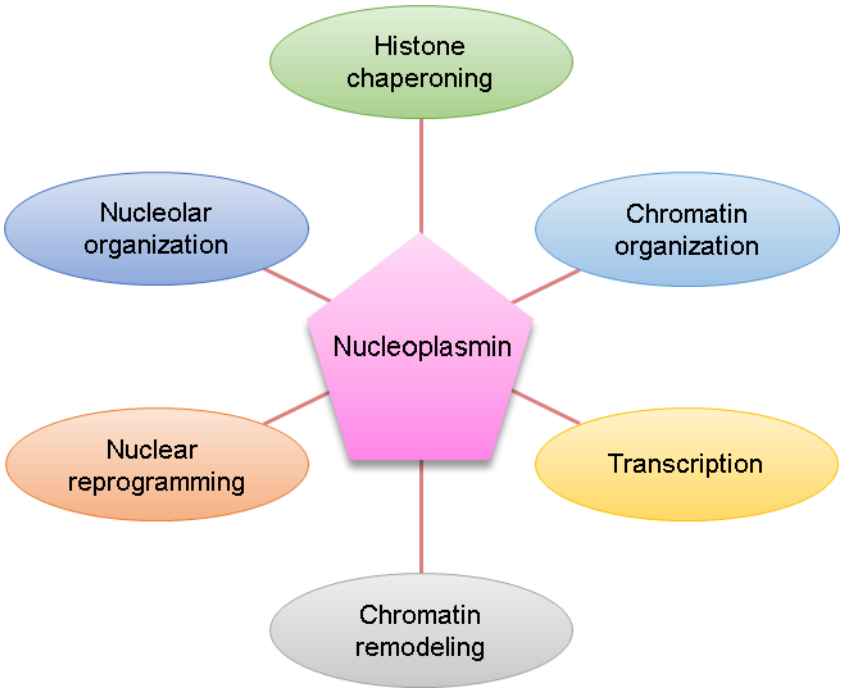



**Figure 1.15: Functional aspects of Nucleoplasmin:** The figure shows the major molecular and cellular processes where the role of Nucleoplasmin and its orthologs has been implicated. See the subsequent sections for details.

#### *1.4.2.2.1. Chromatin remodeling: Implications in oocyte development and fertilization*

The role of Nucleoplasmin has been very well studied in the context of sperm chromatin remodeling post-fertilization. In studies conducted using *Xenopus laevis* and mouse models, it was found that after fertilization of the egg by a sperm, Nucleoplasmin mediated the decondensation of the sperm chromatin in the egg extracts, by the removal of the sperm-

specific basic proteins (SSBPs) or protamines, and deposition of maternal H2A-H2B onto the sperm DNA (Philpott et al. 1991; Inoue et al. 2011). This process was also demonstrated *in vitro* (Huo et al. 2018). At the organismal level, female mice null for *Npm2* (*Npm2*$^{-/-}$) show fertility defects due to failed preimplantation embryo development (Burns et al. 2003). In the oocytes of these animals, in the absence of Npm2 or nucleolus/nucleolus-like bodies of which Npm2 is an important component (Inoue and Aoki 2010; Ogushi et al. 2017), sperm chromatin decondensation is retarded (Burns et al. 2003), which is restored by the microinjection of *Npm2* mRNA in the oocytes (Inoue et al. 2011). Artificial expression of Npm2 in enucleolated oocytes was found to not only reconstitute the nucleolar structure but also rescue the first mitotic division and full-term development of the early embryo (Ogushi et al. 2017).

#### *1.4.2.2.2. Chromatin organization: Implications in nuclear reprogramming*

NPM2 plays a role in the organization of the chromatin in the nucleus. This was evident in the mice null for *Npm2* where abnormalities were observed in the oocyte and early embryonic nuclei (Burns et al. 2003). The abnormalities included an absence of coalesced nucleolar structures and loss of heterochromatin and deacetylated histone H3 which are generally found around the nucleoli in oocytes and early embryos (Burns et al. 2003; De La Fuente et al. 2004). Nucleoplasmin was found to be capable of decondensing the chromatin in undifferentiated mouse cells into a more open structure characterized by changes in multiple histone H3 phosphorylation, increase in H3K14ac, and release of heterochromatin proteins HP1β and TIF1β from the nuclei, thereby playing a role in nuclear reprogramming (Tamada et al. 2006). A similar outcome was observed upon overexpression of NPM2 in HEK-293 cells with respect to epigenetic changes (histone acetylation and methylation) and expression of pluripotency genes, indicating its potential role in somatic cell reprogramming (Sylvestre et al. 2010).

#### *1.4.2.2.3. Regulation of transcription*

Nucleoplasmin and its orthologs have been implicated in cellular transcription under specific contexts. At the organismal level, in mouse and zebrafish models, absence or knockdown of Npm2 was found to result in the reduced expression of early zygotic genes

(Burns et al. 2003; Bouleau et al. 2014). In a different study, it was found that upon pretreating mouse nuclei microinjected into *Xenopus laevis* oocytes, with Nucleoplasmin, there was an enhancement in the expression of certain oocyte-specific genes in the microinjected nuclei (Tamada et al. 2006). Further, overexpression of NPM2 in HEK-293 cells was found to result in the transcriptional upregulation of certain pluripotency genes such as OCT4 (Sylvestre et al. 2010). Fairly recently it was shown that human phospho-mimic NPM2 could enhance induced pluripotent stem cell generation by regulating the expression of genes involved in naïve stem cell stage, in combination with oocyte-enriched histone variants TH2A and TH2B after lentiviral transduction in dermal fibroblasts (Huynh et al. 2016). Such studies indicate that Nucleoplasmin plays a role in context-specific regulation of transcription. However, the mechanism of Nucleoplasmin/NPM2-mediated transcription regulation is unclear. In a very early study, it was demonstrated that *Xenopus* Nucleoplasmin could stimulate the binding of transcription factors like GAL4 and USF onto nucleosomes, leading to factor-induced nucleosome disassembly *in vitro* (Chen et al. 1994) (Figure 1.16). It was proposed that the occurrence of such a mechanism in the cell could potentially result in the activation of transcription from the chromatin *in vivo*.

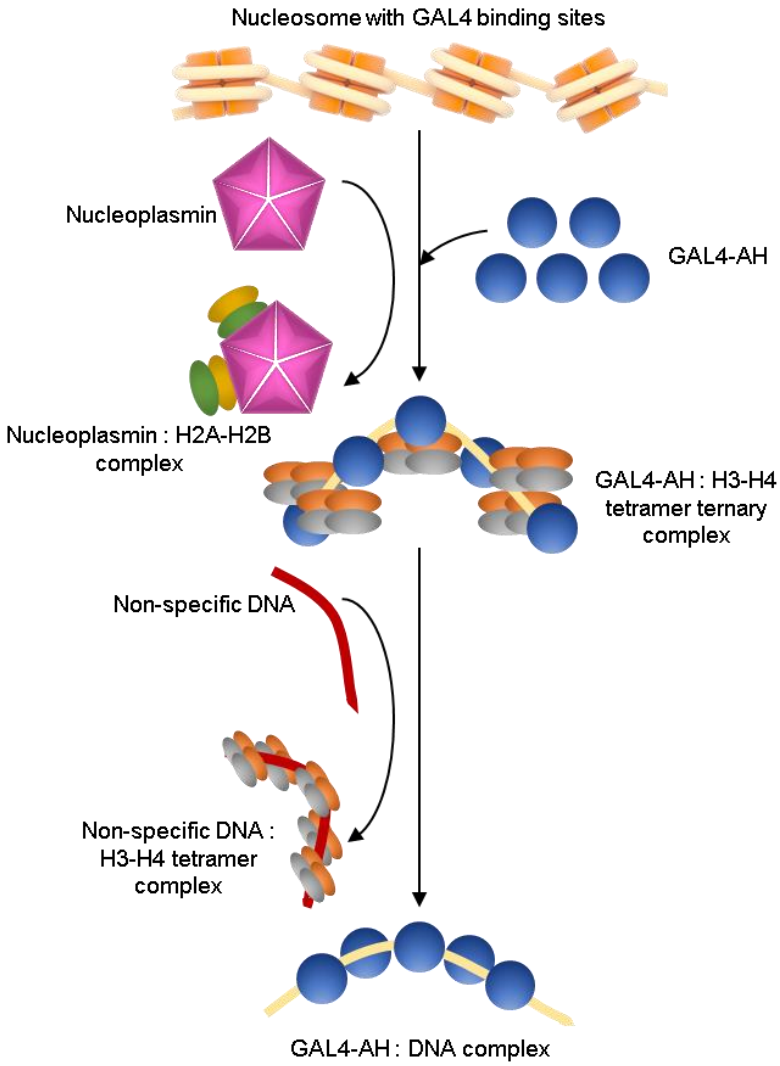



**Figure 1.16. Role of Nucleoplasmin in stimulating transcription factor binding to the nucleosomal template:** The model depicts that in this *in vitro* system, the binding of transcription factor GAL4-AH to nucleosome cores is stimulated in the presence of Nucleoplasmin. GAL4-AH, on the other hand, stimulates the binding of Nucleoplasmin to histone H2A-H2B dimers. This combined reaction of GAL4-AH binding and the displacement of H2A-H2B dimers results in the formation of a Nucleoplasmin : H2A-H2B dimer complex and a GAL4-AH : H3-H4 tetramer : DNA ternary complex. The H3-H4 tetramer in this ternary complex is weakly bound and is readily competed out by non-specific DNA. This second step completes the nucleosome disassembly, resulting in the formation of a GAL4-AH : DNA complex, a mechanism that might lead to activation of gene transcription from the chromatin template. In cells, the removal of H3-H4 tetramers may also be mediated by proteins which specifically interact with these histones. The figure has been adapted from (Chen et al. 1994).

### *1.4.2.3. NPM2 expression and its regulation during development and in disease*

Nucleoplasmin has been proposed to be essentially a maternally expressed gene, having a predominantly ovary-specific expression as observed in *Xenopus*, *Bos taurus*, mouse, and

fish (Burns et al. 2003; Frehlick et al. 2006; Lingenfelter et al. 2011; Bouleau et al. 2014; Onikubo et al. 2015; Cheung et al. 2018). Within the ovary, it is expressed specifically in the oocytes (Burns et al. 2003; Bouleau et al. 2014). In the context of development, Nucleoplasmin mRNA is detected during oogenesis, reaching the highest levels at the pre-antral follicle stage and then rapidly declining at fertilization and vanishing during embryogenesis. The Nucleoplasmin protein persists throughout early embryogenesis and exhibits increased phosphorylation during oocyte maturation and after mid-blastula transition (Cotten et al. 1986; Burglin et al. 1987; Litvin and King 1988; Vitale et al. 2007; Sanchez et al. 2009). Eventually, the cellular levels of Nucleoplasmin decrease in the course of development which has been linked to the regulation of developmental nuclear size scaling since Nucleoplasmin was recently shown to be a limiting component in the scaling of nuclear size with cytoplasmic volume (Chen et al. 2019). Together, these observations suggest Nucleoplasmin as an important factor in ovarian development and early embryogenesis.

Few discrete studies have addressed the regulation of Nucleoplasmin's tissue-specific expression. Mouse *Npm2* promoter harbors certain cis-acting elements namely the Enhancer-box (E-Box) element and the Nobox element (NBE) which were found to be required for its oocyte-specific expression (Tsunemoto et al. 2008). The bovine NPM2 expression has been shown to be regulated in part by miR-181a, indicating post-transcriptional regulation of NPM2 expression as well (Lingenfelter et al. 2011).

In the context of diseases such as cancer, not many studies have been performed to date addressing the expression of NPM2. Reports from databases such as the Human Protein Atlas, GeneCards, and NCBI, have suggested NPM2's expression in certain tissues such as the brain (cerebellum, hippocampus), respiratory epithelia, endocrine glands, melanocytes, and Leydig cells, besides oocytes. Most of the other cells are NPM2 negative under the normal condition. In cancer, higher NPM2 expression has been reported in liver and ovarian cancers compared to expression levels in adjacent normal tissues, as per data from the Human Protein Atlas and GeneCards databases. On the contrary, reduced expression of NPM2 due to promoter DNA methylation has been reported in melanoma and leukemia in comparison to its expression in normal cells (Kroeger et al. 2008; Koga et al. 2009; Fujiwara et al. 2018). Reduced expression of NPM2 due to gene deletion has been observed in colorectal cancer (Xi et al. 2016).

### *1.4.3. NPM3*

NPM3 is the third member of the Nucleoplasmin family which was first identified and characterized in *Xenopus laevis* as NO29 (Zirwes et al. 1997). The mouse and human orthologs of NO29 were subsequently discovered and named as Npm3 and NPM3 respectively (MacArthur and Shackleford 1997; Shackleford et al. 2001).

The human *NPM3* gene is located in chromosome 10 in the region 10q24.32 and consists of 6 exons according to the current GRCh38.p13 assembly of the human genome. To date, no multiple splice variants have been identified or predicted for *NPM3*. As a result, there is no other isoform for this protein. In comparison to its paralogs NPM1 and NPM2, NPM3 lacks much of the C-terminal region of the proteins. As shown in Figure 1.9, the third acidic stretch is absent in NPM3 which is present in the other family members. Consequently, NPM3 has the ability to bind to histones but lacks an intrinsic histone chaperone activity. However, it can enhance activator dependent transcription (Gadad et al. 2010). It cannot homo-oligomerize like NPM1 and NPM2 but can hetero-oligomerize with NPM1 (Okuwaki et al. 2012). NPM3 has a ubiquitous expression pattern and is localized in the nucleus of the cell including the nucleolus (Shackleford et al. 2001). It has multiple potential sites for phosphorylation by several kinases (Shackleford et al. 2001).

NPM3 has been associated with functions such as inhibition of ribosome biogenesis through its interactions with NPM1 (Huang et al. 2005), enhancement of cell proliferation and inhibition of differentiation in embryonic stem cells (Motoi et al. 2008), and regulation of transcription (Gadad et al. 2010; Okuwaki et al. 2012). NPM3 could also have a role in mammalian spermiogenesis by virtue of its interaction with transition protein TP2 (Pradeepa et al. 2009). Reports from databases and a few studies have suggested an association (upregulation or gene deletion) of NPM3 with cancer (Sakhinia et al. 2007; Hallor et al. 2009; Kang et al. 2010; Nishio et al. 2011). However, in-depth research in this aspect has not been conducted to date. NPM3 expression has been reported to be regulated by the transcription factor Sp1 in HeLa cells (Oleaga et al. 2012).

## *1.5. Aims and scope of the study*

Histone chaperones have emerged as important players in various chromatin templated processes such as DNA replication, repair, and transcription. Their critical roles in regulating the histone traffic in the cell are manifested when a loss or gain of their functions results in various developmental and disease conditions including cancer (Burgess and Zhang 2013; Gurard-Levin et al. 2014). Nucleophosmin (NPM1), which belongs to the family of the classical histone chaperone Nucleoplasmin, is involved in multiple cellular processes such as ribosome biogenesis, cell cycle regulation, apoptosis, DNA replication, DNA repair, and transcription, among others (Okuwaki 2008). NPM1 exists in a variety of post-translationally modified forms which modulate its properties and functions, and are implicated in the regulation of different cellular processes. In the context of RNA Pol II-driven chromatin transcription, acetylation of NPM1 was found to increase its histone chaperone activity and transcription activation potential *in vitro* (Swaminathan et al. 2005). The enhanced binding affinity of acetylated NPM1 (AcNPM1) towards acetylated histones suggested that it could potentially induce nucleosome disassembly required for transcription through the chromatin template (Swaminathan et al. 2005). The cellular localization of AcNPM1 is nuclear as opposed to the predominant nucleolar localization of NPM1 (Shandilya et al. 2009). The colocalization of AcNPM1 with RNA Pol II foci in the nucleoplasm provided the first indication of the role of acetylation of NPM1 in regulating RNA Pol II-driven transcription in the cell (Shandilya et al. 2009). However, the genome-wide localization of AcNPM1 and its gene targets were largely unknown. Further, the biochemical mechanisms of NPM1/AcNPM1-mediated regulation of transcription were not clearly defined. Oligomerization of NPM1 which is mediated through its N-terminal core domain was found to play a role in its histone chaperone activity as well as activation of chromatin transcription (Swaminathan et al. 2005). In addition, NPM1/AcNPM1 might also act as a recruiter of RNA Pol II transcription machinery through potential interactions with RNA Pol II subunits, transcription factors, and transcriptional coactivators (Senapati P, Ph.D. thesis, 2014). However, the experimental validation of these hypotheses and the delineation of the stage(s) of the transcription process where NPM1/AcNPM1 is involved, are yet to be determined.

NPM1 is an essential protein for the cell and is required for the normal physiological development of the organism. Its absence or marked reduction has been shown to result in embryonic lethality or developmental disorders (Grisendi et al. 2005; Sportoletti et al.

2008). On the other hand, its overexpression has been observed in tumor cells of various histological origins, which often correlate well with cancer prognosis (Grisendi et al. 2006). In oral cancer, the levels of NPM1 and its acetylated form were found to be elevated in the tumor tissues compared to matched normal samples, which also correlated positively with increasing grades of the tumors (Shandilya et al. 2009). Similar observations made in other types of cancers suggest that NPM1 overexpression could be a cause as well as a consequence of cancer. In support of the former hypothesis, it has been seen in several cancers that targeting NPM1 levels or its functions has resulted in the reduction of cancer progression. However, downregulation of NPM1 expression has also been reported in a few cancer types in specific patient cohorts (Karhemo et al. 2011; Leal et al. 2014). These context-dependent tumor-suppressive and tumor-promoting functions of NPM1 indicate that its expression could be regulated through multiple mechanisms under specific conditions which would determine the physiological or pathophysiological state of the cell. Hence, it is important to understand the mechanisms of regulation of NPM1 expression having implications in the manifestation of diseases such as cancer.

While NPM1 has been an actively researched protein due to its multiple functions and association with cancer manifestation, its homolog NPM2 has been significantly less studied. An important reason for this is the tissue-specific expression pattern of NPM2 (in non-human model organisms) in contrast to the ubiquitous expression of NPM1, which poses technical restrictions for studying the biological functions of NPM2. Its functions, tissue distribution in humans and any association with cancer are largely unknown. NPM2 is the mammalian ortholog of *Xenopus* Nucleoplasmin, which was the first histone chaperone to be described (Laskey et al. 1978). Its role in histone storage and sperm chromatin remodeling in oocytes and eggs is well documented. Having a similar structural and domain organization as that of NPM1, NPM2 can be expected to perform molecular functions analogous to NPM1, besides being a classical histone chaperone. Like NPM1, Nucleoplasmin can also potentially induce nucleosome disassembly, implying its role in the activation of chromatin transcription (Chen et al. 1994). Indeed, artificial downregulation or overexpression of Nucleoplasmin or its orthologs was found to result in the altered expression profile of specific genes *in vivo* (Burns et al. 2003; Bouleau et al. 2014; Huynh et al. 2016). However, it is unclear if NPM2 can directly activate transcription from the chromatin template. Further, the molecular determinants and the cellular contexts which would be required for the functioning of NPM2 as a transcription regulator, are also

unknown. For example, the histone chaperone activity mediated by the oligomerization of the Nucleoplasmin homolog NPM1 appears to be important for its transcription activation potential (Swaminathan et al. 2005). On the other hand, the other Nucleoplasmin family member NPM3, can enhance activator-dependent transcription inspite of lacking intrinsic histone chaperone activity (Gadad et al. 2010). It is not known if the histone chaperone activity and transcription activation property of NPM2 are linked or functionally distinct. Understanding these aspects of NPM2 could provide interesting insights into the functional evolution of the Nucleoplasmin homologs, especially in the context of regulation of transcription.

Based on this background, the objectives of the present study were developed as follows:

1. Regulation of NPM1 expression: Implications in cancer.
    a. Role of transcription factor c-fos/AP-1 and mutant R175H p53 in the regulation of NPM1 expression in cancer.
    b. Role of transcription factor YY1 in the regulation of NPM1 expression in cancer.
2. Role of NPM1 in the regulation of RNA Polymerase II-driven transcription: Implications in oral tumorigenesis.
    a. Identification of genome-wide targets of AcNPM1 in cancer.
    b. Effect of NPM1 downregulation on the transcriptional profile of oral cancer cells and its implications on tumor progression.
3. Functional characterization of mammalian Nucleoplasmin (NPM2) as a potential regulator of transcription.
    a. Analysis of transcription activation potential of NPM2.
    b. Role of Aurora Kinase-mediated phosphorylation of NPM2 on its histone chaperone activity.
    c. Expression analysis of NPM2 in mouse tissues and human cancer cells.

# *Chapter 2: Materials and Methods*

## *2.1. General Methods*

### *2.1.1. Preparation of E. coli competent cells*

Different strains of *E. coli* were used to prepare competent cells meant for purification of DNA or different kinds of proteins. The procedures for preparing the different competent cells are majorly similar with minor variations in the buffer compositions and protocol as described below.

Frozen glycerol stocks (stored at −80°C) of the different strains of *E. coli,* such as DH5α and XL10-Gold (used for DNA purification) or BL21 (DE3) and Rosetta (used for protein purification), were revived by streaking them on a freshly prepared plate of LB (10 g/l tryptone or casein enzymic hydrolysate, 5 g/l yeast extract and 10 g/l NaCl) (HiMedia, Mumbai, India) with agar (1.5% w/v) (HiMedia), and incubating at 37°C overnight. For culturing Rosetta cells, the media in the culture plate contained an additional 34 µg/ml of chloramphenicol antibiotic (HiMedia). For culturing XL-10 Gold cells, the media in the culture plate contained 34 µg/ml of chloramphenicol and 10 µg/ml of tetracycline (HiMedia).

A single colony from the streaked plate was inoculated in 5 – 10 ml of LB containing 10 mM $MgCl_2$ and 10 mM $MgSO_4$. For culturing Rosetta cells, this media contained an additional 34 µg/ml of chloramphenicol antibiotic, whereas for culturing XL-10 Gold cells, the media was supplemented with 34 µg/ml of chloramphenicol and 10 µg/ml of tetracycline. This primary culture was grown overnight at 37°C, 180 rpm. 1% (v/v) of this primary culture was inoculated in 100 – 200 ml of LB medium supplemented with 10 mM $MgCl_2$ and 10 mM $MgSO_4$ (and additional 34 µg/ml of chloramphenicol antibiotic for Rosetta cells) in a large conical flask having at least three-fourths of this volume empty. The secondary culture was grown at 37°C, 160 – 180 rpm until the $O.D._{600}$ of the culture is 0.35 – 0.4. In the case of XL-10 Gold, the $O.D._{600}$ of the primary culture was measured and the calculated volume of this culture was inoculated in 50 ml of LB medium to attain an initial $O.D._{600}$ of 0.1. This secondary culture was grown at 37°C, 180 rpm until its $O.D._{600}$ comes around 1. The secondary XL-10 Gold culture was then grown at 18°C, 180 rpm for 1 h for acclimatization at a lower temperature. The calculated volume of this secondary

culture was then inoculated into 200 ml of LB to get an initial $O.D._{600}$ of 0.05. The tertiary culture was then grown overnight at 18°C, 180 rpm until its $O.D._{600}$ is about 0.45.

The secondary or tertiary cultures were kept on ice for 30 – 45 min to stop the growth of the *E. coli* cells. The cells were pelleted by centrifuging at 4°C, 4000 rpm for 10 min. The supernatant medium was discarded, and the cells were gently resuspended in chilled, filter sterilized, specific transformation buffers as mentioned below:

15 – 30 ml of transformation buffer – I (30 mM $CH_3COOK$, 100 mM KCl, 10 mM $CaCl_2$, 50 mM $MnCl_2$, 50% v/v glycerol, pH adjusted to 5.8 by 0.2% v/v acetic acid) for BL21 (DE3), Rosetta and DH5α cells.

80 ml of transformation buffer (10 mM PIPES, 15 mM $CaCl_2$, 250 mM KCl and 55 mM $MnCl_2$ for XL-10 Gold cells.

The cell suspension was incubated on ice for about 1 – 1.5 h after which the cells were collected by centrifugation at 4°C, 4000 rpm for 10 min. This pellet of cells (for BL21 (DE3), Rosetta and DH5α) was gently resuspended in 4 – 5 ml of chilled, filter sterilized, transformation buffers – II (10 mM Na-MOPS, pH 6.8, 10 mM KCl, 75 mM $CaCl_2$ and 15% v/v glycerol). In the case of XL-10 Gold cells, the pellet was again resuspended in 80 ml of the specific transformation buffer mentioned above, incubated on ice for about 1 h and collected by centrifugation at 4°C, 4000 rpm for 10 min. This final pellet of XL-10 Gold was resuspended in 10 ml of its specific transformation buffer containing 9% v/v of sterilized DMSO.

The cell suspension was aliquoted as 50 – 100 µl aliquots into pre-chilled microfuge tubes, snap-frozen in liquid $N_2$ and stored at −80°C for future use. The above procedures generally yielded a transformation efficiency of $10^4$ – $10^5$ transformants per µg of DNA transformed using a single aliquot of competent cells.

### *2.1.2. Transformation of E. coli competent cells*

Frozen *E. coli* competent cells (50 – 100 µl aliquot) were thawed on ice. 30 – 200 ng of purified DNA was added into the cell suspension. The cells were then incubated with the added DNA, on ice for 30 min followed by a brief heat shock of 1.5 min at 42°C. The cells were immediately cooled on ice for 5 min. About 1 ml of LB was added to the cell

suspension and incubated at 37°C, 180 rpm for 45 min – 1 h for recovery. The cells were then collected by centrifugation at 6000 rpm, RT for 3 min and plated on LB agar plates having the appropriate selection marker. The culture plates were incubated at 37°C overnight for the growth of the colonies.

### *2.1.3. DNA purification*

Plasmid DNA needed for different experiments such as bacterial transformation and mammalian cell transfection were purified from transformed *E. coli* DH5α or XL-10 Gold cells by alkaline lysis method using commercially available plasmid prep kits such as GenElute Plasmid Miniprep Kit (Sigma, St. Louis, MO, USA), QIAGEN Plasmid Mini (QIAGEN, Hilden, Germany) and NucleoSpin Plasmid Miniprep kit (Macherey Nagel, Düren, Germany) as per manufacturers' protocols. Plasmids that were needed for large scale transfections in mammalian cells and for *in vitro* nucleosome and chromatin assembly techniques were purified on a larger scale using the QIAGEN Plasmid Midi Kit and the EndoFree Plasmid Maxi Kit, following the protocols supplied by the manufacturer.

PCR amplicons and restriction enzyme-digested products were purified from agarose gels using the GenElute Gel Extraction Kit (Sigma) and NucleoSpin Gel and PCR Clean-up kit (Macherey Nagel), according to the manufacturers' protocols.

The purified DNA concentration and quality were assessed using the NanoDrop 1000 Spectrophotometer (Thermo Scientific, Waltham, MA, USA) by measuring the ratio of the absorbance at 260 and 280 nm ($A_{260/280}$) and 260 and 230 nm ($A_{260/230}$) of 1 µl of the DNA, and by agarose gel electrophoresis. An $A_{260/280}$ value of 1.8 – 2 and an $A_{260/230}$ value of 2 – 2.5 indicated good and acceptable quality of the DNA.

### *2.1.4. Agarose gel electrophoresis*

0.8 – 1.2% w/v agarose (Sigma) gels depending on the size and forms of the DNA needed to be resolved, were prepared in 1X TBE buffer (100 mM Tris, 100 mM boric acid and 2 mM EDTA, pH 8.3). The nucleic acid sample was mixed with 6X loading dye (0.25% w/v bromophenol blue, 0.25% w/v xylene cyanol and 30% v/v glycerol) or commercially available 5X (Bioline, Boston, MA, USA) or 6X (New England Biolabs, Ipswich, MA,

USA) loading dyes to a final concentration of 1X. The samples were then loaded in the wells of the agarose gel submerged in 1X TBE buffer in the horizontal electrophoresis apparatus and electrophoresed at 50 V – 200 V as per requirement. The gels were stained in 0.5 µg/ml of ethidium bromide (EtBr) solution post-electrophoresis, followed by de-staining in distilled water. The gels were visualized and imaged using a UV illuminator in the Gel Doc XR+ Gel Documentation System (Bio-Rad, Hercules, CA, USA), Omega Lum G Imaging System (Aplegen Inc., Pleasanton, CA, USA) or the ChemiDoc Gel Imaging System (Bio-Rad).

### *2.1.5. SDS-polyacrylamide gel electrophoresis (SDS-PAGE)*

8 – 15% SDS-polyacrylamide gels were prepared for resolving purified proteins and lysates, depending on the molecular weight of the protein to be resolved. The resolving gel (8 – 15%) was prepared by adding calculated volume of 30% w/v acrylamide solution (29% w/v acrylamide (Sigma), 1% w/v bisacrylamide (Sigma)), 0.375 M Tris-HCl, pH 8.8, 0.1% w/v sodium dodecyl sulphate (SDS), 0.1% w/v ammonium persulphate (APS) (Sigma) and 0.04% v/v N,N,N',N'-tetramethylethane-1,2-diamine (TEMED) (Sigma). The stacking gel was prepared by mixing 5% acrylamide, 0.125M Tris-HCl, pH 6.8, 0.1% SDS, 0.1% APS and 0.1% TEMED. The protein samples were mixed with 5X SDS gel loading buffer to a final concentration of 1X SDS gel loading buffer (50 mM Tris-HCl, pH 6.8, 2% w/v SDS, 0.1% w/v bromophenol blue, 10% v/v glycerol and 1% v/v beta-mercaptoethanol) (Sigma), boiled at 90°C for 5 – 10 min and loaded into the wells of the SDS-PAGE gels mounted on the vertical electrophoresis apparatus. The electrophoreses were carried out in Tris-glycine electrophoresis buffer (25 mM Tris, 250 mM glycine, pH 8.3 and 0.1% SDS) at 100 – 150 V. The protein bands were visualized by staining the gels with Coomassie Brilliant Blue (CBB) (Sigma) staining solution (50% v/v methanol, 10% v/v glacial acetic acid and 0.25% w/v Coomassie Brilliant Blue R-250) followed by removal of background stain using the de-staining solution (30% v/v methanol and 10% v/v glacial acetic acid). The gels were photographed and stored after drying in a gel drier as per requirement.

### *2.1.6. Western blotting analysis*

After performing SDS-PAGE for resolving proteins, the gel was rinsed in the transfer buffer (27 mM Tris, 192 mM glycine, 0.036% w/v SDS and 20% v/v methanol) for about 5 min preferably on a reciprocal shaker. Suitable size of a 0.45 µm pore-sized PVDF membrane (Millipore, Burlington, MA, USA) was cut to match the dimensions of the gel to be used for the western transfer process. The membrane was then activated in methanol for 2 – 5 min, followed by equilibration in the transfer buffer for about 5 min. The western transfer was set up in the Trans-Blot SD Semi-Dry Transfer Cell (Bio-Rad) using extra-thick transfer pads as per manufacturer's instructions. The transfer process was carried out at 25 V for the specific duration of time (20 – 50 min) as per the molecular weights of the proteins needed to be probed for. As an optional step, to confirm the efficient transfer of the proteins from the gel on to the membrane, the membrane was stained with Direct Blue (DB) (Sigma) solution (40% v/v ethanol, 10% v/v glacial acetic acid and 0.008% w/v Direct Blue 71) for about 5 – 10 min on a reciprocal shaker, followed by rinsing in PBS for a few min to wash off the background stain. This stained membrane was photographed for documentation if required. The membrane was then blocked with 5% w/v skimmed milk solution in PBS for at least 30 min at RT. The blocked membrane was subsequently incubated with the appropriate dilution of the primary antibody in a 1 – 2.5% w/v skimmed milk solution in PBS overnight at 4°C under gentle shaking. The blot was then washed for 5 – 10 min with 1 – 2 changes in between, with the wash buffer (0.05 – 0.1% v/v Tween-20 in PBS), on a reciprocal shaker to remove the loosely bound proteins and antibodies on the membrane. The washed blot was then incubated in the appropriate secondary antibody conjugated with Horse Radish peroxidase in the suitable dilution, in 1% w/v skimmed milk solution in PBS for 1 – 2 h at RT under mild shaking. The membrane was again washed with the wash buffer as described above. The blot was then developed using the SuperSignal West Pico PLUS Chemiluminescent Substrate (Thermo Scientific) or the Clarity Western ECL Substrate (Bio-Rad), as per manufacturers' directions. The blots were exposed to TMX films (Kodak, Rochester, NY, USA) on an analog X-ray cassette and screen (Kiran, Mumbai, India) for different time points (depending on the intensity of the chemiluminescent signal) and developed using the developer – fixer kit. The developed films were washed with water, dried, marked and photographed for documentation.

### 2.1.7. Dot-Blot and peptide competition assays

For the peptide dot-blot assays, increasing amounts of peptides were spotted equidistantly onto a thin strip of 0.45 μm nitrocellulose membrane (Thermo Scientific). The spots were dried in air and the membrane was blocked in 5% w/v skimmed milk solution in PBS for 30 min, followed by incubation with the primary antibody and subsequent processing as western blots similar to what has been described above.

For peptide competition assays, 2 ml of the primary antibody of the same dilution which was used for dot-blot assays, was pre-blocked with 5 or 10 μg of either the specific or non-specific peptides in 1% w/v skimmed milk solution in PBS for about 12 h at 4°C and about 12 rpm. This blocked antibody was then used to probe the blots spotted with increasing amounts of the peptide. Further processing of the blots was carried out similar to western blotting.

## 2.2. Plasmids and Constructs

### 2.2.1. Cloning and subcloning

The general procedure followed for cloning (where a specific DNA from a pool of DNA was inserted into a vector) and sub-cloning (where a specific DNA was transferred from one vector to another) is as follows. To clone CDS of specific genes into expression vectors, primers were designed having suitable restriction enzyme sites (present in the MCS region of the vectors) which would generate overhangs in the inserts after their digestion with the specific restriction enzymes. It was kept in mind that there was no site within the coding sequence, of those specific restriction enzymes. Also, the compatibility of the enzymes in the same digestion buffer was ensured. The usual length of the cloning primers would be between 18 – 25 nucleotides, having a GC content of about 50%, Tm of about 70 – 80°C, and ending with a G or a C at the 3’ end. It was desirable that the primers have a weak to moderate propensity of secondary structure formation and no primer dimer formation. All primers were procured in lyophilized form custom synthesized at Sigma. The lyophilized primers were reconstituted in an appropriate volume of sterilized de-ionized water to form 100 μM stocks.

The PCR mix for the amplification of the insert sequence contained 0.2 mM of dNTP mix, 0.5 μM of forward primer, 0.5 μM of reverse primer, 50 ng of template DNA, 0.02 U/μl of

Phusion High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA, USA) in the appropriate 1X concentrate buffer supplied by the manufacturer. The PCR program was standardized based on the size of the amplicon and the primer parameters to minimize non-specific amplification. The PCR amplicons were then purified using commercial kits described in Section 2.1.3. The purified PCR amplicons and plasmid vectors were digested sequentially with the specific restriction enzymes (New England Biolabs) to generate the sticky ends. Purified double digested vectors and inserts were then ligated in 1:3 or 1:5 molar ratio using the T4 DNA Ligase enzyme (New England Biolabs) in the suitable ligation buffer at 25°C for 30 min. This mixture was then used for the transformation of *E. coli* DH5α or XL-10 competent cells. Colonies obtained after transformation were screened and positive colonies were confirmed by insert released after double digestion of the isolated plasmids with the specific restriction enzymes, and Sanger sequencing.

The plasmids used in this study which were previously generated in the laboratory or requested from different scientific groups are listed in Appendix Table A.1. The constructs which were generated during this study by cloning from a cDNA pool or sub-cloning from one vector to another, are described in the following sections. The specific primers used for the purposes of cloning and subcloning are listed in Appendix, Table A.2.

#### *2.2.1.1. Bacterial expression construct of wild-type (WT) human NPM2*

To generate bacterial expression construct of C-terminal $His_6$-tagged WT human NPM2 by sub-cloning, the specific insert was PCR amplified (Figure 2.1B) from the N-terminal $His_6$-tagged human NPM2 clone in pET14b (Okuwaki et al. 2012) mentioned in Appendix Table A.1, using specific primers listed in Appendix Table A.2. Since the *Nco* I site itself has the nucleotide sequence ATG which is also the start codon, the forward primer was designed such that two nucleotides, AT, from the start codon of the NPM2 insert sequence, was replaced with G to remove a second start codon in the translating sequence as well as maintain the frame of the sequence. The reverse primer did not have any stop codon so as to obtain a fused C-terminal $His_6$ tag in the translated protein. The insert was subcloned in pET28b (Novagen, Madison, WI, USA, Figure 2.1A) between *Nco* I and *Xho* I restriction sites using the previously generated clone of C-terminal $His_6$-tagged WT human NPM1 in pET28b vector (Swaminathan et al. 2005). The clone was confirmed by insert release (Figure 2.1C) and sequencing.

To generate bacterial expression construct of N-terminal FLAG-tagged WT human NPM2 by sub-cloning, WT human NPM2 insert sequence codon-optimized for bacterial expression, was PCR amplified (Figure 2.2B) from the N-terminal $His_6$-tagged human NPM2 plasmid in pET14b (Okuwaki et al. 2012), using the primers listed in Appendix Table A.2. The reverse primer sequence contained the stop codon CTA after the *Xho* I restriction site. The insert was subcloned in pET21b vector (Novagen, Figure 2.2A) between *Hind* III and *Xho* I restriction sites using a previously generated clone of N-terminal FLAG-tagged WT human NPM1 in pET21b vector (Senapati P, Ph.D. thesis, 2014). The clone was confirmed by insert release (Figure 2.2C) and sequencing.

#### *2.2.1.2. Bacterial expression construct of wild-type (WT) human NPM2 fragment*

Primers for amplification of the coding sequence (CDS) for a peptide fragment (sequence: LEGKQSCRLLLHTICLGEKAKEEMHRVEILPPANQEDKKMQPVTIASLQA) of human NPM2 were designed (Appendix Table A.2). The desired sequence of DNA was PCR amplified (Figure 2.1D) from the full-length NPM2 CDS (from the cDNA pool of HepG2 cells). The insert was double digested with the specific restriction enzymes and cloned within *Nde* I and *Hind* III restriction enzyme sites of pET28b vector (Novagen, Figure 2.1A) to generate an N-terminal $His_6$-tagged NPM2 fragment. The clone was confirmed by insert release (Figure 2.1E) and sequencing.

#### *2.2.1.3. Bacterial expression construct of wild-type (WT) Xenopus Nucleoplasmin (NP)*

The bacterial expression clone for untagged *Xenopus* Nucleoplasmin in pET11b vector was a kind gift from Prof. Arturo Muga of Biofisika, Spain (Hierro et al. 2001). The *Xenopus* Nucleoplasmin (NP) insert was amplified by PCR (Figure 2.1F) from this clone using specific primers listed in Appendix Table A.2. The insert was sub-cloned into pET28b vector (Novagen, Figure 2.1A) between *Nde* I and *Hind* III restriction sites using the prior generated clone of N-terminal $His_6$-tagged NPM2 fragment in pET28b vector described above. The clone was confirmed by insert release (Figure 2.1G) and sequencing.

### *2.2.1.4. Bacterial expression construct of wild-type (WT) mouse Nucleoplasmin (Npm2)*

Npm2 CDS was PCR amplified (Figure 2.1H) from the cDNA pool of adult mouse ovary tissue using specific primers listed in Appendix Table A.2. The insert was double digested with the specific restriction enzymes and cloned within *Nde* I and *Hind* III restriction enzyme sites of pET28b vector (Novagen, Figure 2.1A) to generate N-terminal $His_6$-tagged Npm2. The clone was confirmed by insert release (Figure 2.1I) and sequencing.

### *2.2.1.5. Mammalian expression construct of wild-type (WT) human NPM2*

NPM2 CDS was amplified (Figure 2.3B) using specific primers (Appendix Table A.2) from total poly-A tailed cDNA pool synthesized by reverse transcription of total mRNA isolated from HepG2 (human hepatocellular carcinoma) cells. The double digested insert was cloned within *Hind* III and *BamH* I restriction enzyme sites of p3xFLAG-CMV10 vector (Sigma, Figure 2.3A). The clone was confirmed by insert release (Figure 2.3C) and sequencing.

**Characterization of the 3xFLAG-NPM2 mammalian expression construct:** To check the constitutive expression of the FLAG-tagged NPM2 CDS (from the CMV promoter in the vector), 1 µg of the plasmid was transfected in HEK-293 cells. The cells were harvested post-24 h of transfection and western blotting analysis was carried out with the anti-FLAG antibody which confirmed the expression of FLAG-tagged NPM2 in the transfected cells (Figure 2.4A). To check the localization of this ectopically expressed 3xFLAG-NPM2, immunofluorescence analysis was performed in HEK-293 cells transfected with the plasmid, using anti-FLAG antibody. 3xFLAG-NPM2 showed nucleoplasmic localization (Figure 2.4B) which was in accordance with reported observations (Burns et al. 2003; Okuwaki et al. 2012).

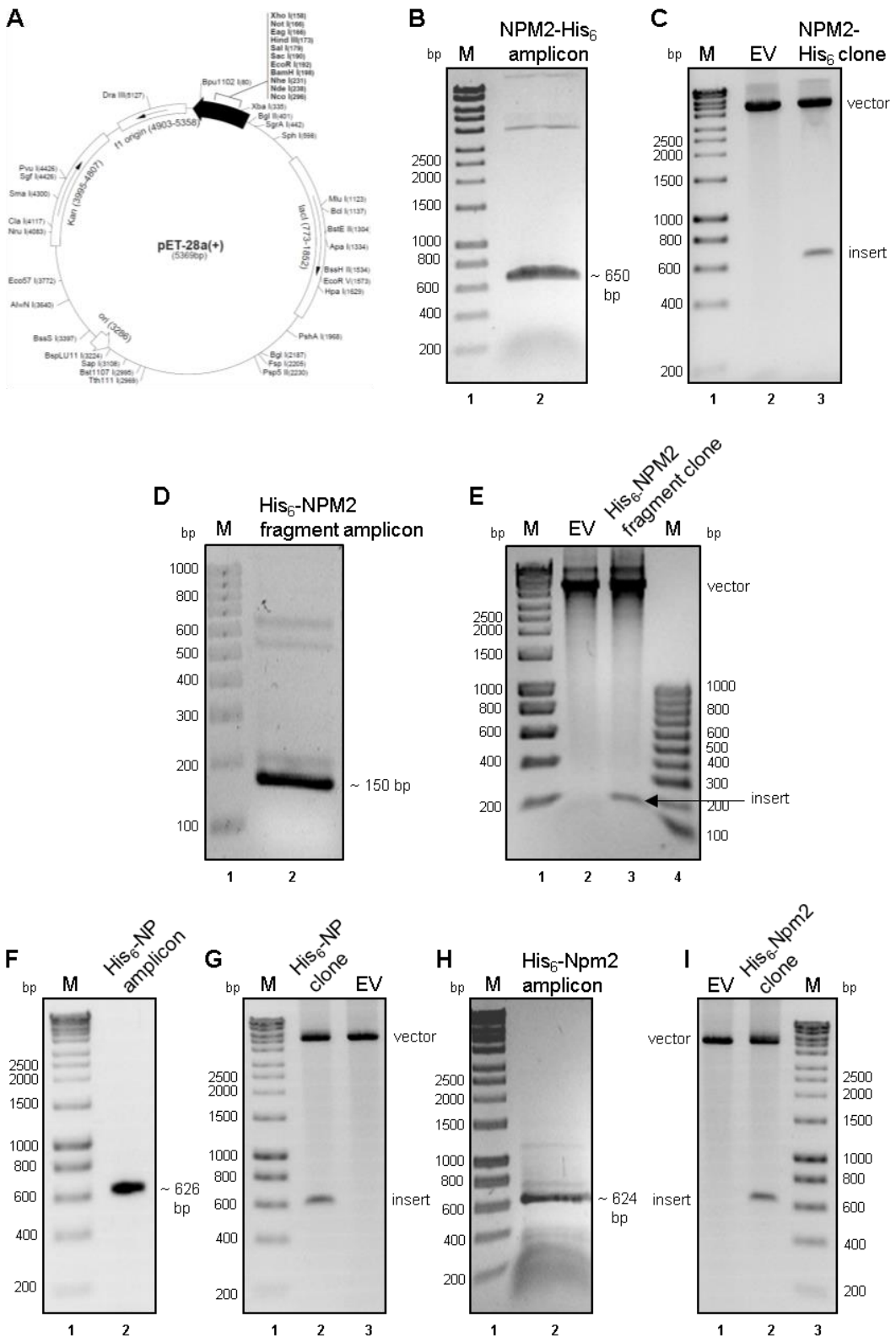



**Figure 2.1. Constructs of NPM homologs and fragments generated using the pET28b vector:** (A) Map of pET28b vector used for cloning. Amplicons of (B) NPM2-$His_6$, (D) $His_6$-NPM2 fragment, (F) $His_6$-NP, and (H) $His_6$-Npm2, generated after PCR. Confirmation of the clones of (C) NPM2-$His_6$, (E) $His_6$-NPM2 fragment, (G) $His_6$-NP, and (I) $His_6$-Npm2 by visualization of the inserts released after double digestion with respective restriction enzymes. M: DNA molecular weight marker, EV: empty vector.

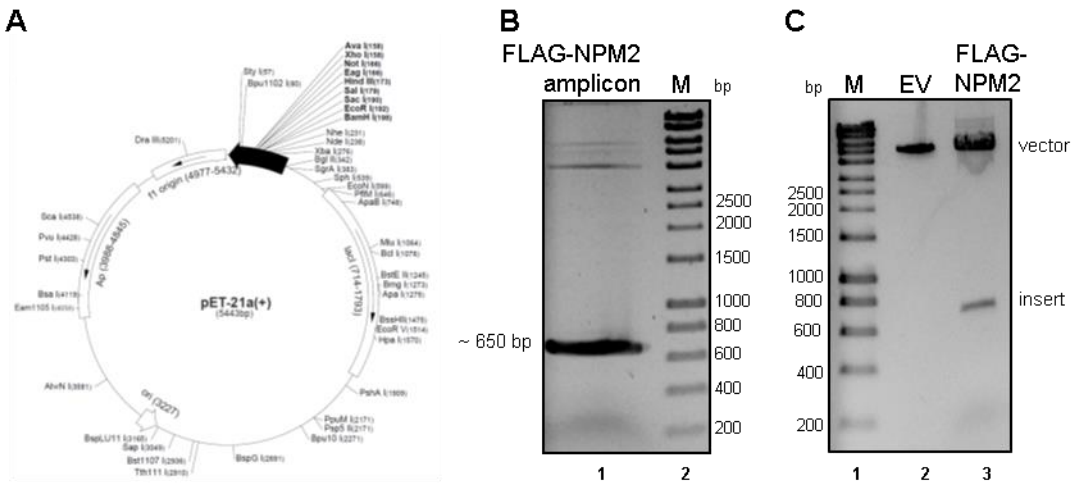



**Figure 2.2. Construct of NPM2 generated using the pET21b vector:** (A) Map of pET21b vector used for cloning. (B) The amplicon of FLAG-NPM2 generated after PCR. (C) Confirmation of the clone of FLAG-NPM2 by visualization of the insert released after double digestion with respective restriction enzymes. M: DNA molecular weight marker, EV: empty vector.

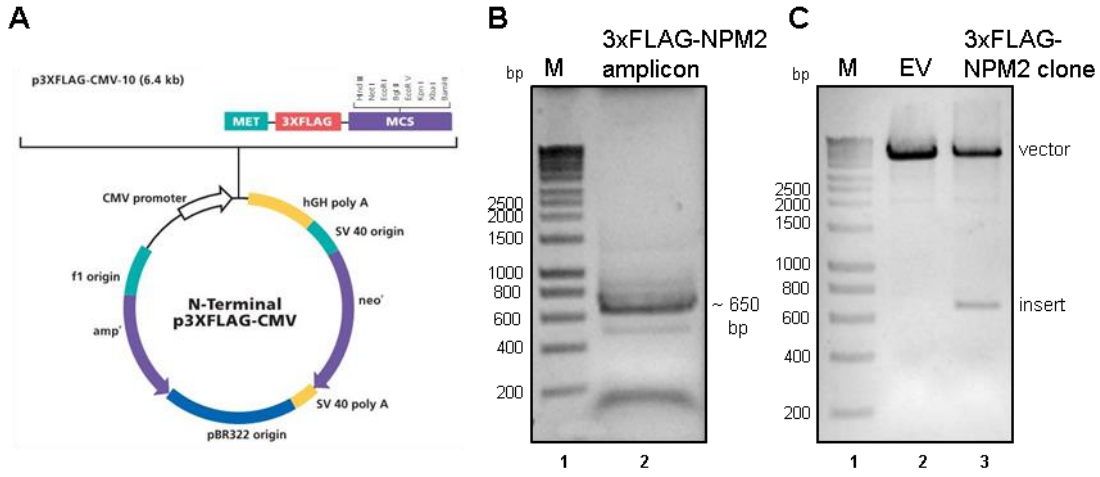



**Figure 2.3. Construct of NPM2 generated using the p3xFLAG-CMV10 vector.** (A) Map of p3xFLAG-CMV10 vector used for cloning. (B) The amplicon of 3xFLAG-NPM2 generated after PCR. (C) Confirmation of the clone of 3xFLAG-NPM2 by visualization of the insert released after double digestion with respective restriction enzymes. M: DNA molecular weight marker, EV: empty vector.

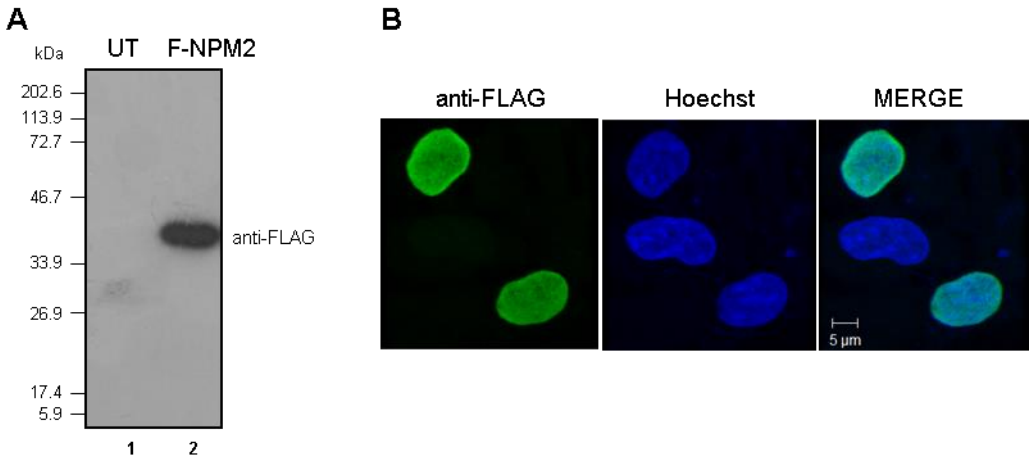



**Figure 2.4. Characterization of the 3xFLAG-NPM2 mammalian expression construct:** (A) Western blot analysis with anti-FLAG antibody showing constitutive expression of FLAG-tagged NPM2 post-24 h of transfection of HEK-293 cells with 1 µg 3xFLAG-NPM2 construct. UT: untransfected. (B) Immunofluorescence analysis using anti-FLAG antibody (green) in HEK-293 cells transfected with 1 µg 3xFLAG-NPM2 construct for 24 h. Nuclei have been stained using Hoechst (blue). Magnification is 63X and the scale bar is 5 µm.

### *2.2.1.6. $G_5$ML array template for histone transfer and in vitro transcription assays*

The $G_5$ML plasmid of about 5.5 kb size, was used in histone transfer and *in vitro* transcription assays. This construct consists of the transcription template flanked on both sides by 5 repeats of the sea urchin 5S rDNA nucleosome positioning sequences. The transcriptional cassette also consists of 5 repeats of the activator Gal4 binding site ($G_5$) upstream to the Adenovirus Major Late (ML) promoter which controls the transcription of a 380 bp G-less cassette (Kundu et al. 2000) (Figure 2.5).

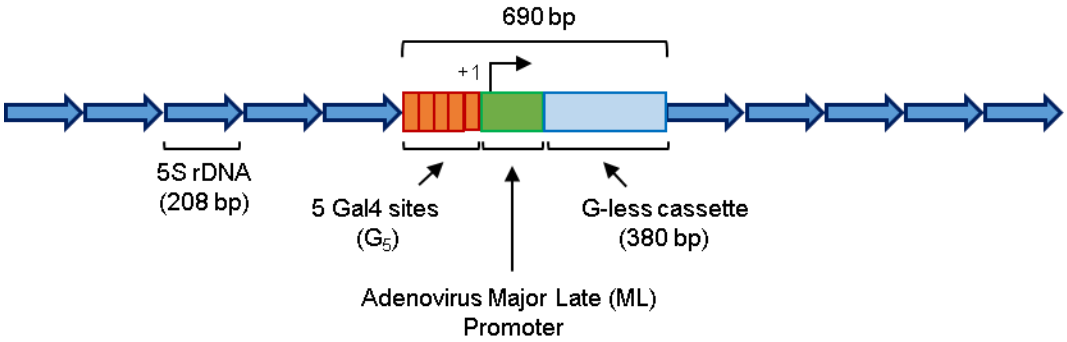
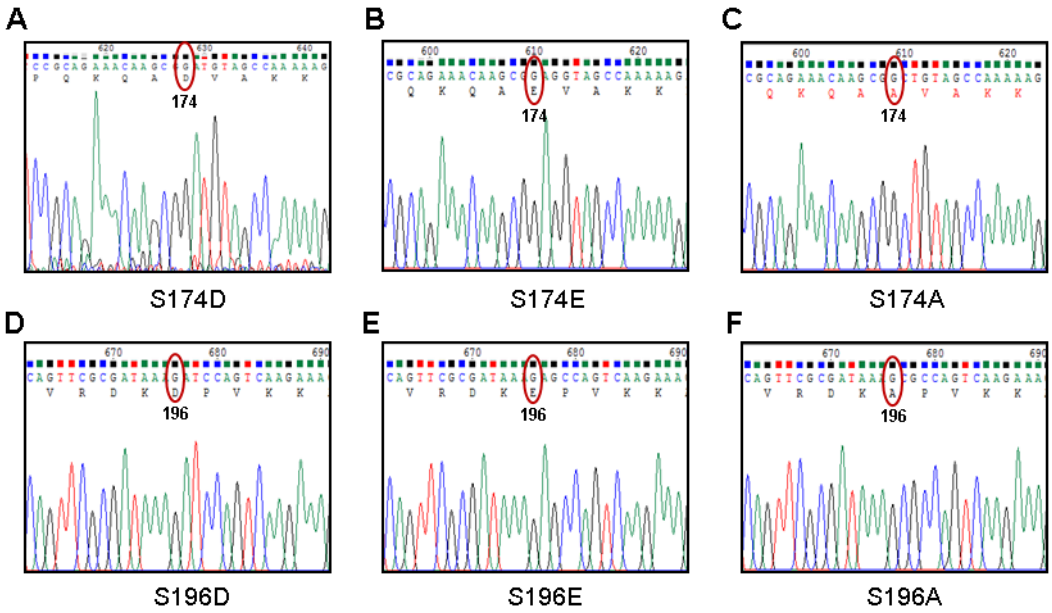



**Figure 2.5. Schematic representation of the $G_5$ML array template:** The central 690 bp region contains 5 Gal4 binding sites ($G_5$, denoted in orange boxes) upstream to the Adenovirus Major Late (ML) promoter (denoted by the green box) that drives the transcription of a 380 bp long G-less cassette (denoted by the light blue box). This region is flanked on each side by 5 sea urchin 5S

rDNA nucleosome positioning sequence (denoted by blue arrows), each being 208 bp long (Kundu et al. 2000).

### *2.2.2. Site-directed mutagenesis*

Various point mutations in the specific clones were generated using the site-directed mutagenesis technique using the QuikChange II XL Site-Directed Mutagenesis Kit (Stratagene, San Diego, CA, USA) according to the manufacturer's protocol. Briefly, 20 ng of the template plasmid DNA, 125 ng each of the forward and reverse mutagenic primers (Appendix Table A.3), 0.2 mM of dNTPs and 5U of PfuUltra High-Fidelity DNA Polymerase (Agilent, Santa Clara, CA, USA) in 1X reaction buffer were used for PCR as recommended in the manufacturer's protocol. The final reaction mixture was incubated with the restriction enzyme *Dpn* I (NEB) at 37°C for 1 – 1.5 h to digest the parental methylated plasmid DNA. 10 µl of this mixture was transformed into *E. coli* XL-10 Gold ultra-competent cells. The cells were recovered in LB media enriched with 10 mM glucose, 12.5 mM $MgSO_4$ and 12.5 mM $MgCl_2$ pre-heated to 42°C, and plated against the respective resistance marker containing LB agar media. Plasmids were isolated from screened colonies and confirmed for the mutation by sequencing (Figure 2.6).



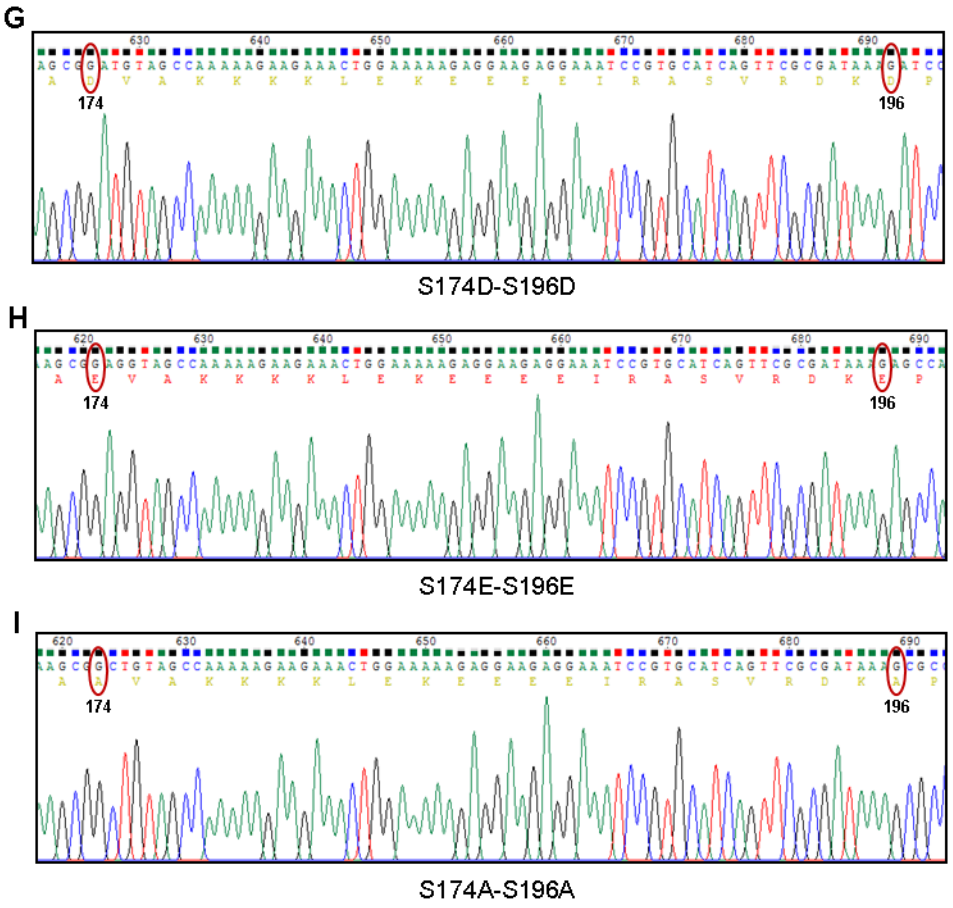



**Figure 2.6. Confirmation of the point mutations:** Chromatograms highlighting the point mutations in $His_6$-NPM2 generating (A) S174D $His_6$-NPM2 (B) S174E $His_6$-NPM2 (C) S174A $His_6$-NPM2 (D) S196D $His_6$-NPM2 (E) S196E $His_6$-NPM2 (F) S196A $His_6$-NPM2 (G) S174D-S196D $His_6$-NPM2 (H) S174E-S196E $His_6$-NPM2, and (I) S174A-S196A $His_6$-NPM2. The mutated residues (circled) are numbered. The chromatograms were visualized using Chromas software (version 2.6.5, Technelysium Pty Ltd).

## *2.3. Cell culture and Animals*

### *2.3.1. Mammalian cell culture*

The various cell lines used in this study and their culture requirements are listed in Appendix, Table A.4. All the media except BEGM were supplemented with 10% (or 20% in case of the mouse myoblast cell line C2C12 for proliferation and maintenance) v/v fetal bovine serum (FBS) (Life Technologies, Bengaluru, India) and 1X antibiotics containing penicillin, streptomycin, and amphotericin B (HiMedia). For differentiating the C2C12 myoblasts into myotubes, the proliferation medium (DMEM + 20% v/v FBS) was changed to differentiation medium (DMEM + 2% v/v horse serum (HiMedia)) and grown for 1 – 5

days depending on the experimental requirement. The cells were grown at 37ºC and 5% $CO_2$ in a humidified chamber. All cell lines used in this study were routinely tested for *Mycoplasma* contamination using the PCR-based Mycoplasma Detection Kit (Applied Biological Materials Inc., Vancouver, Canada, Cat. No. G238) and used for no more than 10 passages. The proper morphology and homogeneity of the cells were ensured by visualization under the microscope (Axiovert 200M or Axiovert 40 CFL (Carl Zeiss AG, Oberkochen, Germany) and IX73 (Olympus, Shinjuku, Tokyo, Japan)) (Figure 2.7).

To freeze and preserve cells stocks, cells were collected after trypsinization and centrifugation at 1000 rpm for 3 min, and were resuspended in a freezing mixture (40% v/v medium, 50% v/v FBS, and 10% v/v cell culture grade DMSO (Sigma)) and immediately transferred to cryovials (Corning, Corning Inc., NY, USA). The stocks were slowly frozen in a cryo-cooler with isopropanol at −80°C for at least 24 h, and later transferred to liquid nitrogen cylinders and stored in the liquid phase of liquid nitrogen. Usually, one and three stocks were prepared from a 90% confluent culture in a T25 and T75 flask respectively.

### *2.3.2. Transfection of plasmids in mammalian cells*

HEK-293 and H1299 cells were transfected with mammalian expression plasmids of specific proteins using the Lipofectamine 2000 transfection Reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. In brief, cells were cultured in dishes until 70 – 90% confluency. With respect to a 30 mm dish, the desired amount of plasmid DNA was diluted in 150 µl of Opti-MEM Reduced Serum Medium (Gibco) while Lipofectamine was diluted separately in another 150 µl of Opti-MEM. For cells grown in larger or smaller sized dishes or wells, the volume of the Opti-MEM to be used was proportionately modified. The DNA to Lipofectamine ratio was standardized for a specific cell line. For HEK-293, 2 µl of Lipofectamine were taken per µg of DNA, while for H1299, 3 µl of Lipofectamine were taken per µg of plasmid DNA to be transfected. The diluted DNA was then added to the diluted Lipofectamine, tap mixed and incubated at RT for 5 min. Meanwhile, the spent medium on the cells to be transfected was removed and fresh complete medium was added. The DNA-lipid complex was then added dropwise over the cells which were then incubated at 37°C. The results of the transfection were analyzed at least after 24 h post-transfection.

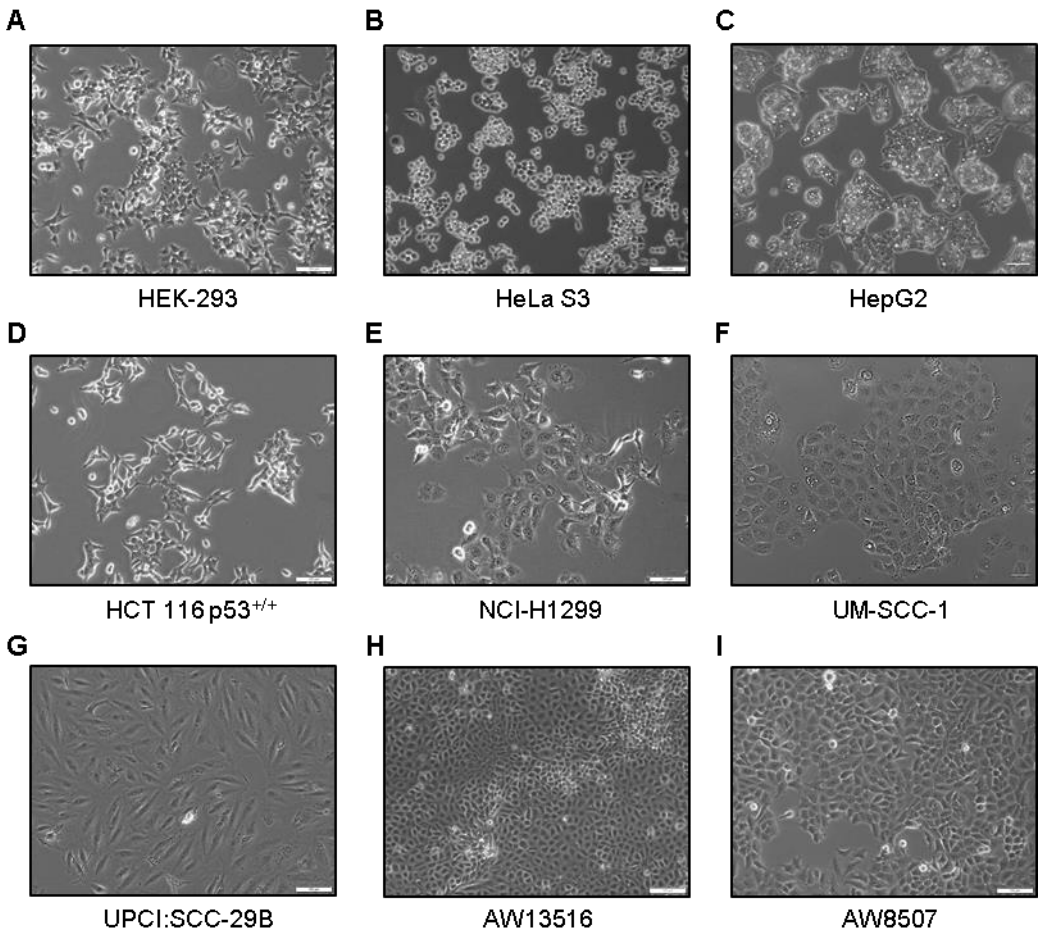



**Figure 2.7. Human cancer or immortalized cell lines commonly used in this study:** (A – I) Photomicrographs of human cancer/transformed cell lines (A) HEK-293, (B) HeLa S3, (C) HepG2, (D) HCT 116 $p53^{+/+}$, (E) NCI-H1299, (F) UM-SCC-1, (G) UPCI:SCC-29B, (H) AW13516, and (I) AW8507 cell lines are shown. Magnification is 5X or 10X and the scale bar is 100 µm.

### *2.3.3. Transfection of silencing RNA (si-RNA) in mammalian cells*

si-RNA against human c-fos (Santacruz, Dallas, TX, USA, Cat No. sc-29221, Lot No. C1417), human c-jun (Santacruz, Cat No. sc-29223, Lot No. H3117) and scrambled negative control (Ambion, Austin, TX, USA, Cat No. AM4611, Lot No. AS0240KM) were transfected in SCC-29B cells using Lipofectamine RNAiMAX or Lipofectamine 2000 Transfection Reagents (Invitrogen) in 6-well format as per manufacturer's protocol. si-RNAs procured in lyophilized form were dissolved in a specific volume of RNase-free water by shaking at RT for 30 min to get a stock concentration of 10 µM. The procedure of the transfection was essentially the same as that for the transfection of plasmid DNA described above (Section 2.3.2). In a 6-well format, for transfection of 30 nM of si-RNA, 9 µl of the transfection reagent was taken. Transfection was done twice at 24 h intervals

and cells were harvested after 48 h of the first transfection, for analyses of protein and mRNA expressions.

### *2.3.4. Stable cell line generation and characterization*

#### *2.3.4.1. Stable cell line generated using pEBTetD vector*

##### *2.3.4.1.1. Stable cell line generated in H1299 p53$^{-/-}$ background for doxycycline-inducible expression of FLAG-tagged mutant R175H p53*

H1299 p53$^{-/-}$ cells (Figure 2.8B upper panel) were transfected with pEBTetD-3xFLAG-R175H p53 construct (Figure 2.8A) for 24 h, followed by selection with 1.1 μg/ml of puromycin for 4 days. The positive colonies (Figure 2.8B lower panel) were selected and characterized for the induction of FLAG-tagged R175H p53 expression by western blotting (Figure 2.8C) and immunofluorescence (Figure 2.8D) analyses after treatment of the cells with 1 μg/ml of doxycycline.

#### *2.3.4.2. Stable cell line generated using p3xFLAG-CMV10 vector*

##### *2.3.4.2.1. Stable cell line in HEK-293 background for constitutive expression of FLAG-tagged NPM2*

HEK-293 cells (Figure 2.9B upper panel) were transfected with the 3xFLAG-NPM2 construct in the p3xFLAG-CMV10 vector (Figure 2.9A) for 24 h followed by selection with 800 μg/ml of G418 antibiotic for 7 days. The antibiotic-resistant colonies (Figure 2.9B lower panel) were trypsinized and pooled together and characterized for the expression of FLAG-tagged NPM2 by western blotting (Figure 2.9C) and immunofluorescence (Figure 2.9D) analyses.

##### *2.3.4.2.2. Stable cell line in UM-SCC-1 p53$^{-/-}$ background for constitutive expression of FLAG-tagged mutant R175H p53 or empty vector*

UM-SCC-1 p53$^{-/-}$ cells were transfected with 3xFLAG-R175H p53 or the empty vector construct (Figure 2.9A) for 24 h, followed by selection with 600 μg/ml of G418 antibiotic for 7 days. The positive colonies from the two cell lines (Figure 2.9E) were selected and

characterized for the expression of FLAG-tagged R175H p53 by western blotting (Figure 2.9F).

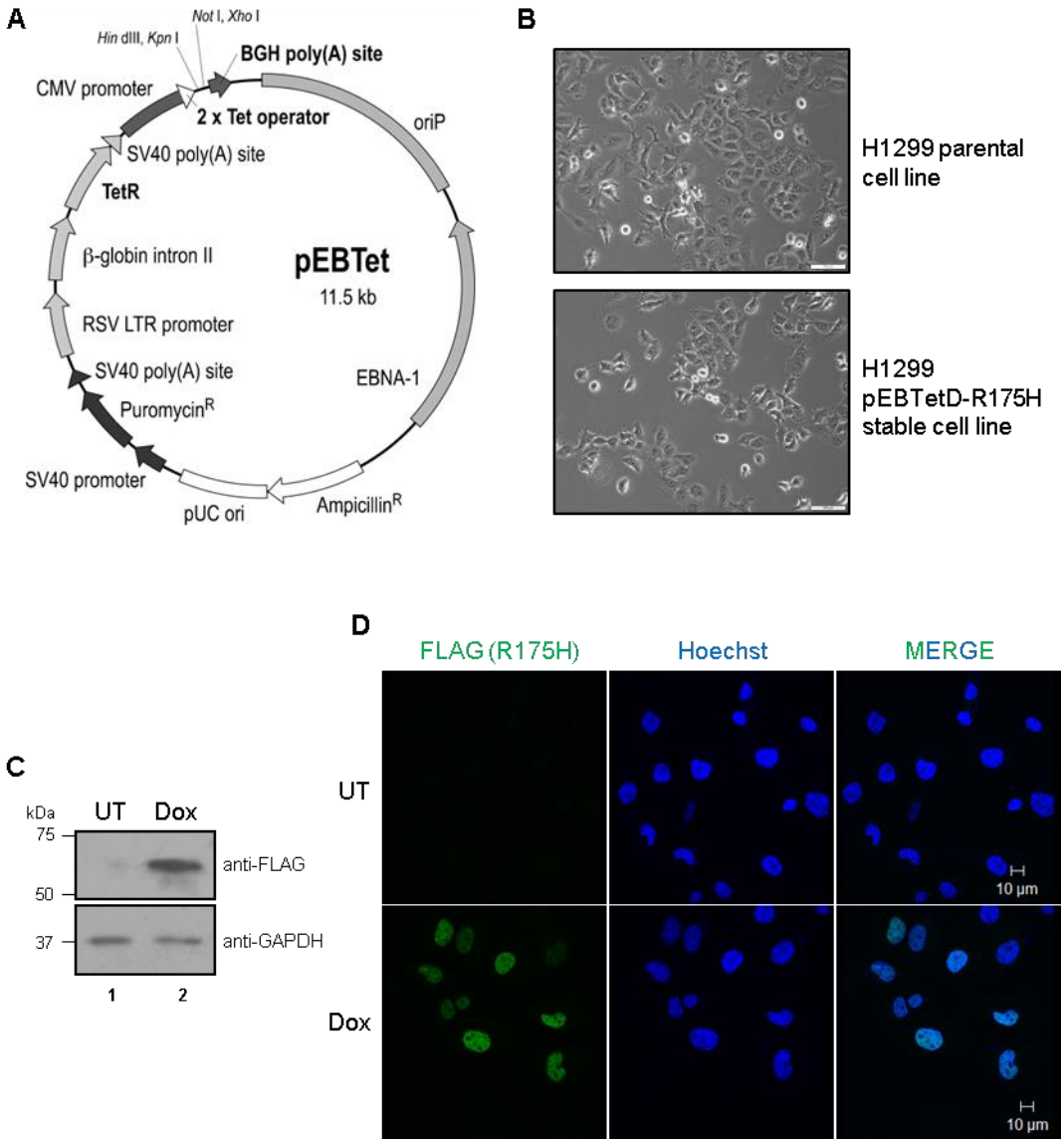



**Figure 2.8. Characterization of the stable cell line in H1299 for doxycycline-inducible expression of 3xFLAG-tagged R175H p53:** (A) Map of pEBTet, the parent vector of the pEBTetD vector used for generating the doxycycline-inducible expression construct of 3xFLAG-tagged R175H p53 (Bach et al. 2007). CMV promoter: cytomegalovirus promoter, 2X Tet operator: two repeats of tetracycline operator sequence, BGH poly(A) site: bovine growth hormone poly(A) site (absent in pEBTetD vector), oriP: Epstein-Barr virus (EBV) origin of replication, EBNA-1: Epstein-Barr nuclear antigen 1, Ampicillin$^R$: ampicillin resistance marker gene (for selection in *E. coli*), pUC ori: pUC vector's origin of replication sequence (for replication in *E. coli*), SV40 promoter: simian virus 40 promoter, Puromycin$^R$: puromycin resistance marker gene (for selection in mammalian cells), SV40 poly(A) site: simian virus polyadenylation sequence and terminator of transcription, RSV LTR promoter: Rous sarcoma virus long terminal repeat promoter, TetR: tetracycline repressor. (B) Photomicrographs of the H1299 parental cells (upper panel) and H1299 stable cells for doxycycline-inducible expression of R175H generated using the pEBTetD vector (H1299-pEBTetD-R175H) (lower panel). Magnification is 10X and the scale bar is 100 µm. (C) Western blot analysis with anti-FLAG antibody (upper panel) showing doxycycline-inducible (Dox) expression of

FLAG-tagged R175H p53 in the H1299 stable cells. The bottom panel shows western blot with anti-GAPDH antibody. (D) Immunofluorescence analysis using anti-FLAG antibody (green) in H1299 stable cells treated with 1 µg/ml of doxycycline (Dox) compared to untreated (UT). Nuclei have been stained using Hoechst (blue). Scale bar is 10 µm.

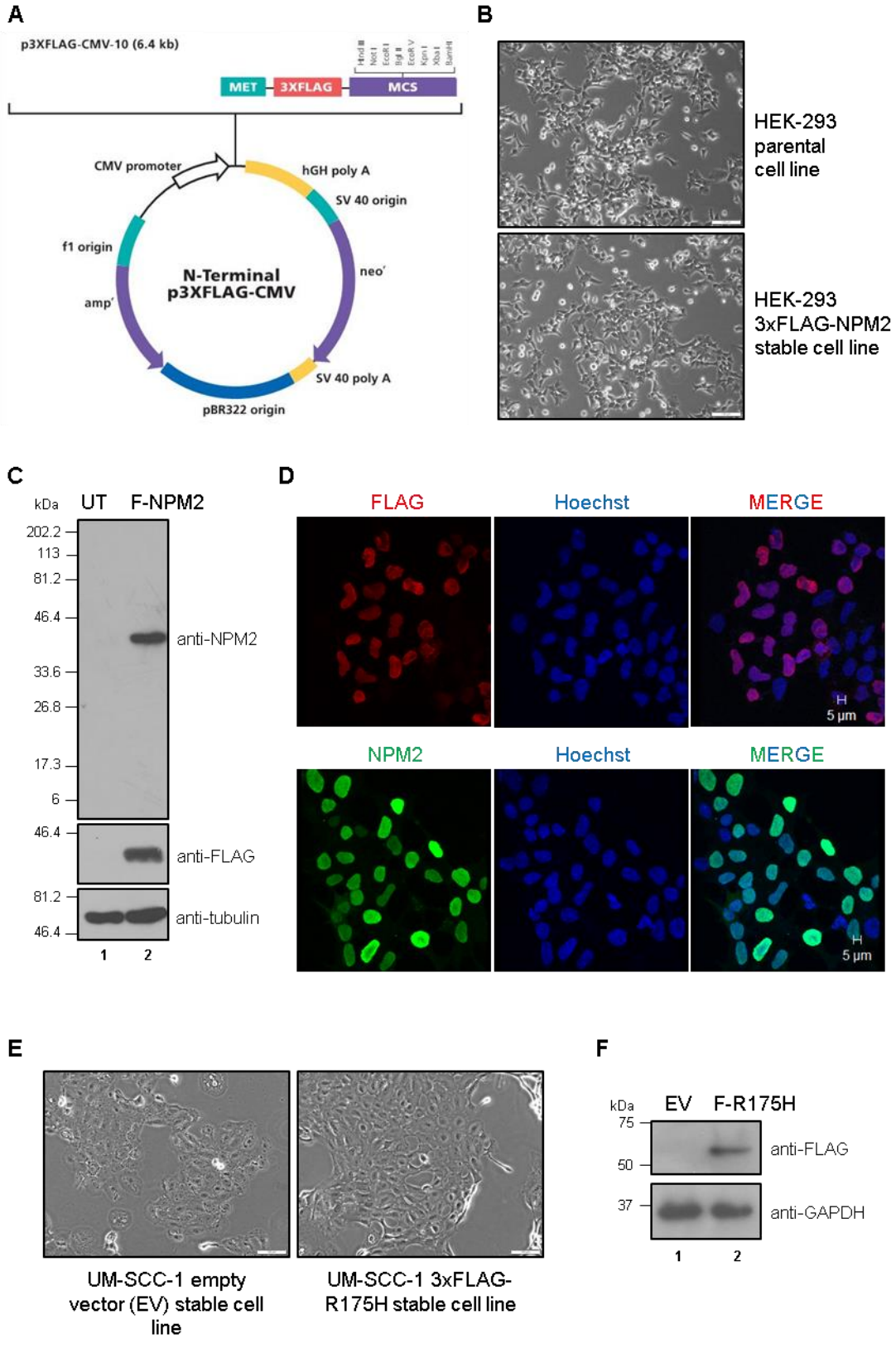


**Figure 2.9. Characterization of the stable cell lines generated using p3xFLAG-CMV10 vector in HEK-393 and UM-SCC-1 cells for constitutive expressions of 3xFLAG-tagged NPM2, R175H p53 or the empty vector:** (A) Schematic representation of the p3xFLAG-CMV10 vector used to generate the constitutive expression constructs of 3xFLAG-NPM2 and 3xFLAG-R175H p53. MET: Start codon for amino acid residue methionine, 3xFLAG: three repeats of FLAG octapeptide sequence, MCS: multiple cloning site, hGH poly A: Human growth hormone polyadenylation signal, SV 40 origin: simian virus 40 origin of transcription, $neo^r$: neomycin resistance marker gene (for selection in mammalian cells), SV 40 poly A: simian virus polyadenylation sequence and terminator of transcription, pBR322 origin: pBR322 vector's origin of replication sequence, $amp^r$: ampicillin resistance marker gene (for selection in *E. coli* cells), f1 origin: origin of replication from f1 phage, CMV promoter: cytomegalovirus promoter. (B) Photomicrographs of the HEK-293 parental cells (upper panel) and HEK-293 stable cells for constitutive expression of FLAG-tagged NPM2 generated using the p3xFLAG-CMV10 vector (HEK-293 3xFLAG-NPM2) (lower panel). Magnification is 10X and the scale bar is 100 µm. (C) Western blot analysis showing expression of FLAG-tagged NPM2 in the HEK-293 stable cells for constitutive expression of 3xFLAG-NPM2, compared to untransfected (UT) control. The upper panel shows western blot with in-house raised anti-NPM2, the middle panel shows western blot of the stripped blot with anti-FLAG and bottom panel with anti-tubulin antibody respectively. (D) Immunofluorescence analysis showing expression and localization of FLAG-tagged NPM2 after immunostaining the HEK-293 stable cells with anti-FLAG (red) and anti-NPM2 (Sigma, Cat. No. SAB1400381) (green) antibodies. Nuclei were stained with Hoechst. Scale bar is 5 µm. (E) Photomicrographs of the UM-SCC-1 stable cells for the constitutive expression of the empty vector ((UM-SCC-1 EV) (left panel) and FLAG-tagged R175H (UM-SCC-1 3xFLAG-R175H) (right panel) generated using the p3xFLAG-CMV10 vector. Magnification is 10X and the scale bar is 100 µm. (F) Western blot analysis with anti-FLAG antibody (upper panel) showing expression of FLAG-tagged R175H p53 compared to the empty vector (EV) control, in the UM-SCC-1 stable cells. The bottom panel shows western blot with anti-GAPDH antibody.

### *2.3.4.3. Stable cell line generated using pTRIPZ vector*

#### *2.3.4.3.1. Stable cell line in AW13516 background for doxycycline-inducible knockdown of NPM1*

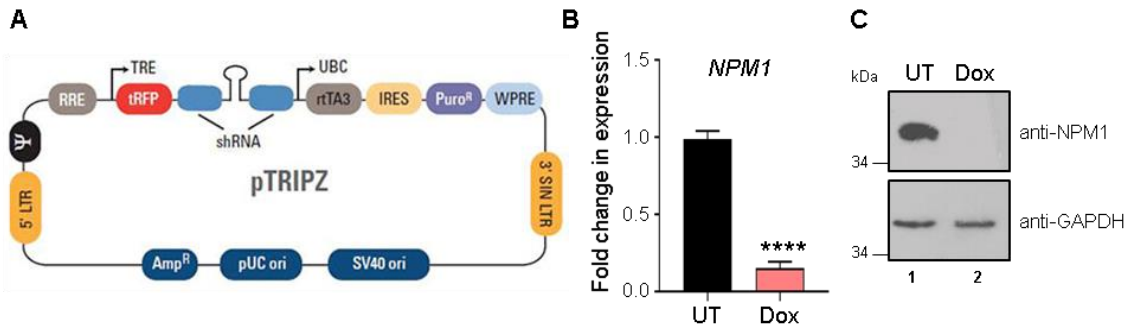
AW13516 cells were transfected with NPM1 shRNA construct in pTRIPZ vector (Thermofisher Scientific, Waltham, MA, USA) (Figure 2.10A) for 24 h, followed by selection with 1.6 µg/ml puromycin antibiotic for 4 days. The puromycin resistant colonies were selected and cultured under antibiotic selection. The stable cell lines thus generated were then characterized for the expression of the shRNA located downstream to the tetracycline-inducible promoter and turbo RFP reporter, by treating the cells with 2 µg/ml of doxycycline every 24 h and scoring for the expression of TurboRFP in a fluorescence microscope (IX73, Olympus). Clones with high expression of the shRNA after doxycycline induction were expanded and sorted using a BD FACSAria™ III (BD Biosciences-US, Franklin Lakes, NJ, USA) cell sorter to select for the high TurboRFP expressing cells. These cells were maintained in complete MEM supplemented with L-glutamine and 1.1 –

1.6 µg/ml of puromycin. The efficient knockdown of NPM1 at the mRNA and protein levels upon doxycycline treatment was ensured by RT-qPCR (Figure 2.10B) and western blotting (Figure 2.10C) analyses.

AW13516-shNPM1 stable cells were transfected with pGL4 luc Neo construct (Promega) for 24 h followed by selection with G418 antibiotic (800 µg/ml) for 1 week. Colonies thus obtained were picked and cultured under antibiotic selection. The stable cell lines thus generated were then characterized for expression of Luciferase by performing Luciferase reporter assays. The clones with high Luciferase activities were expanded, the NPM1 shRNA expression was induced by treatment with doxycycline (2 µg/ml) every 24 h, and sorted using a BD FACSAria™ III (BD Biosciences-US) cell sorter to select for the high TurboRFP expressing cells (Figure 2.10D). The cells were maintained in MEM complete medium supplemented with L-glutamine, 800 µg/mL of G418 and 1.1 µg/ml of puromycin. The knockdown of NPM1 in these cells upon doxycycline treatment was confirmed by RT-qPCR (Figure 2.10E) and western blotting (Figure 2.10F) analyses.

### *2.3.4.3.2. Stable cell line in AW8507 background for doxycycline-inducible knockdown of YY1*

The stable cell line in the AW8507 background harboring doxycycline-inducible construct of shRNA against YY1 (shYY1) was previously generated in the laboratory (Behera et al. 2019) in a similar way as described above (Section 2.3.4.3.1). The RFP expression in the sorted cells was checked under a fluorescence microscope (IX73, Olympus) (Figure 2.10G) and the knockdown of YY1 in these cells upon doxycycline treatment was confirmed by RT-qPCR (Figure 2.10H) and western blotting (Figure 2.10I) analyses.



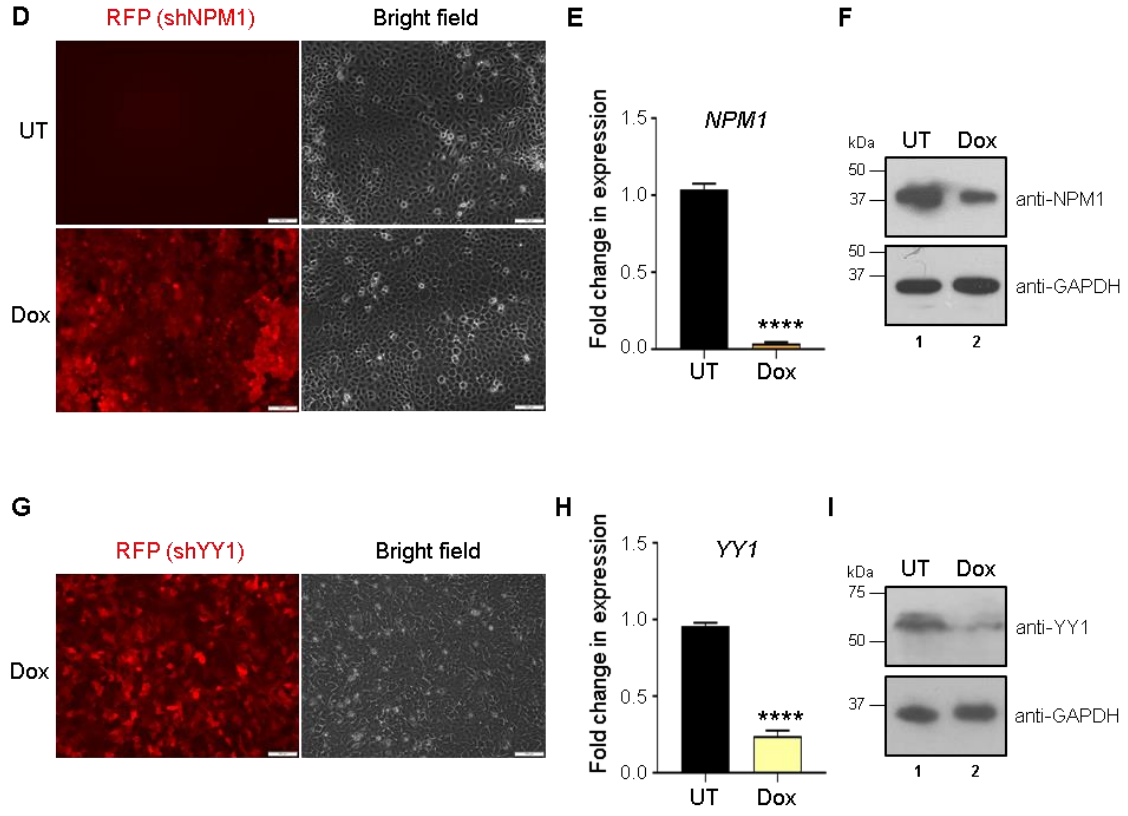



**Figure 2.10. Characterization of the stable cell lines generated using pTRIPZ vector in AW13516 and AW8507 cells for doxycycline-inducible expressions of shRNA against NPM1 and YY1:** (A) Schematic representation of the pTRIPZ vector having the doxycycline-inducible shRNA against NPM1 and YY1. TRE: tetracycline response element (tetracycline-inducible promoter), tRFP: TurboRFP reporter (for visual tracking of transduction and shRNA expression), shRNA: short hairpin RNA, UBC: human ubiquitin C promoter (for constitutive expression of rtTA3 and puromycin resistance genes), rtTA3: reverse tetracycline transactivator 3 (for tetracycline-dependent induction of the TRE promoter), Puro$^{R}$: puromycin resistance marker gene (for selection in mammalian cells), IRES: internal ribosomal entry site (allows expression of rtTA3 and puromycin resistance genes in a single transcript), 5' LTR: 5' long terminal repeat, 3' SIN LTR: 3' self-inactivating long terminal repeat (for increased lentivirus safety), Ψ: psi packaging sequence (allows viral genome packaging using lentiviral packaging systems), RRE: rev response element (enhances titre by increasing packaging efficiency of full-length viral genomes), WPRE: woodchuck hepatitis post-transcriptional regulatory element (enhances transgene expression in the target cells), Amp$^{R}$: ampicillin resistance marker gene (for selection in *E. coli*), pUC ori: pUC vector's origin of replication sequence (for replication in *E. coli*), SV40 ori: simian virus 40 origin of transcription. (B, E) Bars represent fold change in expression of NPM1 mRNA as analyzed by RT-qPCR after doxycycline treatment (Dox) to (B) AW13516 cells stably harboring the shRNA against NPM1 (AW13516-shNPM1) or (E) AW13516-shNPM1 cells stably harboring the pGL4 luc Neo plasmid (AW13516-shNPM1-luc+), compared to untreated (UT) control. Values are mean + SEM from two independent experiments having three technical replicates per experiment. (C, F) Western blot analysis showing expression of NPM1 (upper panel) with (Dox) or without (UT) doxycycline treatment to (C) AW13516-shNPM1 and (F) AW13516-shNPM1-luc+ cells. The bottom panel shows western blot with anti-GAPDH antibody. (D) Photomicrographs showing expression of RFP (red) in doxycycline-treated (Dox) AW13516-shNPM1-luc+ cells (bottom left panel) compared to untreated (UT) control (upper left panel). Corresponding bright-field images are shown in the right. (G) Photomicrographs showing expression of RFP (red) in doxycycline-treated (Dox) AW8507 cells stably harboring the shRNA against YY1 (AW8507-shYY1) (left). Corresponding bright-field image is shown in the right.

(D, G) Magnification is 10X and the scale bar is 100 µm. (H) Bars represent fold change in expression of YY1 mRNA as analyzed by RT-qPCR after doxycycline treatment (Dox) to AW8507-shYY1 cells, compared to untreated (UT) control. Values are mean + SEM from four independent experiments. (B, E, H) Statistical significance was calculated using Student's *t*-test. ****$P < 0.0001$. (I) Western blot analysis showing expression of YY1 (upper panel) with (Dox) or without (UT) doxycycline treatment to AW8507-shYY1 cells. The bottom panel shows western blot with anti-GAPDH antibody.

### *2.3.5. Insect cell culture*

The insect ovarian cell line Sf21 derived from *Spodoptera frugiperda* was procured from Invitrogen (Cat. No. B821). The frozen cell stock was thawed at 37°C in a water bath quickly for about a minute and was diluted ten times with pre-warmed complete Grace's medium (Gibco) supplemented with 10% v/v FBS (Life Technologies, Bengaluru, India) and 1X antibiotics containing penicillin, streptomycin, and amphotericin B (HiMedia). No centrifugation was performed. The cells were seeded in a T25 flask (Corning or Eppendorf, Hamburg, Germany) and kept undisturbed for 45 min – 1 h for them to adhere to the flask, following which the medium was changed to remove the cryoprotectant DMSO. The cells were cultured in complete Grace's medium at 27°C in a BOD (biochemical oxygen demand) incubator. At about 70 – 80% confluency, the spent media was discarded and the adherent cell monolayer was scraped in 5 ml of fresh Grace's complete medium. The scraping was done gently and carefully so as to not lyse the cells by mechanical pressure, and keeping the cells always in contact with the liquid medium. Subsequent subculturing was done at a 1:3 ratio. Pelleting of the cells by centrifugation is avoided. The morphology of the cells was confirmed under an inverted microscope (CKX41 or IX73, Olympus) (Figure 2.11).

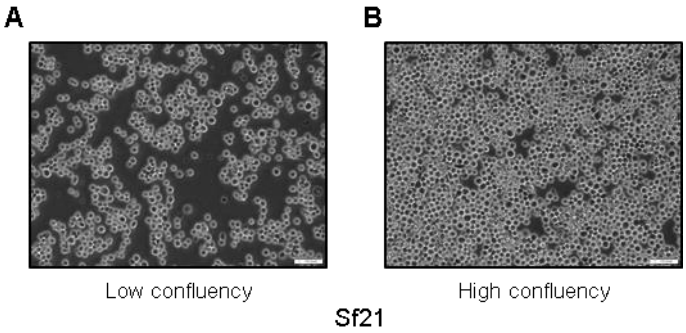



**Figure 2.11. Insect ovarian cell line Sf21:** (A – B) Photomicrographs of Sf21 cells at (A) low and (B) high confluency. Magnification is 10X and the scale bar is 100 µm.

To freeze and preserve cells stocks, Sf21 cells were collected after scraping and centrifugation at 1000 rpm 3 min and were resuspended in a freezing mixture (40% v/v Grace's medium, 50% v/v FBS, and 10% v/v cell culture grade DMSO (Sigma)). The subsequent steps were the same as described previously (Section 2.3.1).

### *2.3.6. Infection of Sf21 cells by baculovirus*

For setting up an infection of the Sf21 cells with genetically engineered baculovirus for recombinant protein expression, the cells were seeded in 150 mm dishes (Corning or Eppendorf). Prior to seeding, the cells were counted using Trypan blue dye and a hemocytometer (Neubauer's Chamber) and about 9 million cells were seeded per 150 mm dish. The cells were cultured in TC-100 insect medium (HiMedia) (20 ml per 150 mm dish). After allowing the seeded cells to adhere to the dish by keeping them undisturbed for about 30 min, the cells in each dish were infected with 500 µl of recently amplified baculovirus suspension (pre-filtered through a 0.22 µm pore-sized membrane filter), added dropwise on the medium in the dish. The steps involving the handling of the baculovirus were performed in dark. After the addition of the virus, the dishes were sealed with plastic covers and other sterilization measures were taken to minimize the spread and cross-contamination of the viruses. The infected cells were monitored after 48 h for enlarged or elongated morphology with more granular nuclei indicating infection, and were harvested within 60 – 70 h post-infection for protein isolation.

### *2.3.7. Mouse and rabbit strains*

Expression analysis of mouse *Npm2* in different tissues of mice was done majorly using adult mice of BALB/c strain in the age group of 4 – 6 months obtained from the Animal House Facility, JNCASR. Wild-type (outbred) and CD-1 strains of mice were also used to confirm if the results obtained in BALB/c were general or strain-specific.

The anti-AcNPM1 and anti-NPM2 polyclonal antibodies were raised in-house in 3-month old rabbits (New Zealand white strain).

## 2.4. Cell culture- and animal-based assays

### 2.4.1. Immunofluorescence (IF) staining

HEK-293 or H1299 cells were grown on poly-L-lysine (Sigma)-coated coverslips in 6 or 24-well plates or 30 mm dishes at 37°C in a 5% $CO_2$ incubator. The medium was aspirated out and the cells were gently rinsed in PBS to remove the medium completely. The cells were fixed in 4% w/v paraformaldehyde solution in PBS for 20 min at RT. Cells were then gently washed in Wash Buffer I (recently autoclaved PBS) thrice, 5 min each, under gentle shaking. At this stage, the fixed samples were temporarily stored at 4°C for at most 24 h in PBS or further processed for immunostaining. The washed cells were then permeabilized in permeabilization buffer (1% v/v Triton-X-100 in PBS) for 5 min at RT. This was followed by blocking in 5% v/v FBS in PBS solution at 37°C for 45 min. Cells were then incubated with the first primary antibody of suitable dilution in Wash Buffer II (1% v/v FBS, 0.1% v/v Triton-X-100 in PBS solution) at RT for 1 h under gentle shaking. This was followed by gentle washing in Wash Buffer II twice for 5 min each at RT under mild shaking and subsequent incubation with the second primary antibody of the desired dilution in Wash Buffer II for 1 h under slow rocking at RT. The subsequent steps were performed in the absence of prolonged exposure to light. The cells were washed and incubated sequentially with fluor-conjugated first and second secondary antibodies (Invitrogen) (diluted 1:1000 in Wash Buffer II) similarly as described before. Next, the cells were incubated in 1 µg/ml of Hoechst 33258 (Sigma) in PBS for 5 min at RT. Cells were finally washed in Wash Buffer I, twice for 5 min each. The coverslips were mounted on glass slides using 70% v/v glycerol, fixed and viewed under Carl Zeiss confocal laser scanning microscope (LSM 510 META) at the JNCASR Imaging Facility, Bangalore, India. Images were captured, processed and analyzed using LSM 5 Image Examiner software.

### 2.4.2. Luciferase assay

HEK-293 or H1299 cells were grown in 6-well plates and transfected with the specific plasmids at about 70 – 80% confluency. pCMV-LacZ expressing β-Galactosidase from the constitutive CMV promoter was used as the internal transfection control plasmid. After 24 h of transfection, cells were washed in PBS and lysed in 1X Reporter Lysis buffer (Promega, Madison, WI, USA) by incubation on ice for 20 min. 10 µl of the clarified cell

lysate was mixed with equal volumes of 2X Luciferase substrate (Promega) and the Luciferase counts were measured using Wallac 1409 liquid scintillation counter (PerkinElmer, Wellesley, MA, USA). For normalization of the transfection efficiency, the β-Galactosidase assay was done by adding equal volumes of cell lysate and 2X β-Galactosidase substrate (Promega). The samples were incubated at 37°C for 5 min or until the appearance of a yellow color, which indicated the formation of the chromogenic product. The reaction was stopped by adding 1M $Na_2CO_3$ to a final concentration of 500 mM and the amount of product formed was measured spectrophotometrically in a VersaMax ELISA reader (Molecular Devices, Sunnyvale, CA, USA).

### *2.4.3. Growth curve (proliferation) assay*

AW13516-shNPM1 untreated cells were seeded at an initial density of $2.5\times10^4$ cells/well of a 24-well plate, in duplicates. The NPM1 shRNA expression in the test samples was induced by treating the cells with 2 μg/ml of doxycycline (Dox) post-24 h of seeding, and every 24 h until the completion of the experiment. The control samples were left untreated (UT). The cell growth was monitored by counting the number of cells every 24 h for 6 days. The number of cells in each well was then plotted in the form of a line graph.

### *2.4.4. Colony formation assay*

AW13516-shNPM1 untreated cells were seeded at an initial density of 500 cells in 100 mm dishes. The NPM1 shRNA expression in the test samples was induced by treating the cells with 2 μg/ml of doxycycline (Dox) post-24 h of seeding, and every 24 h until the completion of the experiment. The control samples were left untreated (UT). After 10 days of treatment, the dishes were washed with PBS and the cells were fixed with methanol for 10 min at RT. The colonies were stained using a 0.2% v/v solution of crystal violet for visualization and counting. The images of the stained dishes were captured using a regular photographic camera.

### *2.4.5. Wound closure assay*

AW13516-shNPM1 untreated cells were seeded in 30 mm dishes and cultured till the monolayer was confluent. A scratch in the monolayer was made along the diameter of the dish using a 200 µl pipette tip. The medium was changed. While the control sample was left untreated (UT), the test sample was treated with 2 µg/ml doxycycline (Dox). Real-time imaging was carried out for 24 h with images captured every 10 min using an Axiovert 200M microscope (Carl Zeiss AG).

To measure wound closure contributed majorly by the migratory property of the cells, the cells were pre-treated with 5 µg/ml of mitomycin C (Sigma) for 2 h before creating the scratch. Subsequent steps in the experiment were the same as described above.

The captured images were analyzed using AxioVision software (release 4.8.2).

### *2.4.6. Inhibitor treatment*

The HCT116 $p53^{+/+}$ cells were treated with 2.5, 5 or 10 µM of the compound Nutlin-3a (Sigma) for 6, 12 or 24 h, as indicated in the corresponding experiments, after which total protein or RNA was extracted from cells for further analyses.

### *2.4.7. Total RNA extraction*

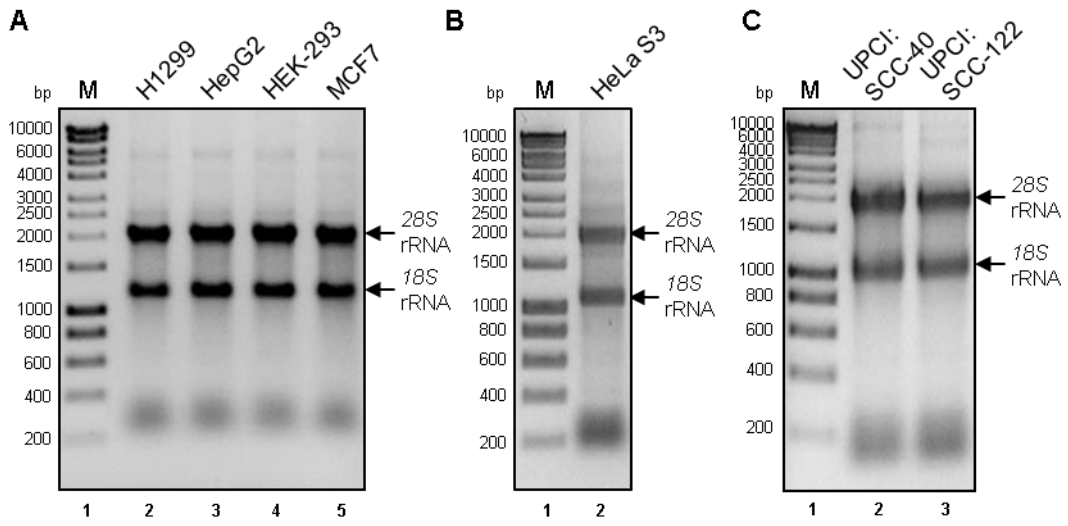
#### *2.4.7.1. Total RNA extraction from cell lines for general experiments*

Cells were cultured in 30 mm or 60 mm dishes as per the requirement of cell number and were harvested at about 90% confluency, in TRIzol reagent (Invitrogen). Lysates were stored at −80°C until RNA isolation which was carried out as per the manufacturer's protocol. All subsequent steps were performed using DEPC-treated RNase free plasticware, and the working area and equipment were decontaminated for the presence of RNase using RNase AWAY Decontamination Reagent (Ambion). The cell lysate prepared in 1 ml of TRIzol was mixed with 200 µl of chloroform and was vigorously shaken for 15 s, followed by incubation at RT for 3 min and subsequent centrifugation at 12000 g, 4°C for 15 min. The upper aqueous phase was collected (~ 500 µl per sample) was collected in a fresh tube and mixed with 500 µl of isopropanol (Sigma), followed by incubation at RT for 10 min. The samples were centrifuged at 12000 g, 4°C for 10 min. The supernatant was discarded

and the RNA pellet was washed with 75% v/v ethanol (~ 1 ml per sample) (by vortexing for a few seconds to dislodge the pellet). The samples were centrifuged at 7500 g, 4°C for 5 min, the supernatants were discarded and the RNA pellets were air-dried. The air-dried RNA pellets were dissolved in 30 – 50 µl of RNase-free water by incubation at 55 – 60°C for 10 min and quantified using NanoDrop (ND1000 spectrophotometer, Thermo Scientific, Waltham, MA, USA).

As an optional step when required, 10 µg RNA was digested with DNase I (New England Biolabs) at 37°C for 20 – 30 min. The enzyme was inactivated with 5 mM of EDTA, pH 8, and heat treatment at 75°C for 20 min. The RNA was re-precipitated with 0.3 M sodium acetate, pH 5.2, followed by the addition of 2.5 times volume of ethanol and incubating at −80°C overnight. The precipitated RNA was recovered after centrifugation at maximum speed (~13000 rpm), 4°C for at least 30 min and finally recovered after a 70% v/v ethanol wash. The DNase I-digested RNA pellet, after air-drying, was finally dissolved in 10 µl of RNase-free water by incubation at 55 – 60°C for 10 min and quantified using NanoDrop (ND1000 spectrophotometer, Thermo Scientific).

The quality of total RNA samples was checked in a 0.8% w/v non-denaturing agarose gel for the presence of distinct bands of 28S and 18S rRNA (Figure 2.12A – F). RNA samples were stored at −80°C.



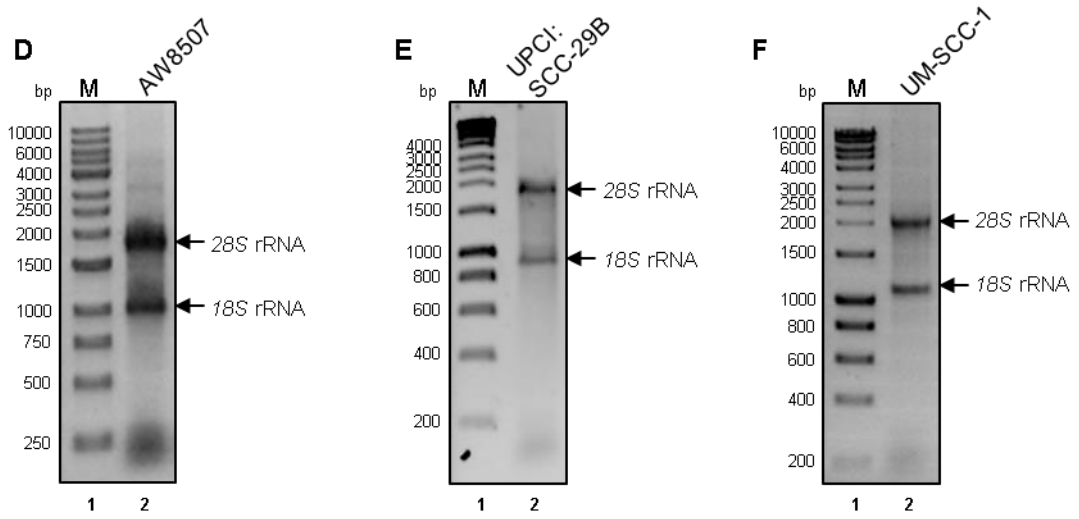



**Figure 2.12. Profile of total RNA isolated from cell lines:** (A – F) Profile of total RNA isolated from cell lines as indicated after non-denaturing agarose gel electrophoresis. Lanes 1 of each figure denotes the DNA molecular weight marker (M). Prominent *28S* and *18S* rRNA bands shown are indicative of the integrity and good quality of the isolated RNA.

### *2.4.7.2. Total RNA extraction from mice tissues*

Mouse organs were freshly collected in RNase-free tubes. The tissue of suitable size (about 50 – 100 mg) was cut out and homogenized in 1 ml of TRIzol reagent (Invitrogen) using a homogenizer (Omni International Tissue Master 125, Marietta, GA, USA). The homogenization was performed at 4°C in a relatively RNase-decontaminated area and using RNase-decontaminated equipment and plasticware. The tissue homogenates were then mixed with chloroform (200 µl per 1 ml of homogenate). Subsequent steps of total RNA isolation were the same as described for the cell lines mentioned above. The integrity of RNA was checked in a 0.8% w/v non-denaturing agarose gel for the presence of distinct bands of 28S and 18S rRNA (Figure 2.13A – B). RNA samples were temporarily stored at −80°C.

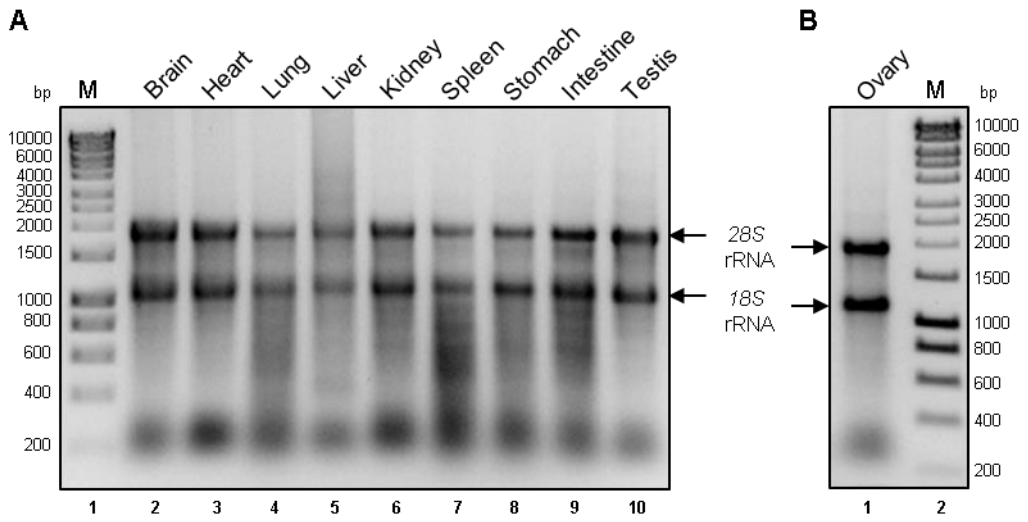



**Figure 2.13. Profile of total RNA isolated from mouse tissues:** (A – B) Profile of total RNA isolated from different organ tissues of a (A) male and (B) female mouse as indicated, after non-denaturing agarose gel electrophoresis. The DNA molecular weight marker (M) is shown. Prominent *28S* and *18S* rRNA bands shown are indicative of the integrity and good quality of the isolated RNA.

### *2.4.7.3. Total RNA extraction from cell lines for RNA-seq and its quality control*

To identify the differentially expressed genes by doxycycline-induced shRNA-mediated knockdown of NPM1 in the human oral cancer cell line AW13516, cells were cultured in biological quadruplicates, in 60 mm dishes under untreated (UT, control) and doxycycline-treated (Dox, 1 µg/ml, NPM1 knockdown) conditions for 6 days. Cells were harvested and RNA was isolated using the GenElute Mammalian Total RNA Miniprep Kit (Sigma) following the manufacturer's protocol. Briefly, prior to harvesting the cells, 10 µl of β-mercaptoethanol (Sigma) were added per 1 ml of the Lysis Solution to inactivate the RNases. Cells from each 60 mm dish were lysed directly in the culture vessels, in 500 µl of this reconstituted Lysis Solution and stored temporarily at −80°C. For RNA isolation, the lysate was loaded in the filtration column and centrifuged at maximum speed (12000 – 16000 g) for 2 min to remove the cell debris and sheared DNA. The filtrate was mixed thoroughly with equal volumes of 70% v/v ethanol solution, loaded into the binding column in volumes of 700 µl and centrifuged at maximum speed for 15 s. The flow-through was discarded and the column was washed in 250 µl of Wash Solution 1 by centrifugation at maximum speed for 15 s. To the column, 80 µl of DNase I/Digest Buffer (10 µl of DNase I plus 70 µl of DNase I Digest Buffer, invert-mixed prior to the step; Sigma) were added and incubated at RT for 15 min. The column was then washed with 250 µl of Wash Solution

1 by centrifugation at maximum speed for 15 s. The column was then washed with 500 µl of ethanol containing Wash Solution 2 and centrifuged at maximum speed for 15 s. This step was repeated a second time with centrifugation for 2 min followed by a free spin for 1 min to remove any residual ethanol. The RNA was eluted in 50 µl of Elution Buffer by centrifugation at maximum speed for 1 min and collected in an RNase-free tube. The quantity and quality of the RNA were assessed by estimating the concentration using the NanoDrop 1000 spectrophotometer (Thermo Scientific) and electrophoresing the samples in a 0.8% w/v non-denaturing agarose gel (Table 2.1 and Figure 2.14). The samples were shipped to Quick Biology Inc, Pasadena, CA, USA for RNA-seq assay and had passed the quality control test (Figure 2.15A – D and Table 2.2).

| Sl. No. | Sample Name | $A_{260/280}$ | $A_{260/230}$ | Conc. (ng/µl) |
|---|---|---|---|---|
| 1 | Untreated (UT, replicate – 1) | 2.06 | 2.12 | 1088.5 |
| 2 | Doxycycline-treated (Dox, replicate – 1) | 2.06 | 2.16 | 862.6 |
| 3 | Untreated (UT, replicate – 2) | 2.07 | 2.24 | 768.4 |
| 4 | Doxycycline-treated (Dox, replicate – 2) | 2.06 | 2.00 | 1322.1 |

**Table 2.1. Quality control determinants of RNA samples for RNA-seq as measured by spectrophotometer NanoDrop 1000 (Thermo Scientific):** $A_{260/230}$ (absorbance at 260 nm wavelength by absorbance at 280 nm) and $A_{260/230}$ (absorbance at 260 nm wavelength by absorbance at 230 nm) values of > 2 indicate good quality of RNA.

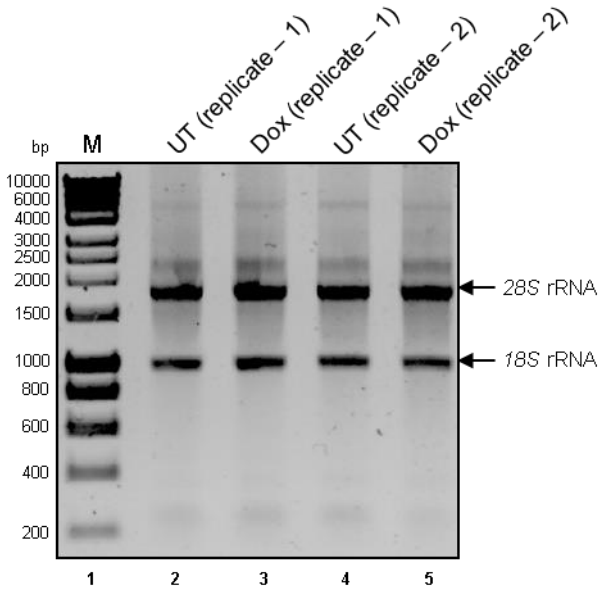


**Figure 2.14. Profile of total RNA used for RNA-seq assay:** Profile of total RNA isolated from AW13516-shNPM1-luc+ cells, with (Dox) and without (UT) doxycycline treatment as indicated, after non-denaturing agarose gel electrophoresis. The DNA molecular weight marker (M) is shown. Prominent *28S* and *18S* rRNA bands shown are indicative of the integrity and good quality of the isolated RNA.

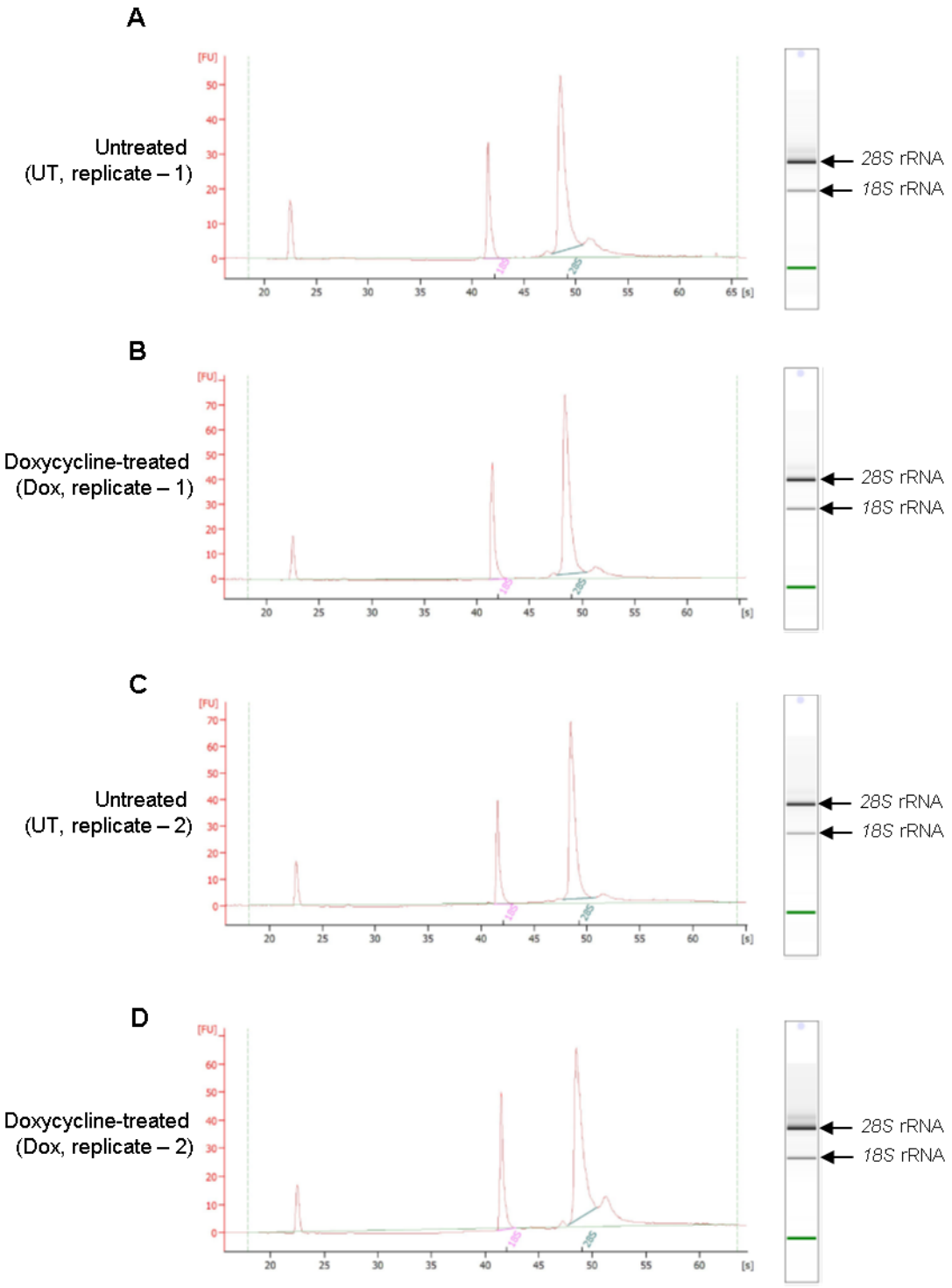



**Figure 2.15. Electropherograms of total RNA used for RNA-seq assay:** (A – D) Electropherograms of total RNA isolated from biological duplicates (replicate-1 and replicate-2) of

untreated (UT, control) and doxycycline-treated (Dox, NPM1 knockdown) AW13516-shNPM1-luc+ cells.

| Sl. No. | Sample Name | rRNA ratio (*28S*/*18S*) | RNA integrity number (RIN) | Conc. (ng/µl) |
|---|---|---|---|---|
| 1 | Untreated (UT, replicate – 1) | 2.5 | 10 | 143 |
| 2 | Doxycycline-treated (Dox, replicate – 1) | 2.4 | 10 | 163 |
| 3 | Untreated (UT, replicate – 2) | 2.6 | 10 | 146 |
| 4 | Doxycycline-treated (Dox, replicate – 2) | 2.2 | 10 | 200 |

**Table 2.2. Quality control determinants of RNA samples for RNA-seq as measured by Agilent 2100 Bioanalyzer System:** rRNA ratio (*28S*/*18S*) of > 2 and RNA integrity number (RIN) in the range of 7 – 10 indicate good quality of RNA.

### *2.4.8. cDNA synthesis*

For cDNA synthesis of poly-A tailed mRNAs, 2 µg of total RNA was used. 10 µl reaction mixture consisting of 2 µg RNA, 1 mM dNTPs and 3.5 µM oligo dT primers (Sigma) was incubated at 70°C for 10 min. The mixture was then chilled on ice, followed by addition of Moloney Murine Leukemia Virus reverse transcriptase (MMLV-RT) Buffer (Sigma) to a final concentration of 1X, 1 µl MMLV reverse transcriptase enzyme (Sigma), and 0.5 µl RNase inhibitor cocktail, RNase Out (Invitrogen) in a final volume of 20 µl. cDNA synthesis carried out at 37°C for 50 min, followed by a final denaturation and enzyme inactivation at 94°C for 10 min. The cDNAs were stored at −20ºC until further use.

### *2.4.9. mRNA expression analysis*

#### *2.4.9.1. Semi-quantitative PCR*

0.5 µl of respective cDNA (half diluted in water) was used per PCR reaction mixture containing 0.5 µM each of the forward and reverse primers for the specific genes (Appendix Table A.5), 0.2 U of Phusion HF DNA Polymerase (New England Biolabs), 0.2 mM dNTPs, in 1X Phusion-HF Buffer (New England Biolabs). Reactions were carried out in Bio-Rad MJ-Mini Personal Thermal Cycler. Reaction program was set as follows: Initial

denaturation at 98°C for 30 s; PCR cycle consisting of denaturation at 98°C for 10, annealing for 30 s at temperature standardized for specific primer set, extension at 72°C for 15 s; final extension at 72°C for 10 min. Cycle numbers were standardized for specific PCR products. The expression level was analyzed from the signal intensity of the respective PCR product electrophoresed in a 1% w/v agarose gel and stained by ethidium bromide. β-actin was used as the internal control.

#### *2.4.9.2. Reverse transcriptase quantitative PCR (RT-qPCR)*

Quantitative real-time PCR was carried out using either the Bio-Rad CFX96 instrument or the StepOnePlus Real-Time PCR Detection System (Applied Biosystems (ABI), Waltham, MA, USA).

In the case of Bio-Rad CFX96 instrument, mRNA expression analysis was carried out in the following way: 1µl of respective cDNA (half diluted in water) was used per PCR reaction mixture containing 0.2 µM each of forward and reverse primers for specific genes in 1X KAPA SYBR FAST Master Mix Universal (Kapa Biosystems, Wilmington, MA, USA). The reaction program was followed as per the manufacturer's protocol and standardized for each set of gene primers used.

In the case of the StepOnePlus Real-Time PCR Detection System, the reactions were set up using either KAPA SYBR FAST qPCR Master Mix (Kapa Biosystems), Power SYBR Green PCR Master Mix (ABI), or the TB Green Premix Ex Taq II (Tli RNase H Plus) (Takara Bio Inc., Kusatsu, Shiga Prefecture, Japan) as per manufacturer's protocol and using 2 µl of appropriately diluted DNA (1 in 40 for mRNA expression analysis).

Fold changes were calculated using the formula $2^{-(Ct_{test}-Ct_{control})}$ considering β-actin or *18S* as the housekeeping genes. The sequences of the primers are given in Appendix Table A.5. The sensitivity and specificity of the primers were determined by melt curve analysis and gel profile of the amplicons.

## *2.5. Antibodies*

### *2.5.1. In-house raised antibodies*

Polyclonal antibodies against human NPM1 acetylated at sites K229, K230, and human NPM2 were generated in 3 months old New Zealand white rabbits following the standard

protocol for priming and booster doses. Before immunization, about 5 ml of pre-immune bleed was collected from each rabbit. Immunization was done to two rabbits simultaneously for the generation of one type of antibody. Each rabbit was primed with 250 µg of the immunogen (peptide or protein in a total volume of 500 µl of PBS) in Freund's Complete Adjuvant (500 µl) (Genei, Bangalore, India) prepared as an emulsion (1 ml) and boosted after every 2 weeks with 100 – 125 µg of the immunogen (peptide or protein in a total volume of 500 µl of PBS) in Freund's Incomplete Adjuvant (500 µl) (Genei). The injections were done subcutaneously to the rabbits by a trained professional at the Animal House facility of JNCASR. Test bleeds of 2 – 3 ml were periodically collected and tested for the presence of the specific antibody in the sera. Upon getting a specific antibody, a major bleed of 10 – 30 ml was taken. For carrying out further booster immunizations to the rabbits from where the major bleeds were taken, a recuperation period of at least one week was given. The collected blood was stored at 4°C for about 24 h for separation of the serum from the other components of the blood, followed by centrifugation at 2000 rpm, 4°C for 10 min. The clear serum was collected, aliquoted and stored at −20°C for future use. Smaller aliquots of the crude sera (about 500 µl) were used to purify the antibody by peptide affinity purification or Protein G-Sepharose (Amersham Biosciences, Little Chalfont, UK) based affinity purification (Section 2.6.9) for use in experimental techniques.

#### *2.5.1.1. Generation and characterization of polyclonal anti-AcNPM1 (K229, K230) antibody raised in rabbit*

A peptide containing acetylated Lys229 (K229) and Lys230 (K230) residues was designed, custom synthesized and conjugated with Keyhole Limpet Hemocyanin (KLH) (Genemed Synthesis Inc., San Antonio, TX, USA). The sequence of the peptide is KLH-C-KGQESFK(Ac)K(Ac)QEKTP (residues 223 to 235 of NPM1). The lyophilized peptide was reconstituted in PBS to make a usual concentration of 1 – 2 µg/µl. The immunization was carried out as described above. Specific antibody against AcNPM1 was obtained after the third booster to one of the two rabbits (rabbit # 340) and a major bleed was taken at that time. Characterization of the purified antibody was done by western blotting, dot-blot assays, immunofluorescence, and immunoprecipitation experiments to confirm the specificity of the antibody (Figure 2.16A – B and Senapati P, Ph.D. thesis, 2014).

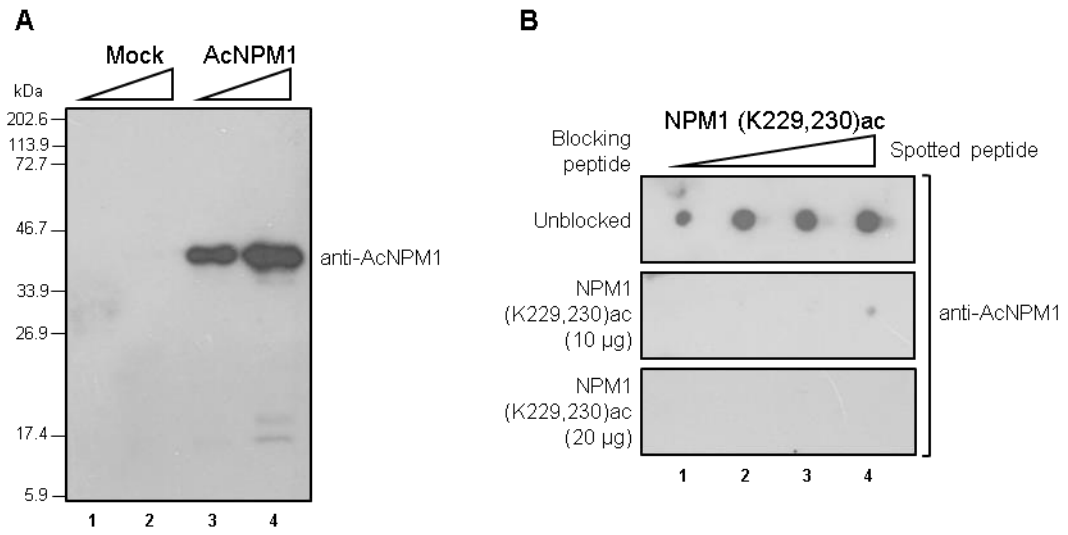
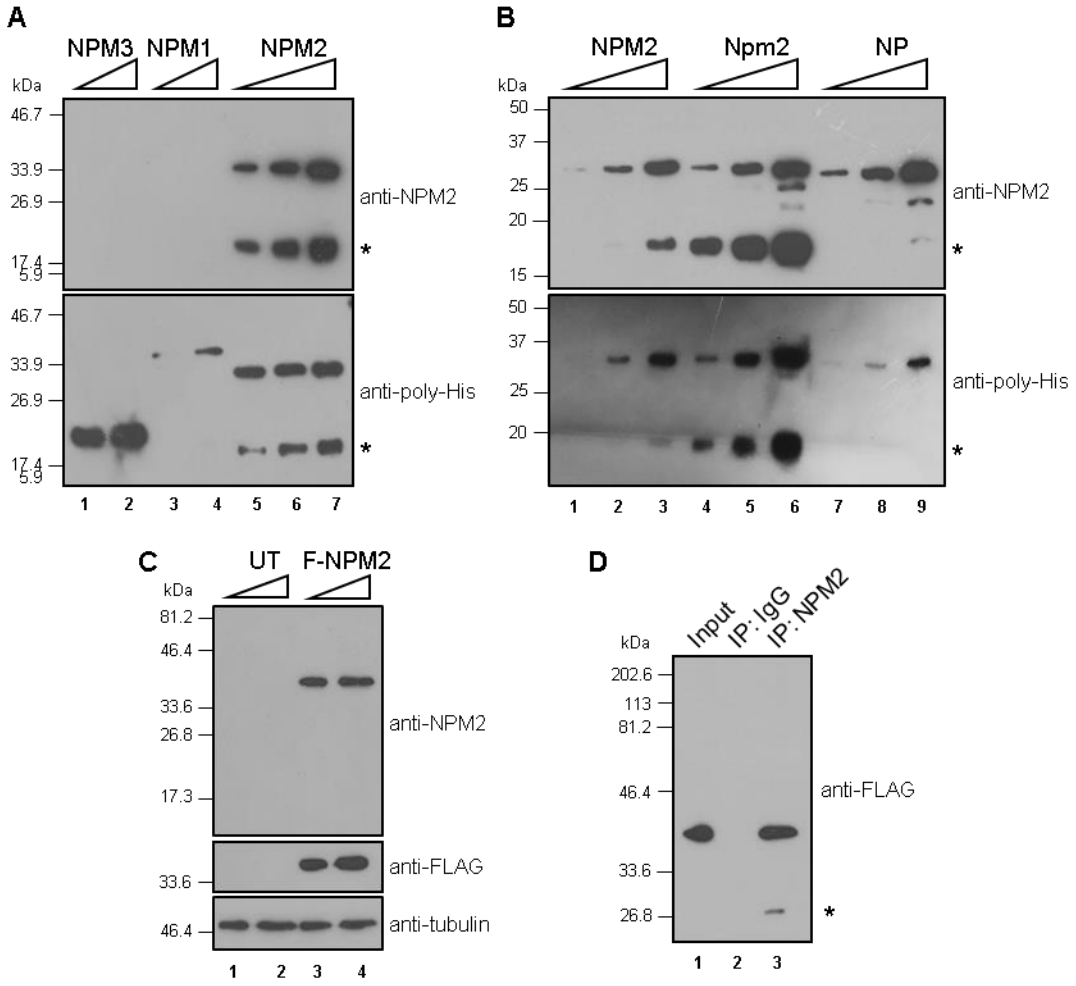



**Figure 2.16. Basic characterization of the anti-AcNPM1 antibody:** (A) Western blot analysis with 500 ng (lanes 1 and 3) and 1 µg (lanes 2 and 4) of mock acetylated (without p300 enzyme, lanes 1 – 2) or acetylated (with p300 enzyme, lanes 3 – 4) NPM1 using anti-AcNPM1(K229, K230) antibody. (B) Dot-blot analysis with 25, 50, 100 and 200 ng of NPM1 (K229, K230)ac peptide spotted, using anti-AcNPM1(K229, K230) antibody pre-blocked by 10 and 20 µg of NPM1 (K229, K230)ac peptide as indicated. Results show that the anti-AcNPM1 antibody is specific to NPM1 acetylated at sites K229, K230.

### *2.5.1.2. Generation and characterization of polyclonal anti-NPM2 antibody raised in rabbit*

Polyclonal antibody against human NPM2 was raised in a rabbit immunized with recombinantly expressed and affinity-purified $His_6$-tagged NPM2 peptide fragment from transformed *E. coli* BL21 cells (Figure 2.19E). This fragment of NPM2 protein (sequence: LEGKQSCRLLLHTICLGEKAKEEMHRVEILPPANQEDKKMQPVTIASLQA) shares a 63% sequence identity with mouse Npm2 and hence could possibly cross-react with the mouse ortholog. The protein was dialyzed in PBS after purification for injecting in the rabbits. The immunization procedure was essentially the same as that described in Section 2.5.1. Priming was done with 350 – 500 µg of immunogen injected in each rabbit while booster doses were given using ~ 200 µg of the immunogen. A specific antibody against NPM2 was obtained after the third booster to one of the two rabbits (rabbit # 359) and after the fourth booster to the other rabbit (rabbit # 358). Major bleeds were taken from these rabbits at these time points. Crude serum from rabbit # 358 was used for peptide affinity-based purification of anti-NPM2 antibody.

The specificity of this antibody against human NPM2 protein was checked by western blot, immunofluorescence, and immuno-pull-down assays. The antibody was specific to NPM2 and was not cross-reactive to NPM1 and NPM3 (Figure 2.17A). It cross-reacted with recombinant mouse and *Xenopus* Nucleoplasmin proteins (Figure 2.17B). It recognized a specific band at the expected position in lysates of HEK-293 cells transfected with the 3xFLAG-NPM2 plasmid (Figure 2.17C). The antibody could pull down FLAG-tagged NPM2 in 3xFLAG-NPM2 transfected HEK-293 cells (Figure 2.17D) and also immunostain FLAG-tagged NPM2 in the same (Figure 2.17E). The results were also compared with a commercially procured polyclonal antibody raised in mouse (Sigma) which showed similar profiles (Figure 2.18).



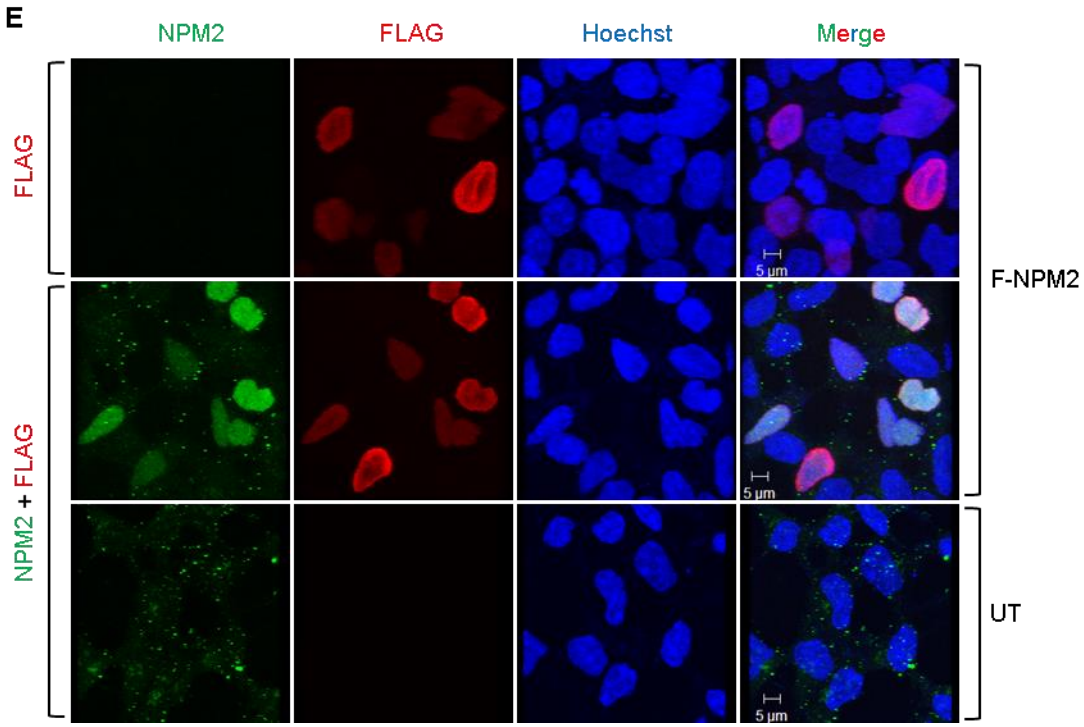



**Figure 2.17. Characterization of the anti-NPM2 antibody:** (A) Western blot analysis taking increasing doses of $His_6$-tagged NPM proteins, 100 and 200 ng of NPM3-$His_6$ (lanes 1 – 2), 100 and 200 ng of NPM1-$His_6$ (lanes 3 – 4), and 50, 100 and 200 ng of $His_6$-NPM2 (lanes 5 – 7). The upper panel shows western blot with anti-NPM2 and the lower panel shows western blot with anti-poly-His antibodies respectively. * denotes degradation product of NPM2. (B) Western blot analysis taking increasing doses (10, 20 and 40 ng) of $His_6$-tagged Nucleoplasmin ortholog proteins as indicated. NPM2: human Nucleoplasmin, Npm2: mouse Nucleoplasmin, and NP: *Xenopus* Nucleoplasmin. The upper panel shows western blot with anti-NPM2 and lower panel shows western blot with anti-poly-His antibodies respectively. * denotes degradation product of NPM2/Npm2. (C) Western blot analysis taking increasing amounts of untransfected (UT, lanes 1 – 2) and 3xFLAG-NPM2-transfected (F-NPM2, lanes 3 – 4) HEK-293 whole-cell lysates. The upper panel shows western blot with anti-NPM2, the middle panel with anti-FLAG, and the lower panel shows western blot with anti-tubulin antibodies respectively. (D) Western blot analysis with anti-FLAG antibody after immuno-pull-down (IP) of FLAG-tagged NPM2 from transfected HEK-293 cells by anti-NPM2 antibody (2 µg). Input (lane 1) was 2% of the lysate used for the pull-down and IP with pre-immune IgG (lane 2) was taken as a negative control. * denotes the position of IgG light chain. (E) Immunofluorescence analysis showing nucleoplasmic localization of NPM2 after transfection of HEK-293 cells with 3xFLAG-NPM2 plasmid. NPM2 (green) and FLAG (red) signals colocalize in the transfected cells (F-NPM2, middle panel). Transfected (F-NPM2) cells stained without the NPM2 primary antibody (upper panel) and untransfected (UT) cells stained with both anti-NPM2 and anti-FLAG antibodies (lower panel) were kept as negative controls. Nuclei were stained with Hoechst. Scale bar is 5 µm.

#### *2.5.1.3. anti-NPM1 monoclonal antibody*

anti-NPM1 mouse monoclonal antibody was previously generated in-house (hybridoma clone 28M1) in association with Abexome Biosciences, Bangalore, India and characterized in the laboratory (Senapati P, Ph.D. thesis, 2014).

#### *2.5.1.4. anti-GAPDH antibody*

Rabbit polyclonal anti-GAPDH antibody was previously generated and characterized in the laboratory (Gadad et al. 2011a).

### *2.5.2. Commercial antibodies*

The various commercial antibodies used in this study and their relevant information are listed in Appendix Table A.6. The validation and characterization of the commercially procured anti-NPM2 antibody (Sigma) is described in the following section.

#### *2.5.2.1. anti-NPM2 raised in mouse and its characterization*

Polyclonal anti-NPM2 antibody raised in mouse was procured from Sigma (Cat No. SAB1400381). The antibody was characterized along with the in-house raised anti-NPM2 antibody. The antibody was specific to NPM2 and did not cross-react with NPM1 or NPM3 (Figure 2.18A). It showed weak cross-reaction towards mouse Nucleoplasmin but no cross-reaction at the doses tested towards *Xenopus* Nucleoplasmin (Figure 2.18B). The antibody also did not recognize the degradation product of human and mouse Nucleoplasmin proteins (Figure 2.18A – B). The antibody recognized FLAG-tagged NPM2 in the whole-cell lysates prepared from 3xFLAG-NPM2 HEK293 stable cells (Figure 2.18C). Both in-house raised and commercial anti-NPM2 proteins could pull-down NPM2 in the lysates prepared from 3xFLAG-NPM2 HEK293 stable cells (Figure 2.18D – E), as well as show characteristic nucleoplasmic staining for NPM2 in these cells (Figure 2.18F). This commercial anti-NPM2 antibody (Sigma) was more sensitive in its reactivity towards human NPM2 than the in-house raised anti-NPM2 antibody (Figure 2.18D – E, compare the 'Input' lanes) and could detect as less as 10 ng of recombinant human NPM2 by western blotting (Figure 2.18B).

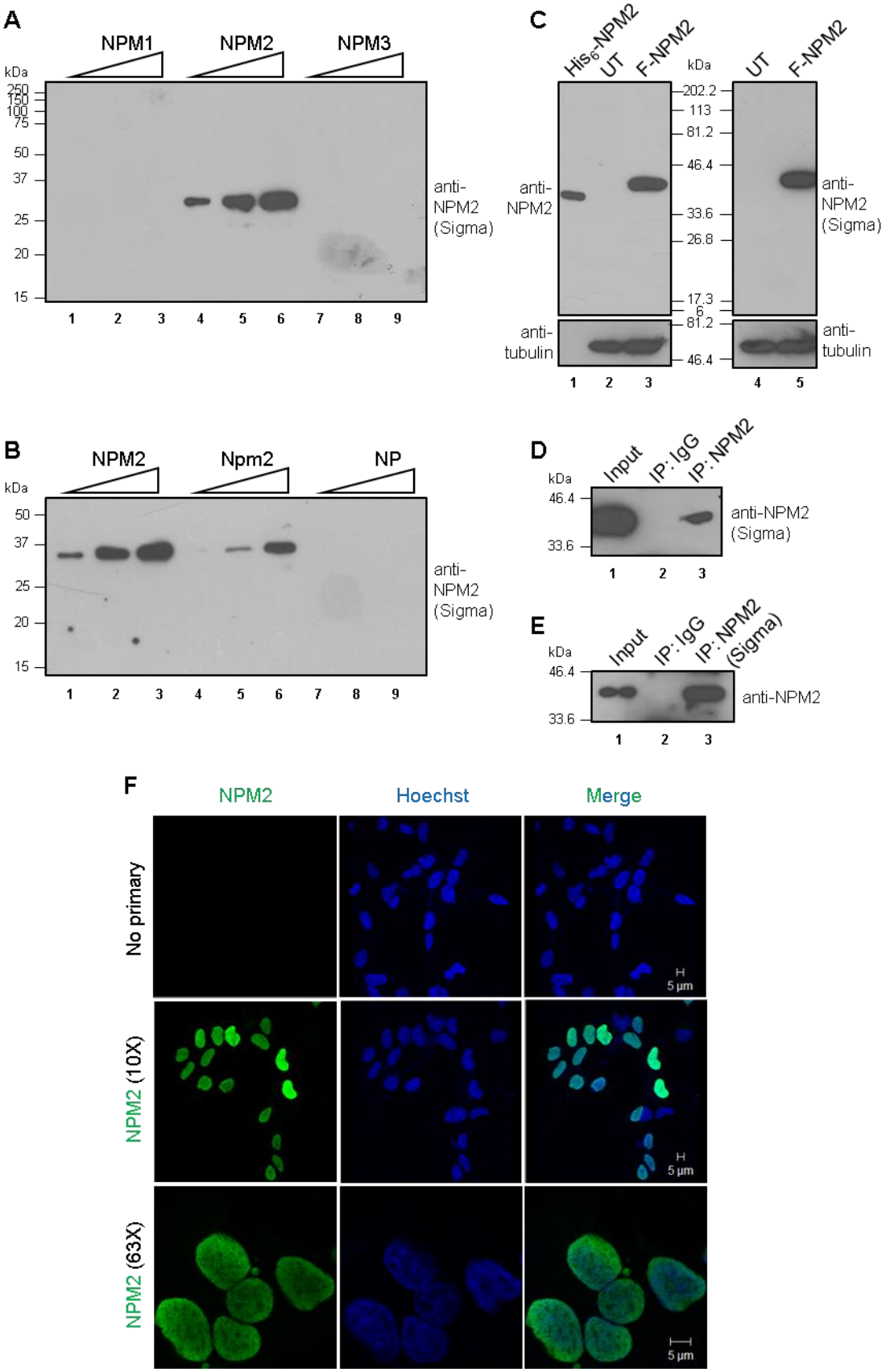



**Figure 2.18. Characterization of the anti-NPM2 antibody (Sigma, Cat. No. SAB1400381):** (A) Western blot analysis taking increasing doses (25, 50, and 75 ng) of $His_6$-tagged NPM proteins as

indicated, using anti-NPM2 (Sigma) antibody. (B) Western blot analysis taking increasing doses (10, 20 and 40 ng) of $His_6$-tagged Nucleoplasmin ortholog proteins as indicated. NPM2: human Nucleoplasmin, Npm2: mouse Nucleoplasmin, and NP: *Xenopus* Nucleoplasmin. (C) Western blot analysis with in-house raised anti-NPM2 antibody (upper panel, lanes 1 – 3) and the commercial anti-NPM2 antibody (Sigma) (upper panel, lanes 4 – 5) as indicated. The $His_6$-NPM2 recombinant protein was taken as the positive control for the western blot (lane 1), untransfected (UT) HEK-293 whole-cell lysate was taken as the negative control (lanes 2 and 4), while whole-cell lysates from 3xFLAG-NPM2 HEK-293 stable cell line (F-NPM2, lanes 3 and 5) serve as the test. The bottom panel shows western blot with anti-tubulin antibody. (D) Western blot analysis with anti-NPM2 antibody (Sigma) after immuno-pull-down (IP) of NPM2 from 3xFLAG-NPM2 HEK-293 stable cells by anti-NPM2 antibody (1 µg). (E) Western blot analysis with anti-NPM2 antibody (in-house raised) after immuno-pull-down (IP) of NPM2 from 3xFLAG-NPM2 HEK-293 stable cells by anti-NPM2 antibody (Sigma, 1 µg). (D – E) Inputs (lane 1) were 2% of the lysates used for the pull-downs and IP with pre-immune IgG (lane 2) was taken as a negative control. (F) Immunofluorescence analysis showing nucleoplasmic localization of NPM2 (green) in 3xFLAG-NPM2 HEK-293 stable cells. Cells stained without the NPM2 primary antibody (upper panel) serve as negative control. The magnified image of the cells (bottom panel) shows the absence of NPM2 staining in the nucleoli. Nuclei were stained with Hoechst. Scale bar is 5 µm.

## *2.6. Protein purification*

### *2.6.1. Purification of recombinant $His_6$-tagged NPM proteins, AP4, POLR2K, PC4, and SNAI1 from E. coli*

The different NPM proteins were purified by the Ni-NTA-based affinity purification method as per protocol previously standardized (Swaminathan et al. 2005; Gadad et al. 2010). Briefly, primary cultures of *E. coli* BL21 (DE3) cells, transformed with the respective plasmids, were grown at 37°C, 180 rpm for about 12 h. All the culture media (LB) were supplemented with 50 µg/ml of kanamycin antibiotic. The secondary culture medium was inoculated with 10% v/v of the overnight-grown primary culture and induced with 0.5 mM of IPTG when its $O.D._{600}$ = 0.3 – 0.5. Cultures were grown for an additional 3 h at the same conditions and harvested by centrifugation at 6000 rpm, 4°C for 10 min. Pellets were temporarily stored at −80°C.

Cell pellets were resuspended in 30 ml of homogenization buffer (20 mM Tris-HCl (Sigma), pH 7.5, 10% v/v glycerol, 0.2 mM Na-EDTA (Sigma), pH 8, 300 mM KCl (Sigma), 0.1% v/v Nonidet P-40 (Sigma), 20 mM imidazole (Sigma), 2 mM PMSF (Sigma), 2 mM β-mercaptoethanol (Sigma)) and lysed on ice by sonication (3 cycles, each with bursts of 3 s with a gap of 5 s for a total time of 3 min at 35% amplitude and a gap of 5 min between two cycles) using a Vibra-Cell ultrasonic liquid processor (Sonics & Materials Inc., Newtown, CT, USA). The supernatant fraction was collected after

centrifugation of the lysate at 12000 rpm, 4°C for 20 – 30 min and incubated with equilibrated Ni-NTA beads (Calbiochem) (about 500 µl of 50% v/v slurry for a litre of secondary culture with high expression of the $His_6$-tagged protein) at 12 rpm, 4°C for 3 h. Beads were collected by centrifugation at 1500 rpm, 4°C for 10 min and washed with wash buffer (20 mM Tris-HCl, pH 7.5, 10% v/v glycerol, 0.2 mM Na-EDTA, pH 8, 300 mM KCl, 0.1% v/v Nonidet P-40, 40 mM imidazole, 2 mM PMSF, 2 mM β-mercaptoethanol). Washed beads were packed in an Econo-column (Bio-Rad) and protein was eluted in elution buffer (20 mM Tris-HCl, pH 7.5, 10% v/v glycerol, 0.2 mM Na-EDTA, pH 8, 100 mM KCl, 0.2% v/v Nonidet P-40, 250 mM imidazole, 2 mM PMSF, 2 mM β-mercaptoethanol). The purity and yield of the protein were analyzed in 12% or 15% (for NPM2 fragment) SDS-PAGE (Figure 2.19A – J). Accordingly, suitable fractions were pooled and dialyzed against BC-100 buffer (20 mM Tris-HCl, pH 7.5, 10% v/v glycerol, 0.4 mM Na-EDTA, pH 8, 100 mM KCl, 0.1% v/v Nonidet P-40, 0.5 mM PMSF, 9.8 mM β-mercaptoethanol) or PBS (137 mM NaCl, 2.7 mM KCl, 10 mM $Na_2HPO_4$, 1.8 mM $KH_2PO_4$) (in case of NPM2 fragment for injecting into rabbits) to remove the imidazole. Concentrations of the proteins were determined by gel estimation taking BSA as standard. Proteins were aliquoted and snap-frozen in liquid $N_2$ and stored at −80°C.

The $His_6$-tagged AP4, POLR2K, PC4, and SNAI1 proteins were also purified following a similar protocol as described above.

N.B. The lower band in the gels showing $His_6$-tagged NPM2 protein is its degradation product which has been confirmed after protein identification through mass spectrometry (Proteomics Facility at Molecular Biophysics Unit, Indian Institute of Science, Bangalore, India) (Section 2.12.1).

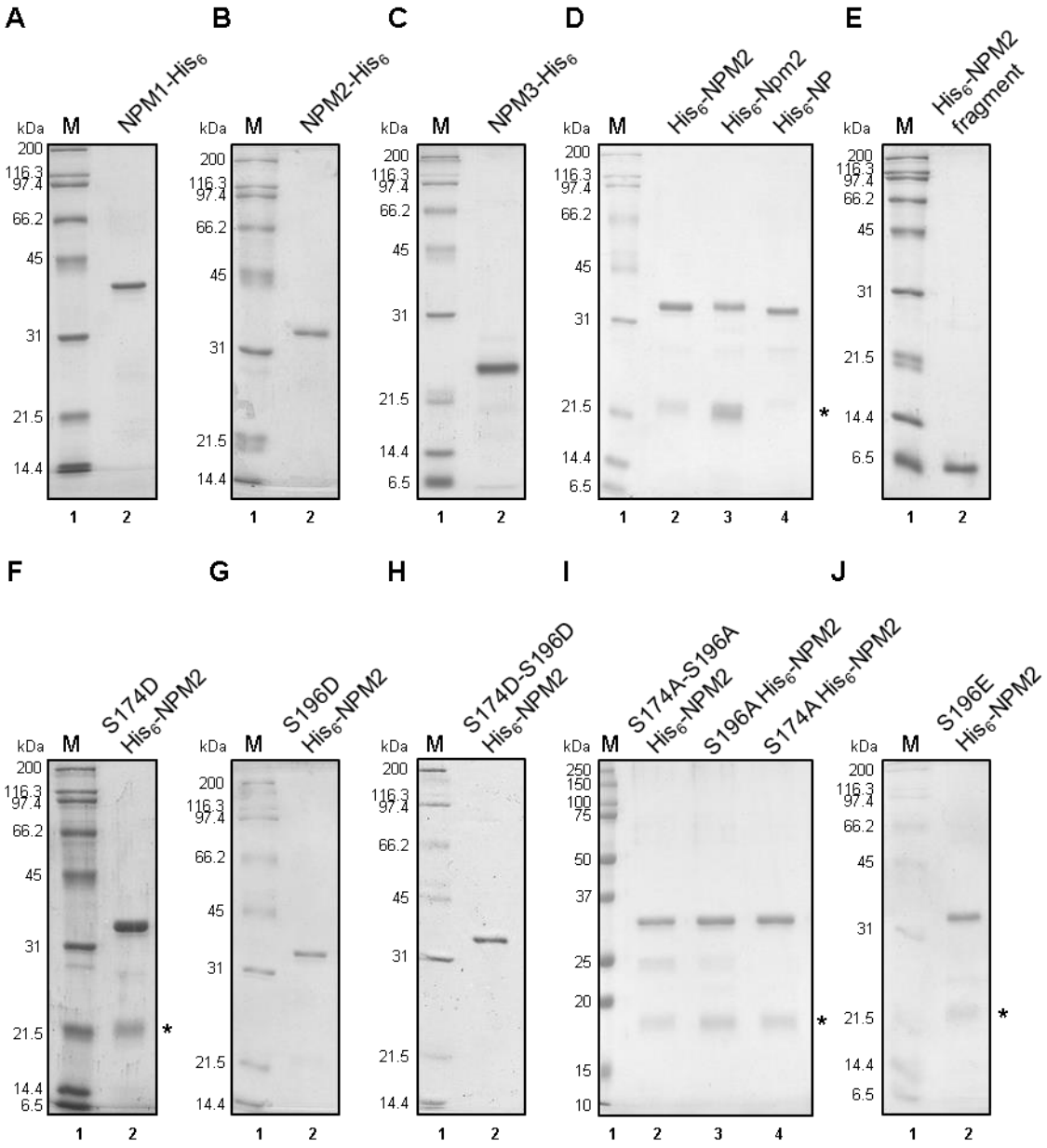



**Figure 2.19. Purification profiles of recombinant $His_6$-tagged NPM proteins:** (A – J) Purification profiles of the different NPM proteins as indicated. NPM1-$His_6$: C-terminal $His_6$-tagged NPM1, NPM2-$His_6$: C-terminal $His_6$-tagged human NPM2, NPM3-$His_6$: C-terminal $His_6$-tagged NPM3, $His_6$-NPM2: N-terminal $His_6$-tagged human NPM2, $His_6$-Npm2: N-terminal $His_6$-tagged mouse Nucleoplasmin, $His_6$-NP: N-terminal $His_6$-tagged *Xenopus* Nucleoplasmin. (F – J) Different point mutants generated in $His_6$-NPM2 are indicated. (A – J) M denotes protein marker and * denotes degradation product of NPM2.

### *2.6.2. Purification of recombinant FLAG-tagged NPM1 and Gal4-VP16 from E. coli*

FLAG-Gal4-VP16 and FLAG-NPM1 were previously purified (Senapati P, Ph.D. thesis, 2014) by immunoaffinity purification with M2 agarose (Sigma) (Kundu et al. 2000). The protocol followed is briefly described below.

*E. coli* BL21 (DE3) competent cells were transformed with the FLAG-Gal4-VP16 or FLAG-NPM1 expression constructs. All media were supplemented with 100 μg/ml ampicillin. The primary culture (100 ml) was grown at 37°C, 180 rpm for about 12 h. The overnight-grown culture was inoculated to make 1 l of secondary culture medium (LB), and grown at the same conditions till the $O.D._{600}$ was about 0.4. The bacterial culture was induced with 0.4 mM IPTG and was grown at the same conditions for an additional 3 h. The cells were then harvested by centrifugation at 6000 rpm, 4ºC for 10 min. The pellet was resuspended in lysis buffer (20 mM Tris-HCl (Sigma), pH 7.4, 20 % v/v glycerol, 2 mM β-mercaptoethanol (Sigma), 0.2 mM EDTA (Sigma), 2 mM PMSF (Sigma), 0.1 % v/v Nonidet P-40 (Sigma) and 300 mM KCl (Sigma)) and sonicated with 4 – 5 cycles of 30 s each at a setting of 32% amplitude using a Vibra-Cell ultrasonic liquid processor (Sonics & Materials Inc.). The supernatant fraction of the lysate was collected after centrifuging it at 16000 rpm, 4°C for 30 min. The supernatant was mixed with FLAG-M2-agarose beads (Sigma) pre-equilibrated with the lysis buffer and incubated on an end-to-end shaker at 12 rpm, 4ºC for 3 h. The beads were collected by centrifugation at 1500 rpm, 4°C for 10 min and were washed three times with the wash buffer (having the same composition as the lysis buffer), followed by four washes with BC-100 (20 mM Tris-HCl, pH 7.4, 20% v/v glycerol, 2 mM β-mercaptoethanol, 0.2 mM EDTA, 0.5 mM PMSF, 0.1% v/v Nonidet P-40, and 100 mM KCl). The washed beads were then packed into a spin column. The protein was eluted with the elution buffer (BC-100 containing 0.2 mg/ml 3xFLAG-peptide (Sigma)). The eluted FLAG-tagged proteins were aliquoted, snap-frozen in liquid $N_2$ and stored at −80ºC.

### *2.6.3. Purification of recombinant untagged Xenopus core histones from inclusion bodies of E. coli*

Recombinant *Xenopus* core histones were purified according to the protocol previously published (Luger et al. 1999), with minor modifications. The pET3d clones expressing untagged histones were a kind gift from Dr. K. Luger. *E. coli* BL21 (DE3) or Rosetta pLysS cells were transformed with the appropriate core histone expression constructs and were grown on LB agar plates containing 100 μg/ml ampicillin (along with 34 μg/ml chloramphenicol if Rosetta pLysS was used) overnight at 37°C. Five different colonies were inoculated into separate small-scale cultures (5 ml) and grown at 37°C, 180 rpm for

4 h or until the $O.D._{600}$ was within 0.3 – 0.6. Glycerol stocks of these cultures were made by mixing 0.5 ml of the growing culture to 0.2 ml of sterilized glycerol and stored at −80°C. The cultures were then induced with 0.2 mM IPTG. Half of each culture was left uninduced. The cultures were allowed to grow for another 2 – 3 h at 37°C, 180 rpm and harvested thereafter by centrifugation at 6000 rpm, 4°C for 5 min. The cell pellets were boiled in 100 µl of 1X SDS gel loading buffer. The induced expression of the histone proteins was analyzed by SDS-PAGE followed by staining of the 12 – 15% gels with Coomassie blue. The glycerol stock of the culture that showed maximum induction of the histone expression was used for restreaking on an LB agar plate containing 100 µg/ml ampicillin and incubated at 37°C overnight. The next morning, 2 – 5 aliquots of 4 m LB media containing 100 µg/ml ampicillin were inoculated with one colony each from the plate and grown at 37°C, 180 rpm for approximately 4 h or until the $O.D._{600}$ reached 0.3. The cultures were combined and inoculated into a 2 l Erlenmeyer flask containing 500 – 750 ml LB with 100 µg/ml ampicillin and grown at the same conditions until the $O.D._{600}$ reached 0.6. The culture was then induced by the addition of 0.2 mM IPTG and grown for another 2 h (for H3 and H4) or 4 h (for H2A and H2B). The cells were subsequently harvested by centrifugation at 6000 rpm at RT (room temperature). The cell pellet was resuspended homogeneously in 25 ml of wash buffer (50 mM Tris-HCl (Sigma), pH 7.5, 100 mM NaCl (Sigma), 1 mM Na-EDTA (Sigma), 1 mM benzamidine (Sigma), 5 mM β-mercaptoethanol (Sigma)), snap-frozen in liquid $N_2$ and stored temporarily at −80°C.

On the day of the purification, the cell suspension was thawed at 37°C in a water bath. The cell suspension was then sonicated by bursts of 3 s each with a gap of 5 s for a total time of 3 min at 35% amplitude using a Vibra-Cell ultrasonic liquid processor (Sonics & Materials Inc.). The sonication cycle was repeated 2 – 3 times with a gap of 5 min between two cycles. The lysate was clarified by centrifugation at 4°C, 23000 g for 10 min, and the supernatant was discarded. The pellet was washed by resuspension and centrifugation thrice in 15 ml of wash buffer containing 1% v/v Triton-X-100 (Sigma). The detergent was removed by repeating the washing thrice with wash buffer (without the Triton-X-100 detergent). The residual pellet containing the inclusion bodies was soaked in 1 ml of DMSO (Sigma) for 30 min at RT followed by unfolding of the proteins in 10 ml of a buffer containing 7 M guanidine HCl (Sigma), 20 mM Tris-HCl, pH 7.5, and 10 mM DTT (Calbiochem), for 1 h at RT under gentle shaking. The mixture was centrifuged at 16000 rpm, RT for 10 min and the supernatant was saved. The pellet was resuspended in 5 ml of buffer containing 7 M

guanidine HCl, 20 mM Tris-HCl, pH 7.5, 10 mM DTT, and unfolding was allowed to proceed for 1 h at RT. The suspension was centrifuged at 16000 rpm, RT for 10 min. The pellet was discarded, and the supernatant was pooled with the previously saved 10 ml of supernatant. The pooled unfolded protein supernatants were dialyzed extensively (overnight) at RT against SAU100 buffer (7 M urea (Sigma), 20 mM sodium acetate (Sigma), pH 5.2, 100 mM NaCl (Sigma), 5 mM β-mercaptoethanol (Sigma), and 1 mM Na-EDTA (Sigma)) using 6 to 8 kDa molecular mass cut-off dialysis bags (Thermo Scientific). The unfolded histone was then incubated with SP sepharose beads (GE Healthcare, Chicago, IL, USA) in SAU 100 buffer for 1 h at RT. The SP-Sepharose beads were then packed into an Econo-column (Bio-Rad) and washed with SAU buffers with increasing salt concentration (100 – 600 mM NaCl). The histones generally get eluted at a salt concentration of 300 mM and above. The elution fractions containing histones were analyzed by SDS-PAGE followed by staining of the gels with Coomassie blue. Suitable fractions were pooled together. The histones were further concentrated using a Centricon concentrator with 3 kDa cutoff (Millipore). The concentration was determined by measuring the UV absorbance of each histone at 276 nm. The following formula was used: $A = \varepsilon cl$ where A is the absorbance of the protein at 276 nm, ε is the molar extinction coefficient of the histone at 276 nm, c is the concentration of the protein and l is the path length of the cuvette. The molar extinction coefficients for the histone proteins are given in Table 2.3. The purity of each histone was checked by SDS-PAGE followed by staining of the gels with Coomassie blue (Figure 2.20). The histones were stored in the denatured form in SAU buffers at −80°C after snap freezing in liquid $N_2$.

| Sl. No. | Histone | Molecular weight | Molar extinction co-efficient (ε)($cm^{-1} M^{-1}$) at 276 nm |
|---|---|---|---|
| 1 | H2A | 13960 | 4050 |
| 2 | H2B | 13774 | 6070 |
| 3 | H3 | 15273 | 4040 |
| 4 | H4 | 11236 | 5040 |

**Table 2.3. Molecular weights and molar extinction coefficients of histone proteins** (Luger et al. 1999).

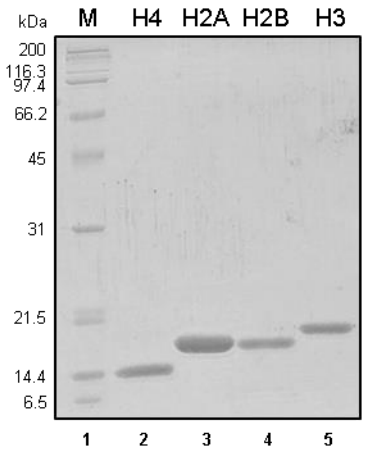



**Figure 2.20. Purification profile of recombinant *Xenopus* core histones:** Gel profile of purified recombinant *Xenopus* core histones H4, H2A, H2B, and H3 are shown. M denotes the protein molecular weight marker.

### *2.6.4. Purification of core catalytic domain of dTopoI (Drosophila Topoisomerase I) from E. coli*

The $His_6$-tagged core catalytic domain of Topoisomerase I from *Drosophila* (Shaiu and Hsieh 1998) was purified by Ni-NTA-based affinity chromatography method from transformed *E. coli* BL21 (DE3) cells. The medium required for the growth of the transformed cells, is LB containing 50 μg/ml of kanamycin. The primary culture (100 ml) was grown overnight at 37°C, 180 rpm. It was inoculated in a final volume of 1 l secondary culture medium and grown at the same conditions until the $O.D._{600}$ reached about 0.5. This culture was induced with 0.42 mM IPTG, and grown at 30ºC for 5 h. The cells were then harvested by centrifugation at 5000 rpm, 4ºC for 10 min. The pellet was resuspended in lysis buffer (50 mM sodium phosphate (Sigma), pH 7.0, 0.5 M NaCl (Sigma), 15% v/v glycerol, 15 mM imidazole (Sigma), 0.1% v/v Nonidet P-40 (Sigma), 0.2 mM PMSF (Sigma) and 0.5 mM benzamidine (Sigma)). The homogenate was sonicated with 4 – 5 cycles of 30 s each at a setting of 32% amplitude using a Vibra-Cell ultrasonic liquid processor (Sonics & Materials Inc.). The lysate was clarified by centrifuging it at 16000 rpm, 4ºC for 30 min. The supernatant fraction was mixed with Ni-NTA beads (Calbiochem) pre-equilibrated with the lysis buffer for protein binding and incubated on an end-to-end shaker at 4°C, 12 rpm for 3 h. The beads were then collected by centrifugation at 1500 rpm, 4°C for 5 – 10 min. The resin was washed five times with the lysis buffer and packed into an Econo-column (Bio-Rad). The protein was eluted using the elution buffer (50 mM sodium phosphate (Sigma), pH 7.0, 0.5 M NaCl (Sigma), 15% v/v glycerol, 500 mM

imidazole (Sigma), 0.1% v/v Nonidet P-40 (Sigma), 0.2 mM PMSF (Sigma) and 0.5 mM benzamidine (Sigma)). The eluted fractions were analyzed by SDS-PAGE followed by Coomassie blue staining. Suitable fractions were then pooled and dialyzed against a dialysis buffer (25 mM HEPES K+ (Sigma) pH 6.7, 0.1 mM EDTA (Sigma), 10% v/v glycerol, 50 mM NaCl, 0.01% v/v Nonidet P-40, 0.5 mM DTT (Calbiochem), 0.2 mM PMSF and 0.5 mM benzamidine). The purified protein was analyzed by SDS-PAGE followed by Coomassie blue staining (Figure 2.21A). The purified protein was aliquoted, snap-frozen in liquid $N_2$ and stored at −80ºC. The activity of the purified Topoisomerase I was assessed by its ability to relax a supercoiled plasmid DNA ($pG_5ML$) by the template relaxation assay as mentioned in Section 2.11.5, and analyzing the plasmids in a 1% w/v agarose gel by agarose gel electrophoresis (Figure 2.21B, lanes 2 and 3). The enzyme was found to be active.

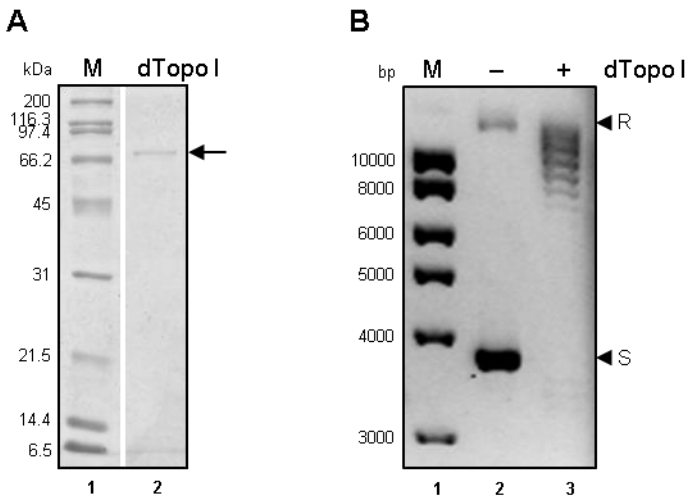



**Figure 2.21. Purification and activity check of core catalytic domain of *Drosophila* Topoisomerase I (dTopo I):** (A) Purification profile of the catalytic domain of dTopo I. The position of the protein is indicated by the arrow. Relevant lanes from the same gel have been cropped and aligned. (B) The topoisomerase activity of purified dTopo I was checked by incubation of a supercoiled DNA template ($pG_5ML$) with (lane 3) or without (lane 2) the enzyme in the template relaxation assay. R denotes the position of the relaxed form and S denotes the position of the supercoiled form of the DNA. M denotes the (A) protein molecular weight marker or (B) DNA ladder.

### *2.6.5. Purification of mouse nucleosome assembly protein (NAP1) from E. coli*

$His_6$-tagged recombinant mouse NAP1 was previously purified (Senapati P, Ph.D. thesis, 2014) by Ni-NTA based affinity chromatography method. The protocol followed is briefly described below.

*E. coli* BL21 (DE3) cells were transformed with the NAP1 expression construct. The medium required for the growth of the transformed cells, is LB containing 100 μg/ml of ampicillin. The primary culture (100 ml) was grown overnight at 37°C, 180 rpm. The overnight-grown culture was inoculated into a final volume of 1 l secondary culture medium and grown at 37ºC, 180 rpm till the $O.D._{600}$ reached 0.4. This culture was induced with 0.4 mM IPTG, and grown at 30ºC, 180 rpm for an additional 3 h. The cells were then harvested by centrifugation at 6000 rpm, 4ºC for 10 min. The pellet was resuspended in lysis buffer (20 mM Tris-HCl (Sigma), pH 7.4, 20% v/v glycerol, 2 mM β-mercaptoethanol (Sigma), 0.2 mM EDTA (Sigma), 2 mM PMSF (Sigma), 0.1% v/v Nonidet P-40 (Sigma), 30 mM imidazole (Sigma) and 300 mM KCl (Sigma)) and sonicated with 4 – 5 cycles of 30 s each at a setting of 32% amplitude using a Vibra-Cell ultrasonic liquid processor (Sonics & Materials Inc.). The lysate was clarified by centrifuging it at 16000 rpm, 4ºC for 30 min. The supernatant was mixed with Ni-NTA beads pre-equilibrated with the lysis buffer for protein binding by incubation at 12 rpm, 4°C on an end-to-end shaker for 3 h. The beads were collected by centrifugation at 1500 rpm, 4°C for 5 – 10 min, washed eight times with the wash buffer (20 mM Tris-HCl, pH 7.4, 20% v/v glycerol, 2 mM β-mercaptoethanol, 0.2 mM EDTA, 2 mM PMSF, 0.1% v/v Nonidet P-40, 30 mM imidazole, and 300 mM KCl), and packed into an Econo-column (Bio-Rad). The protein was eluted with the elution buffer (20 mM Tris-HCl, pH 7.4, 20% v/v glycerol, 2 mM β-mercaptoethanol, 0.2 mM EDTA, 2 mM PMSF, 0.1% v/v Nonidet P-40, 250 mM imidazole and 100 mM KCl). The eluted protein was then dialyzed against dialysis buffer (elution buffer without imidazole). The dialyzed protein was subjected to another round of purification using an anion exchanger (Q-sepharose, GE Healthcare). The dialyzed protein was incubated with Q-sepharose beads pre-equilibrated with the BC-100 buffer previously used, for 3 h at 4°C, 12 rpm on an end-to-end shaker. The beads were then collected by centrifugation as done previously and washed with BC-100 thrice to remove the unbound or loosely bound proteins, and packed into an Econo-column (Bio-Rad). The mouse NAP1 protein was eluted using BC-100 buffers with increasing salt concentration (200 – 500 mM). The eluted fractions were analyzed by SDS-PAGE followed by Coomassie blue staining and suitable fractions were pooled and dialyzed again against BC-100. The purified protein was analyzed by SDS-PAGE followed by Coomassie blue staining. The purified protein was aliquoted, snap-frozen in liquid $N_2$ and stored at −80ºC.

### *2.6.6. Purification of core histones from HeLa S3 cells and rat liver tissue*

The human core histones were purified from the nuclear pellet of HeLa S3 cells. About 2 ml of the nuclear pellet was resuspended in 8 – 10 ml chilled buffer A (100 mM potassium phosphate buffer, pH 6.7, 0.1 mM EDTA (Sigma), 10% v/v glycerol, 0.1 mM PMSF (Sigma), 0.1 mM DTT (Calbiochem) and 630 mM NaCl (Sigma)). The suspension was transferred to a pre-cooled Dounce homogenizer (Wheaton, Millville, NJ, USA) and homogenized using the Pestle B for 30 min (7 – 10 strokes at a time, thrice, with about 5 min gap between two such cycles) on ice at 4°C. The suspension was centrifuged at 14000 rpm, 4°C for 20 min and the supernatant was collected into a fresh tube. The supernatant was mixed with 2 g of swelled Bio-Gel HTP Hydroxyapatite (Bio-Rad), which had been presoaked in 15 ml of 10 mM potassium phosphate buffer pH 6.7, overnight. The mixture was incubated at 4°C, 12 rpm for 3 h on an end-to-end shaker for protein-resin binding. The matrix was collected by centrifugation at 1500 rpm, 4°C for 3 min and was washed thrice with 25 – 40 ml of buffer A. The hydroxyapatite matrix was then packed into a glass Econo-column (Bio-Rad) and was washed overnight with 500 ml – 1 l buffer A. The core histones were eluted in buffer B (same composition as buffer A but with 2 M NaCl instead of 630 mM). Around 15 – 20 elution fractions of 0.5 – 1 ml each were collected and kept on ice or at 4°C temporarily. The fractions were analyzed for the presence of core histones by SDS-PAGE followed by Coomassie blue staining. Suitable fractions were pooled together and dialyzed against a dialysis buffer (1 M HEPES, pH 7.9, 100 mM KCl, 20% v/v glycerol, 0.2 mM EDTA, pH 8.0) or BC-100 (20 mM Tris-HCl, pH 7.5, 20% v/v glycerol, 100 mM KCl and 0.1 mM DTT). The profile (Figure 2.22A) and concentration of the purified histones were analyzed by SDS-PAGE (12 – 15% gel) followed by Coomassie blue staining along with known concentrations of another batch of core histones. The proteins were aliquoted, snap-frozen in liquid $N_2$ and stored at −80ºC.

Core histones from the nuclear pellet of rat liver tissue were previously purified in the laboratory in a similar manner (Modak et al. 2013) (Figure 2.22B).

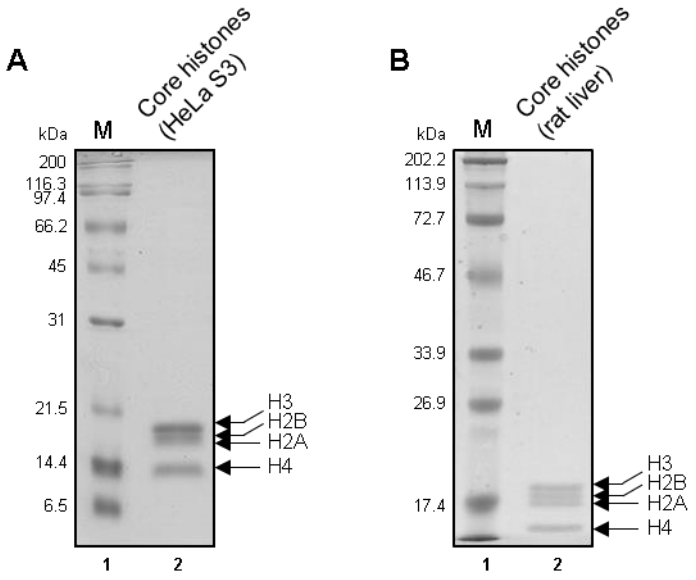



**Figure 2.22. Purification profile of core histones:** Gel profile of purified (A) human core histones (1.5 µg) from Hela S3 and (B) rat core histones (1.5 µg) from rat liver nuclear pellet (lane 2). Positions of the core histones are indicated with arrows and M denotes the protein molecular weight marker.

### *2.6.7. Purification of human lysine acetyltransferase p300 (KAT3B) from Sf21*

Recombinant $His_6$-tagged human p300 protein was purified from Sf21 (*Spodoptera frugiperda*) insect ovary cells infected with baculovirus containing the respective expression construct. The culturing and infection of the cells have been described in Sections 2.3.5 and 2.3.6.

After a specific period of time since the infection of the Sf21 cells, the baculovirus-infected cells in 150 mm dishes were taken out of the BOD incubator at 27 – 28°C, and kept at 4°C for about 30 min. The cells were harvested by gently scraping them off the dishes using a cell scraper (Corning) and collecting them into 50 ml centrifuge tubes. The residual cells in the dishes were further scraped out and collected using PBS mixed with 0.2 mM PMSF (Sigma). The cells were pelleted by centrifugation at 4°C, 2000 rpm for 5 – 10 min. The pellet was resuspended in 10 ml of ice-cold lysis or homogenization buffer (10 mM Tris-HCl (Sigma), pH 7.5, 10% v/v glycerol, 0.1% v/v Nonidet P-40 (Sigma), 2 mM β-mercaptoethanol (Sigma), 0.2 mM PMSF (Sigma), 500 mM NaCl (Sigma), 15 mM imidazole (Sigma) with 50 µg/ml leupeptin and 50 µg/ml aprotinin or 1X protease inhibitor cocktail (Sigma)). The suspension was transferred into a Dounce homogenizer (Wheaton) and homogenized for 30 min by 3 cycles of 10 strokes each using the tight pestle (pestle

B). A gap of 5 min between two cycles was given and the process was carried out on ice at 4°C. The homogenate was clarified by centrifugation at 12000 rpm, 4ºC for 15 min. The supernatant thus obtained was mixed with Ni-NTA resin (Calbiochem) pre-equilibrated with the homogenization buffer and incubated at 12 rpm, 4°C for 2 h on an end-to-end shaker. The beads were collected by centrifugation at 1200 rpm, 4°C for 3 min and washed 8 times (6 times with 10 ml and 2 times with 1 ml of wash buffer) with the wash buffer (10 mM Tris-HCl, pH 7.5, 10% v/v glycerol, 0.2% v/v Nonidet P-40, 2 mM β-mercaptoethanol, 2 mM PMSF, 300 mM NaCl, and 15 mM imidazole). The $His_6$-tagged protein was eluted (three eluates collected) with the elution buffer (10 mM Tris-HCl, pH 7.5, 10% v/v glycerol, 0.1% v/v Nonidet P-40, 2 mM β-mercaptoethanol, 0.2 mM PMSF, 200 mM NaCl, 250 mM imidazole with 50 µg/ml leupeptin and 50 µg/ml aprotinin of 1X protease inhibitor cocktail (Sigma)). The purification profile of the protein was analyzed by SDS-PAGE (8% gel) followed by Coomassie blue staining (Figure 2.23A). The protein was aliquoted, snap-frozen in liquid $N_2$ and stored at −80ºC. The enzymatic (acetyltransferase) activity of the purified p300 was checked by performing *in vitro* filter-binding assay using rat liver core histones (Section 2.11.3). The enzyme was found to be active (Figure 2.23B).

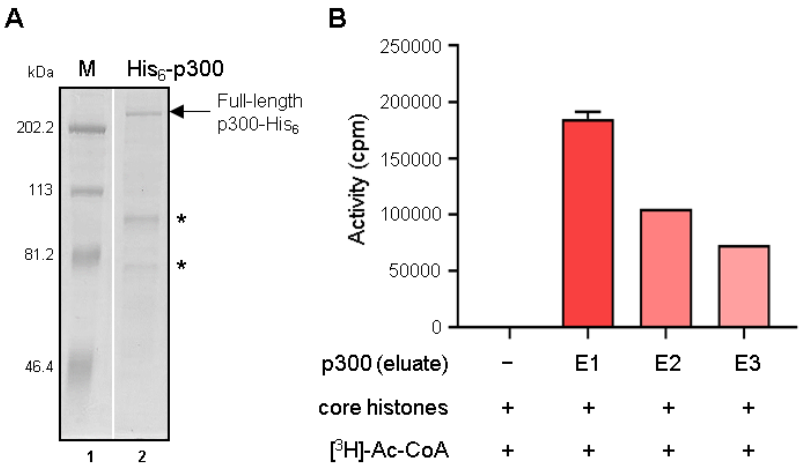



**Figure 2.23. Purification profile and analysis of the enzymatic activity of purified recombinant $His_6$-tagged human p300:** (A) Gel profile of $His_6$-tagged human p300 purified from Sf21 cells after infection with the specific baculovirus. M denotes the protein molecular weight marker. The position of the full-length p300 protein is mentioned and its degradation products are indicated with *. Relevant lanes from the same gel have been cropped and aligned. (B) Bars represent activity of the different eluates (E1, E2, and E3) of the purified p300 protein compared to 'no enzyme control' (Blank), assessed by its ability to transfer tritiated acetyl CoA ([$^3$H]-Ac-CoA) to purified core histones in the *in vitro* filter-binding assay and measured in terms of radioactive emissions per minute (counts per min or cpm). Values corresponding to E1 are mean + SEM from two biological replicates.

### *2.6.8. Purification of Aurora A and Aurora B kinases from Sf21*

$His_6$-tagged recombinant human Aurora Kinase A (AURKA) and Aurora Kinase B (AURKB) were previously purified (Senapati P, Ph.D. thesis, 2014)(Shandilya et al. 2014a) from Sf21 (*Spodoptera frugiperda*) insect ovary cells infected with baculovirus containing the respective expression constructs. The culturing and infection of the cells have been described in Sections 2.3.5 and 2.3.6. The procedure of the harvesting and purification of the recombinant $His_6$-tagged proteins was essentially the same as that followed for p300 purification described in Section 2.6.7. Minor modifications are mentioned below.

The Sf21 cells infected with the respective baculovirus were harvested 60 h post-infection. Post-binding of the protein with the Ni-NTA resin, the beads were washed 5 times with the wash buffer. The purification profile of the protein was analyzed by SDS-PAGE (12% gel) followed by Coomassie blue staining. The enzymatic (kinase) activities of the purified Aurora Kinase A and B were checked by performing an *in vitro* kinase assay (Section 2.11.1) using recombinant histone H3 (Senapati P, Ph.D. thesis, 2014) or the known non-histone substrate of Aurora Kinase that is NPM1 (Shandilya et al. 2014a) (Figure 2.24) in presence of radiolabelled [$\gamma$-$^{32}$P]-ATP.

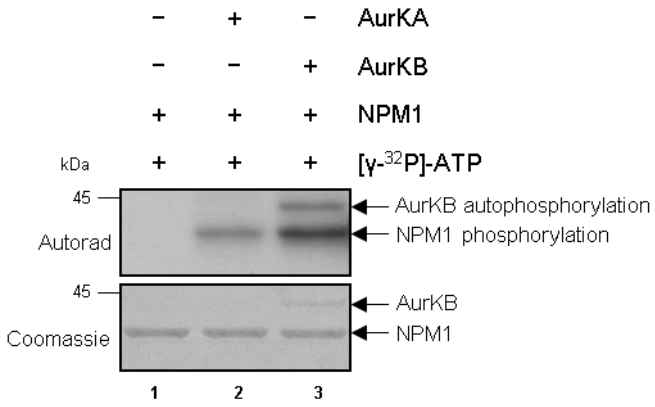



**Figure 2.24. Analysis of the enzymatic (kinase) activity of purified recombinant $His_6$-tagged human Aurora Kinase A and B:** *In vitro* phosphorylation assay demonstrating the kinase activity of Aurora Kinase A (AURKA, lane 2) and Aurora Kinase B (AURKB lane 3) towards NPM1 as a substrate using [$\gamma$-$^{32}$P]-ATP. The 'no enzyme' control (lane 1) serves as the negative control for the assay. Note that the assay is only qualitative as the activities of the kinases have not been normalized.

### *2.6.9. Purification of antibody from crude sera*

#### *2.6.9.1. Purification of anti-AcNPM1 (K229, K230)*

The polyclonal anti-AcNPM1 antibody was purified by peptide affinity chromatography. The NPM1 (K229, K230)ac peptide was coupled to CnBr Sepharose beads (GE Healthcare) in the following way. 52 mg of the beads were soaked in 1 ml of 1 mM HCl at RT for 30 min, after which they were washed with 15 gel volumes of chilled 1 mM HCl by centrifugation thrice at 1500 rpm, 4°C for 5 min. Washed beads were then equilibrated with 1 ml of coupling buffer (100 mM $NaHCO_3$, (Sigma) pH 8.3, 500 mM NaCl (Sigma)) by centrifugation at the same conditions. The beads were incubated with 700 – 750 µg of the NPM1 (K229, K230)ac peptide solution in PBS in a final volume of 4 ml having an effective concentration of 100 mM $NaHCO_3$ and 500 mM NaCl, at 12 rpm, RT on an end-to-end shaker for 2 – 4 h. Beads bound to the ligand were washed with coupling buffer as before and then incubated in 4 ml of 1 M ethanolamine (Sigma), pH 8, at 12 rpm, RT on an end-to-end shaker for 2 h. Beads were then washed alternately with 8 gel volumes of ice-cold low pH buffer (100 mM NaOAc, pH 3 – 4, 500 mM NaCl) and high pH buffer (100 mM Tris-HCl, pH 8 – 9, 500 mM NaCl), 6 times each, and finally with 10 gel volumes of ice-cold PBS. Beads coupled to the peptide were then incubated in 500 µl of crude serum mixed with 500 µl of PBS at 12 rpm, 4°C on an end-to-end shaker for about 12 h. Beads were collected by centrifugation and washed with 20 gel volumes of ice-cold PBS, 4 times. Antibody bound to the peptide-coupled beads were eluted using 200 mM glycine (Sigma), pH 2.8. 100 µl of eluted antibody fractions (5 – 6 fractions were collected) were mixed with 10 µl of 1 M Tris-HCl (Sigma), pH 8, and analyzed by SDS-PAGE (12% gel) followed by Coomassie blue staining. Suitable fractions were pooled and dialyzed against 50% v/v glycerol in PBS at 4°C for 2 – 4 h. Purified antibody thus obtained were analyzed by SDS-PAGE (12% gel) followed by Coomassie blue staining (Figure 2.25). The concentration of the antibody was determined by gel estimation taking known concentrations of BSA as standard. The antibody was aliquoted and stored at −20°C for future use.

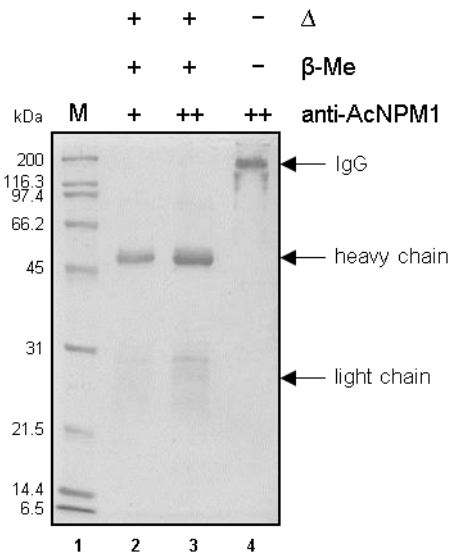



**Figure 2.25. Purification profile of anti-AcNPM1 antibody:** Gel profile of anti-AcNPM1 antibody purified from the crude serum by peptide affinity chromatography. The antibody was loaded in the gel with (lanes 2 – 3) and without (lane 4) treatment with reducing agent β-mercaptoethanol (β-Me) and boiling at 90°C (Δ) to check its integrity. A single higher band in lane 4 indicates that the antibody (IgG) was intact and suitable for use in immuno-pull-down assays. The antibody heavy and light chain bands are indicated by arrows.

### *2.6.9.2. Purification of anti-NPM2*

The polyclonal anti-NPM2 antibody raised in-house in rabbit was purified both by peptide affinity chromatography as well as Protein G affinity chromatography.

The batch of antibody purified by peptide affinity chromatography was done following the same protocol as described in Section 2.6.9.1. The purification profile of the antibody is shown in Figure 2.26A.

Another batch of antibody was purified by Protein G affinity chromatography. For this, about 250 µl of the crude sera mixed with Tris-HCl (Sigma), pH 8, to a final concentration of 100 mM, was passed through an Econo-column (Bio-Rad) packed with 125 µl of 50% v/v slurry of Protein G Sepharose beads (Amersham Biosciences) pre-equilibrated with 100 mM of Tris-HCl, pH 8. Beads bound to the antibody were washed once with 20 gel volumes of 100 mM Tris-HCl, pH 8, and then with 10 mM Tris-HCl, pH 8 in the column itself. The antibody was eluted from the beads using 5 bead volumes of 100 mM glycine (Sigma), pH ~ 3, and collected as 100 µl fractions, which were immediately mixed with Tris-HCl, pH 8 to a final concentration of 100 mM. The fractions were analyzed by SDS-PAGE (12% gel) followed by Coomassie blue staining. Suitable fractions were pooled and dialyzed against

50% v/v glycerol in PBS at 4°C for 2 – 4 h. The purified antibody was analyzed by SDS-PAGE (12% gel) followed by Coomassie blue staining for quality check (Figure 2.26B). The purified antibody was aliquoted and stored at −20°C for future use.

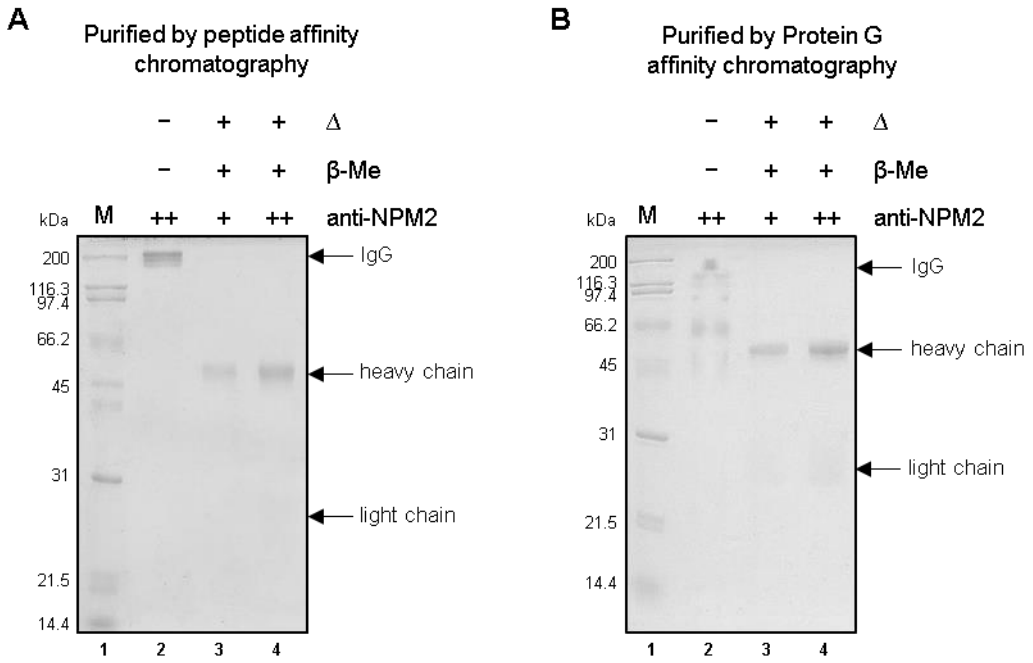



**Figure 2.26. Purification profile of anti-NPM2 antibody:** (A – B) Gel profiles of anti-NPM2 antibody purified from the crude serum by (A) peptide affinity and (B) Protein G affinity chromatography. The antibody was loaded in the gel with (lanes 3 – 4) and without (lane 2) treatment with reducing agent β-mercaptoethanol (β-Me) and boiling at 90°C (Δ) to check its integrity. A single higher band in lane 2 indicates that the antibody (IgG) was intact and suitable for use in immuno-pull-down assays. The antibody heavy and light chain bands are indicated by arrows. M denotes the protein molecular weight marker.

### *2.6.9.3. Purification of pre-immune IgG*

The pre-immune IgG from rabbit serum was purified by Protein G affinity chromatography following the same protocol as described in Section 2.6.9.2. The purification profile of the antibody is shown in Figure 2.27.

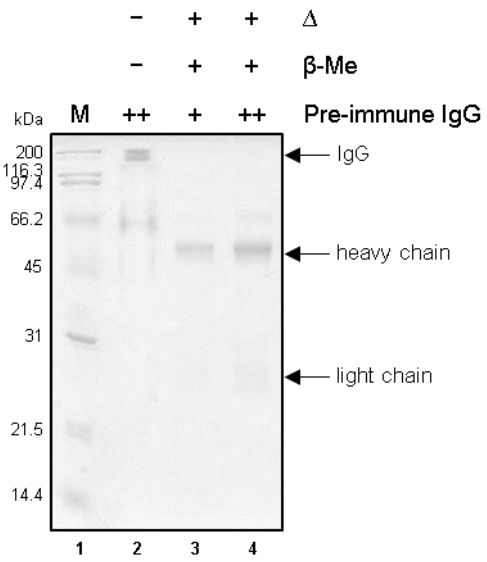



**Figure 2.27. Purification profile of pre-immune IgG from rabbit serum:** Gel profile of pre-immune IgG purified from the rabbit crude serum by Protein G affinity chromatography. The antibody was loaded in the gel with (lanes 3 – 4) and without (lane 2) treatment with reducing agent β-mercaptoethanol (β-Me) and boiling at 90°C (Δ) to check its integrity. A single higher band in lane 2 indicates that the antibody (IgG) was intact and suitable for use in immuno-pull-down assays. The antibody heavy and light chain bands are indicated by arrows. M denotes the protein molecular weight marker.

### *2.6.10. Purification of 3xFLAG-NPM2 from HEK-293 stable cell line*

HEK-293 cells stably expressing 3xFLAG-NPM2 or p3xFLAG-CMV10 empty vector (negative control) were cultured in 150 mm dishes. At around 90% confluency, cells from one 150 mm dish for each type were harvested by scraping and washing with PBS. Cells were lysed in 1 ml of FLAG-Lysis buffer (50 mM Tris-HCl (Sigma), pH 7.4, 150 mM NaCl (Sigma), 1 mM EDTA (Sigma), 1% v/v Triton-X-100 and 10% v/v glycerol along with 1X protease inhibitor cocktail (Sigma)) at 4°C, 12 rpm for 1 h. The lysate was clarified by centrifugation at 13000 rpm, 4°C for 10 min and the supernatant was collected. 40 µl of 50% v/v slurry of M2 agarose beads (Sigma) were added to each sample and kept for binding at 4°C for not more than 3 h. Beads were then collected by centrifugation at 8200 g for 30 s and washed thrice with Tris-buffered saline (TBS; 50 mM Tris-HCl, pH 7.5, 150 mM NaCl) by incubation on an end-to-end rotor at 4°C, 12 rpm for 5 min and centrifugation at the same conditions as mentioned before. 50 µl of BC-100 (20 mM Tris-HCl, pH 7.5, 10% v/v glycerol, 0.4 mM Na-EDTA, pH 8, 100 mM KCl, 0.1% v/v Nonidet P-40, 0.5 mM PMSF, 9.8 mM β-Mercaptoethanol) buffer having 15 µg of 3xFLAG-peptide (Sigma) was added to the beads in each sample and incubated on an end-to-end rotor at 4°C, 12 rpm for

30 min to 1 h for elution of the FLAG-tagged proteins. The beads were spun down at 8200 g, 4°C for 5 min and the supernatant containing the eluted proteins were collected, aliquoted and snap-frozen in liquid $N_2$. Purity and concentration of the proteins were checked by SDS-PAGE and staining with Coomassie brilliant blue (Figure 2.29)

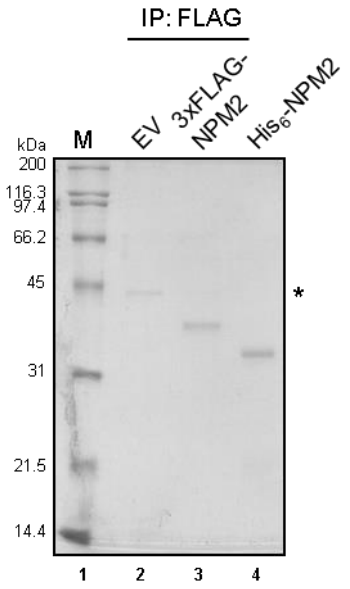



**Figure 2.28. Purification profile of 3xFLAG-NPM2 from HEK-293 stable cells:** Profile of FLAG-tagged NPM2 purified from HEK-293 stable cells (lane 3) and the complex pulled down from cells stably expressing the empty FLAG vector (EV, negative control: lane 2). * denotes non-specific protein pulled down by the M2 agarose beads in the negative control cells. Lane 4 indicates the position of $His_6$-NPM2 purified from *E. coli* cells. M denotes the protein molecular weight marker.

## *2.7. Whole-cell lysate preparation*

### *2.7.1. RIPA whole-cell lysate preparation for immunoprecipitation assays and western blotting*

HEK-293 or H1299 cells grown in 100 mm dishes to about 90% confluency were harvested by gently scraping the cells off the dish using a cell scraper (Corning) and washing in PBS for 2 – 3 times at 2000 rpm, 4°C for 3 min. The cell pellet was resuspended in 10 pellet volumes of RIPA (radioimmunoprecipitation assay) lysis buffer (50 mM Tris-HCl (Sigma), pH 7.4, 150 mN NaCl (Sigma), 1% v/v Nonidet P-40 (Sigma), 0.5% w/v Na-deoxycholate (Sigma), 2 mM Na-EDTA (Sigma), pH 8, 1 mM EGTA (Sigma), 0.1% w/v SDS (Sigma) and 1X protease inhibitor cocktail (Sigma)). Cell lysis was carried out on an end-to-end shaker at 12 rpm, 4°C for 3 h, followed by a short sonication (5 cycles, each of 30 s ON

and 30 s OFF mode) using a Diagenode Bioruptor (Liège, Belgium) to shear the genomic DNA. The lysate was clarified by centrifugation at 13000 rpm, 4ºC for 10 min. The supernatant was collected and used subsequently for immunoprecipitation assays.

For preparing lysates for western blotting, the lysis buffer contained 300 mM of NaCl instead of 150 mM. The rest of the protocol was the same as described above. The protein concentration of the final lysate was estimated by colorimetric assay using the Bradford reagent (Bio-Rad). The lysates were aliquoted, snap-frozen in liquid $N_2$ and stored at −80°C for future use.

### *2.7.2. Laemmli lysate preparation for western blotting*

Mammalian cells grown in culture dishes or flasks were harvested by gently scraping the cells of the dishes or collecting after trypsinization, and washed in PBS at 2000 rpm, 4°C for 3 min, 2 – 3 times. The cell pellet was resuspended in 10 pellet volumes of Laemmli lysis buffer (2% w/v SDS (Sigma), 10% v/v glycerol and 62.5 mM Tris-HCl (Sigma), pH 6.8). Cell lysis was carried out by boiling the suspension at 90°C for 3 min, followed by incubation at RT for 3 min and vortex mixing for a few seconds. This step was repeated for a total of 3 times, followed by a short sonication (bursts of 3 s each with a gap of 5 s for a total time of 1 – 2 min at 35% amplitude) using a Vibra-Cell ultrasonic liquid processor (Sonics & Materials Inc.) to shear the genomic DNA. The lysate was clarified by centrifugation at 13000 rpm, RT for 5 – 10 min. The supernatant was collected, aliquoted, snap-frozen in liquid $N_2$ and stored at −80°C for future use. Normalization of the total protein across samples was done by SDS-PAGE followed by Coomassie blue staining.

### *2.7.3. FLAG (M2) lysate preparation for western blotting and immunoprecipitation assays*

Mammalian cells grown in culture dishes were harvested by gently scraping the cells off the dishes or collecting after trypsinization, and washed in PBS at 2000 rpm, 4°C for 3 min, 2 – 3 times. The cell pellet was resuspended in 10 pellet volumes of FLAG-Lysis buffer (50 mM Tris-HCl (Sigma), pH 7.4, 150 mM NaCl (Sigma), 1 mM EDTA (Sigma), 1% v/v Triton-X-100 (Sigma) and 10% v/v glycerol along with 1X protease inhibitor cocktail (Sigma)). Cell lysis was carried out on an end-to-end shaker at 12 rpm, 4°C for 1 h, followed

by a short sonication (5 cycles, each of 30 s ON and 30 s OFF mode) using a Diagenode Bioruptor (Liège, Belgium) to shear the genomic DNA. The lysate was clarified by centrifugation at 13000 rpm, 4°C for 5 – 10 min. The supernatant was collected and used subsequently for immunoprecipitation assays or aliquoted, snap-frozen in liquid $N_2$ and stored at −80°C for future use in western blotting. The protein concentration of the final lysate was estimated by colorimetric assay using the Bradford reagent (Bio-Rad) or normalization of the total protein across samples was done by SDS-PAGE followed by Coomassie blue staining.

## *2.8. Pull-down assays*

### *2.8.1. In vivo immuno-pull-down with anti-NPM2 or anti-c-fos antibody*

RIPA lysate (having 150 mM NaCl, Section 2.7.1) was prepared from HEK-293 cells grown in 100 mm culture dishes. An input lysate (10% of the total volume of lysate to be used for the immunoprecipitation (IP)) was saved at −80°C. The rest of the lysate was incubated with 1 or 2 µg of purified anti-NPM2 antibody or pre-immune IgG wherever applicable and 40 µl of 50% v/v slurry of pre-blocked and pre-equilibrated Protein G Sepharose beads (Amersham Biosciences) on an end-to-end shaker at 12 rpm, 4°C for about 12 h. Alternatively, the addition of the Protein G beads could also be done post-incubation with antibody, for 2 – 3 h. The pre-blocking of the beads was done with BSA (Sigma) to a final concentration of 1 mg/ml by incubation at 12 rpm, 4°C for 12 h. The beads were then collected by centrifugation at 1500 rpm, 4°C for 3 – 5 min, washed in RIPA lysis buffer thrice, boiled in 1X SDS loading buffer (50 mM Tris-HCl (Sigma), pH 6.8, 2% v/v SDS (Sigma), 0.1% w/v bromophenol blue (Sigma), 10% v/v glycerol, 100 mM β-mercaptoethanol (Sigma)) for 10 min and analyzed by SDS-PAGE followed by western blotting.

Immunoprecipitation assays from AU565 cells were performed by lysing 10 million cells in RIPA buffer followed by incubation with 5 µg anti-c-fos antibody or pre-immune IgG and 25 µl of Dynabeads Protein G (Invitrogen) on an end-to-end shaker at 4°C overnight. Beads were washed thrice with RIPA buffer and eluted using a 1X sample loading buffer. The other steps of the assay were essentially the same as those described above.

### *2.8.2. In vivo immuno-pull-down with anti-FLAG M2-agarose*

*In vivo* immuno-pull-down with anti-FLAG M2-agarose was performed as per manufacturer's protocol (Sigma). Briefly, mammalian cells (HEK-293 or H1299) were cultured in 100 mm dishes and harvested at about 90% confluency by scraping. The cells were pelleted and washed in PBS by centrifugation. The pellet was resuspended in 10 pellet volumes of FLAG-Lysis buffer (50 mM Tris-HCl (Sigma), pH 7.4, 150 mM NaCl (Sigma), 1 mM EDTA (Sigma), 1% v/v Triton-X-100 (Sigma) and 10% v/v glycerol along with 1X protease inhibitor cocktail (Sigma)). Cell lysis was carried out by incubation on an end-to-end shaker at 12 rpm, 4°C for 1 h. The lysate was clarified by centrifugation at 13000 rpm, 4°C for 5 min. An input lysate (10% of the total volume of lysate to be used for the immunoprecipitation (IP)) was saved at −80°C. The remaining volume of the lysate was incubated with an appropriate volume (20 – 40 µl) of 50% v/v slurry of anti-FLAG M2-agarose beads (Sigma) pre-equilibrated with TBS (50 mM Tris-HCl, pH 7.5. 150 mM NaCl), overnight at 12 rpm, 4°C. The beads were then collected by centrifugation at 8200 g, 4°C for 30 s and washed thrice with TBS or the lysis buffer. For the washing, the beads were incubated with the buffer at 12 rpm, 4°C for 5 min followed by centrifugation at 8200 g, 4°C for 30 s. The immunoprecipitated complex was then eluted using 3xFLAG-peptide (Sigma) as described in Section 2.6.10, or in 1X SDS sample loading buffer, followed by heating at 90°C for 10 min. The eluted complex was subjected to SDS-PAGE followed by western blotting.

### *2.8.3. In vitro Ni-NTA pull-down assays*

About 1 µg of recombinant FLAG-NPM1 protein purified from *E. coli*, was used per reaction for mass acetylation by p300 enzyme (20 ng) (Section 2.11.2). The mock acetylated (without acetyl-CoA) sample was kept as control. The acetylation of NPM1 protein by this reaction was confirmed by western blotting with the anti-AcNPM1 antibody. This mock acetylated or acetylated FLAG-NPM1 was incubated with $His_6$-tagged proteins namely AP4, POLR2K, PC4 or SNAI1 in 200 µl of the interaction buffer (20 mM Tris-HCl (Sigma), pH 7.9, 20% v/v glycerol, 0.2 mM EDTA (Sigma), pH 8, 0.1% v/v Nonidet P-40 (Sigma), 2 mM PMSF (Sigma), 150 mM KCl (Sigma) and 30 mM imidazole (Sigma)) along with Ni-NTA resin (Calbiochem). 20% of the sample was saved prior to the addition of Ni-NTA as input. The mixture was incubated on an end-to-end shaker at 12 rpm, 4°C,

for 2 h. The beads were extensively washed with 1 ml interaction buffer, thrice. The proteins bound to the beads were eluted in 1X SDS sample loading buffer and were analyzed by SDS-PAGE followed by western blotting.

## *2.9. Chromatin Immunoprecipitation (ChIP)*

H1299 or AW13516 cells were cultured in 100 mm or 150 mm dishes up to 80 – 90% confluency. Cells were cross-linked using 1% formaldehyde (Sigma) added directly into the culture medium for 10 min at RT on a reciprocal shaker. The reaction was stopped by the addition of 0.125 M glycine (Sigma) for 5 min at RT on the reciprocal shaker. The crosslinked cells adhered to the dishes were rinsed twice with ice-cold PBS containing 2 mM PMSF (Sigma), scraped off the dishes and collected in tubes, followed by centrifugation at 4000 rpm, 4°C for 5 min. The cell pellets were stored temporarily at −80°C.

About 10 – 20 million cells (standardized per ChIP with specific antibody) were lysed in 200 µl of SDS lysis buffer (1% w/v SDS (Sigma), 10 mM EDTA (Sigma), 50 mM Tris-HCl (Sigma), pH 8.1, with 1X protease inhibitor cocktail (Sigma)) and subjected to sonication using a Diagenode Bioruptor (Liège, Belgium) to produce DNA fragments of 100 – 300 bp in length. The number of cycles of sonication (each cycle with of 30 s ON and 30 s OFF mode) was standardized by testing the profile of the sheared DNA obtained after sonicating for 5 – 20 cycles. A gap of 3 – 5 min was allowed after every 5 cycles. The profile corresponding to 15 cycles of sonication generally gave DNA fragments enriched at 300 bp and hence this condition was chosen for performing the assays (Figure 2.29). The subsequent steps involving chromatin was performed using low-retention plasticware. The lysate was then clarified by centrifugation at 13000 rpm, 4°C for 5 – 10 min. About 10% volume of this lysate that would be used for the immunoprecipitation was saved and stored at −80°C temporarily. The lysate having the sheared chromatin, was diluted 10-fold using cold ChIP-dilution buffer (0.01% w/v SDS, 1.1% v/v Triton X-100 (Sigma), 1.2 mM EDTA, 16.7 mM Tris-HCl, pH 8, and 167 mM NaCl (Sigma) with 1X protease inhibitor cocktail) and pre-cleared prior to immunoprecipitation, with pre-blocked Protein G Sepharose (Amersham Biosciences) (30 µl of 50% v/v slurry, per IP sample) on an end-to-end shaker at 12 rpm, 4°C for 1 h. The samples were centrifuged at 2000 rpm, 4°C for 3 – 5 min and the supernatant was collected in a fresh tube. The pull-downs for ChIP assay

were performed with 5 µg of antibody (anti-FLAG, anti-c-fos, anti-AcNPM1 or mouse or rabbit pre-immune IgG (as a negative control) wherever applicable) and pre-blocked Protein G Sepharose beads (60 µl of 50% v/v slurry, per IP sample), and incubated on an end-to-end shaker at 12 rpm, 4°C for about 12 h. The beads were subsequently washed as follows:

1. Low salt buffer (0.1% w/v SDS, 1% v/v Triton X-100, 2 mM EDTA, 20 mM Tris-HCl, pH 8, and 150 mM NaCl) with 1X protease inhibitors: 1 wash.

2. High salt buffer (0.1 w/v % SDS, 1% v/v Triton X-100, 2 mM EDTA, 20 mM Tris-HCl, pH 8, and 500 mM NaCl with 1X protease inhibitors): 1 wash.

3. LiCl buffer (250 mM LiCl (Sigma), 1% v/v Nonidet P-40 (Sigma), 1% w/v sodium deoxycholate, 1 mM EDTA and 10 mM Tris-HCl, pH 8, with 1X protease inhibitors): 1 wash.

4. TE (10 mM Tris-HCl, pH 8, and 1 mM EDTA with 1X protease inhibitors): 2 washes.

To the washed beads and input sample (10% of volume used in IP), elution buffer (0.2% w/v SDS and 100 mM $NaHCO_3$) was added (500 µl per sample) and the protein-DNA complexes were eluted out from the beads on an end-to-end rotor at 12 rpm, RT for 30 min. The DNA-protein complexes were then reverse cross-linked by adding 200 mM NaCl and 20 µg Proteinase K (Sigma) and incubating at 65ºC for 4 h under mild agitation in a Thermomixer (Eppendorf). Subsequently, 20 µg of RNase A (Sigma) were added and the samples were further incubated for 15 min at 37ºC. The immunoprecipitated DNA was extracted by phenol-chloroform extraction, ethanol-precipitated, and used for real-time ChIP-qPCR analysis (Section 2.9.1).

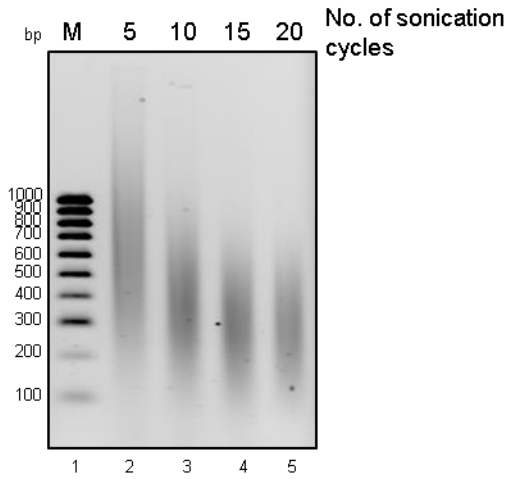


**Figure 2.29. ChIP input DNA profile:** Representative image showing the profile of input DNA for ChIP, isolated after shearing of the chromatin for the different number of sonication cycles as indicated, in the AW13516 shNPM1-luc+ untreated (control) cells. 15 cycles of chromatin sonication gave good enrichment of the fragmented DNA around 300 bp and hence was chosen as the condition for the final experiment.

### *2.9.1. ChIP-qPCR*

The enrichment of a specific protein in regions of the chromatin was analyzed by quantitative PCR following chromatin immunoprecipitation (ChIP-qPCR) using the StepOnePlus Real-Time PCR Detection System (Applied Biosystems (ABI). Reactions were set up using either KAPA SYBR FAST qPCR Master Mix (Kapa Biosystems) or TB Green Premix Ex Taq II (Tli RNase H Plus) (Takara Bio Inc.) as per manufacturer's protocol and using 2 µl of appropriately diluted (1 in 10) DNA. The results of the ChIP were represented in terms of percentage of input or fold enrichment over IgG negative control. The sensitivity and specificity of the primers used were determined by melt curve analysis. The primer pairs used for ChIP-qPCR analysis in this study are listed in Appendix Table A.7.

## *2.10. Immunohistochemistry*

Oral cancer patient tissue samples were collected from Sri Devaraj Urs Medical College, Kolar and Bangalore Institute of Oncology (BIO), Bangalore, in the Indian state of Karnataka. The ethical clearance committees of the institutions agreed to this investigation and informed consent was obtained from all patients. The clinicopathological information of the patients from whom the oral tumor tissues were derived, is listed in Appendix, Table A.8. Tumor and adjacent normal tissue samples collected from the oral cancer patients after surgery were dehydrated and paraffin-embedded, followed by sectioning using a microtome (Leica Biosystems, Wetzlar, Germany). After blocking, these sections were incubated in different primary antibodies: anti-NPM1, anti-c-fos, and anti-p53 primary antibodies. Immunohistochemistry (IHC) was performed with the streptavidin-biotin kit (Dako, Santa Clara, CA, USA). Immuno-reactivity (brown color precipitate) was developed using the chromogen 3,3'-diaminobenzidine (DAB) tetrahydrochloride (Sigma), and counterstaining was done with hematoxylin.

Tissue arrays were prepared from oral cancer samples along with its matched normal tissues by punching a 1 mm core region from the donor blocks. 5 µm paraffin sections were cut for immunohistochemistry. The sections on the array were baked overnight on a dry bath at 56°C and for 1 h at 60°C. The sections were subjected to antigen retrieval in sodium citrate (Sigma) buffer, pH 7.2, for 20 min and cooled down to RT for further processing. The sections were blocked in 5% w/v skimmed milk solution for 2 h at RT. The respective primary antibodies of appropriate dilution were then added and kept in a humid chamber for overnight incubation at RT. The arrays were developed using the streptavidin-biotin kit (Dako) and 3,3'-diaminobenzidine as substrate, and counterstained by diluted hematoxylin. The slides were mounted with dibutylphthalate polystyrene xylene (DPX). Semi-quantitative staining analysis was done using H-score (histo-score) by counting around 400 cells from three different fields in the slides considering the low, moderate and higher intensity of protein expression. Nuclear positivity was considered for all three proteins.

## *2.11. In vitro assays*

### *2.11.1. In vitro phosphorylation assay*

1 µg of the substrate protein was incubated with the specific kinase, human Aurora Kinase A, human Aurora Kinase B or rat liver Casein Kinase II (Sigma) (suitably diluted in BC-100), in 2X Kinase Buffer (100 mM Tris-HCl (Sigma), pH 7.5, 200 mM NaCl (Sigma), 0.2 mM EGTA (Sigma), pH 8, 20 mM $MgCl_2$ (Sigma), and 0.4% v/v β-Mercaptoethanol (Sigma)) making an effective concentration of 1X in the final volume, in the presence of 1 µl of 0.25 µCi/µl γ-$^{32}$P-labeled ATP (Board of Radiation and Isotope Technology, Dept. of Atomic Energy, Govt. of India) at 30°C for 30 min. The reaction was stopped on ice, the samples were boiled in 1X SDS sample loading buffer at 90°C for 10 min and subjected to SDS-PAGE. The gel was stained by Coomassie blue and dried in a gel drier. The dried gel was subjected to autoradiography by exposing to TMX films (Kodak) in an analog X-ray cassette and screen (Kiran) for different time points at RT or −80°C (depending on the intensity of the radioactive signal emitted) and developed using the developer – fixer kit.

For mass phosphorylation, 6 µg of the substrate protein was incubated with the specific kinase in 1X Kinase Buffer in the presence of cold 100 mM cold (non-radiolabeled) ATP (Sigma), pH 7. The reaction mixture was incubated at 30°C for 1.5 h with replenishment

with ATP and the enzyme every 30 min for complete phosphorylation. The reaction was finally stopped on ice and the samples were dialyzed against BC-100 at 4°C for 1 – 2 h to remove the excess ATP and stabilize the proteins. The samples were then snap-frozen in liquid $N_2$ and stored at −80°C for use in subsequent assays later.

### *2.11.2. In vitro acetylation assay*

1 µg of the substrate protein was incubated with the specific acetyltransferase enzyme (suitably diluted in BC-100) in a 30 µl final reaction volume of HAT (histone acetyltransferase) buffer consisting of 50 mM Tris-HCl (Sigma), pH 8.0, 10% v/v glycerol, 1 mM DTT (Calbiochem), 1 mM PMSF (Sigma), 10 mM sodium butyrate (Sigma), 0.1 mM EDTA (Sigma), pH 8.0, and 1 µl of 2.5 Ci/mmol [$^3$H]-acetyl CoA (PerkinElmer, Waltham, MA, USA). The reaction was incubated at 30°C for 30 min, following which the reaction was stopped on ice. As an optional step, the proteins in the reaction mixture were precipitated in 25% v/v trichloroacetic acid (TCA) (Sigma) and incubating on ice for 30 min, followed by centrifugation at 13000 rpm, 4°C for 30 min. The supernatant was discarded and the pellet was washed with ice-cold acetone. The pellet was collected by centrifugation at 13000 rpm, 4°C for 30 min and air-dried. The samples were boiled in a 1X SDS-sample loading buffer at 90°C for 10 min and resolved by SDS-PAGE followed by Coomassie blue staining. To visualize the radiolabeled acetylated proteins, the gel was subjected to fluorography. For this, the gel was equilibrated in DMSO for 30 min at RT under gentle shaking and then incubated in a solution containing 22.5% w/v of 2,5-diphenyloxazole (Sigma) in DMSO for 1 h at RT under gentle shaking. The gel was then kept in water for about 1.5 h, dried, subjected to autoradiography by exposing to TMX films (Kodak) in an analog X-ray cassette and screen (Kiran) for different time points at −80°C (depending on the intensity of the radioactive signal emitted), and developed using the developer – fixer kit.

*In vitro* mass acetylation assay was performed with 1.0 µg of bacterially expressed $His_6$-tagged or FLAG-tagged NPM1 in a 30 µl final reaction volume of HAT buffer (as mentioned above) and 1.2 mM cold (non-radiolabeled) acetyl-CoA (Sigma) with 20 ng of p300 enzyme. The reaction was incubated at 37ºC for 90 min. Acetyl-CoA and the p300 enzyme were replenished at 30 min intervals. The acetylated and the mock acetylated (without acetyl-CoA) proteins were analyzed by SDS-PAGE followed by western blotting

with anti-AcNPM1 antibody (Figure 2.30) and subsequently used for *in vitro* interaction assays.

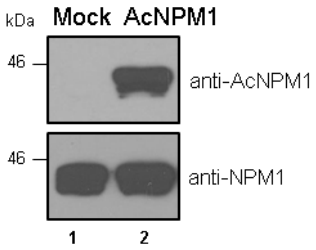



**Figure 2.30. Analysis of mass acetylation of purified recombinant NPM1 by lysine acetyltransferase p300:** Western blot analysis after *in vitro* mass acetylation of FLAG-NPM1 by lysine acetyltransferase (KAT) p300. The upper panel shows western blot with anti-AcNPM1 and bottom panel with anti-NPM1 antibody respectively. The 'Mock' (lane 1) indicates acetylation in the absence of acetyl-CoA and serves as the negative control for the subsequent assay.

### *2.11.3. In vitro filter-binding assay*

The *in vitro* filter binding assay was performed to measure the enzymatic activity of various lysine acetyltransferase (KAT) enzymes. For this, 1 µg of core histones (substrate) was incubated with the KAT enzyme (appropriately diluted to prevent saturation of the detected activity) in 1X HAT buffer consisting of 50 mM Tris-HCl (Sigma), pH 8.0, 10% v/v glycerol, 1 mM DTT (Calbiochem), 1 mM PMSF (Sigma), 10 mM sodium butyrate (Sigma), 0.1 mM EDTA (Sigma), pH 8.0, and 1 µl of 2.5 Ci/mmol [$^3$H]-acetyl CoA (PerkinElmer), at 30°C for 30 min. The reaction was stopped on ice and the samples were spotted uniformly on circular pieces of P81 phosphocellulose filter papers (Whatman plc., Maidstone, UK). The spotted filter papers were air-dried and then washed in wash buffer (46 mM $NaHCO_3$ (Sigma) and 4.7 mM $Na_2CO_3$ (Sigma)). The first wash was for 15 min, while the second and third wash for 10 min at RT on a reciprocal shaker. The washed filter papers were dried at 90°C over a heating block. The dried filter papers were soaked in scintillation fluid (3 ml per sample, in a glass vial) prepared by dissolving 250 mg of 1,4-bis(5-phenyloxazol-2-yl) benzene (POPOP, Sigma) and 2.5 g of 2,5-diphenyloxazole (PPO, Sigma) in 500 ml of toluene. The radioactive signal emitted from the [$^3$H]-acetyl CoA incorporated in the substrate protein by the enzyme was detected by a Wallac 1409 liquid scintillation counter (PerkinElmer) and expressed in terms of counts per min (cpm).

### *2.11.4. In vitro methylation assay*

1 µg of the substrate protein was incubated with the specific methyltransferase enzyme (suitably diluted in BC-100) in a 30 µl final reaction volume of HAT (histone methyltransferase) buffer consisting of 20 mM Tris-HCl (Sigma), pH 8, 4 mM EDTA (Sigma), 200 mM NaCl (Sigma) with 1 µl of S-adenosyl-L-[methyl-$^3$H]methionine (15 Ci/mmol, PerkinElmer NEN), at 30° C for 30 min. The reaction was stopped on ice and the samples were analyzed by SDS-PAGE followed by Coomassie staining and subsequent fluorography as described in Section 2.11.2.

### *2.11.5. In vitro template relaxation assay*

p$G_5$ML purified using the QIAGEN EndoFree Plasmid Maxi Kit (Figure 2.31A), was relaxed using $His_6$-tagged core catalytic domain of *Drosophila* Topoisomerase I (dTopo I) purified from *E. coli* (Section 2.6.4). For a large scale relaxation of the supercoiled plasmid for use in *in vitro* chromatin transcription assay, 20 µg of supercoiled $G_5$ML plasmid was incubated in 1X template relaxation buffer (50 mM Tris-HCl (Sigma), pH 7.5, 50 mM KCl (Sigma), 10 mM $MgCl_2$ (Sigma), 0.1 mM EDTA (Sigma), 0.5 mM DTT (Calbiochem), 30 µg/ml BSA (Sigma)) in the presence of 10 µl of dTopoI protein, at 37°C for 2 h. The mixtures were deproteinized and the DNA was extracted using phenol-chloroform-isoamyl alcohol (25:24:1, pH 8, Sigma). The extracted DNA was analyzed in a 1% w/v agarose gel, electrophoresed at 50 V for at least 5 h (Figure 2.31B). The relaxed DNA was aliquoted and stored at −20°C for future use.

For use in histone transfer or supercoiling assays, the same procedure was followed with minor modifications: 200 ng of the plasmid was incubated with specific dilutions of dTopo I (neat, 1/2, 1/3, 1/4, 1/5) in separate reactions in a final volume of 10 µl. The incubation was carried out at 37°C for 40 min. DNA could be relaxed even in 1/5 dilution of the enzyme (Figure 2.31C) and hence this dilution was used as the standard in the histone transfer assay.

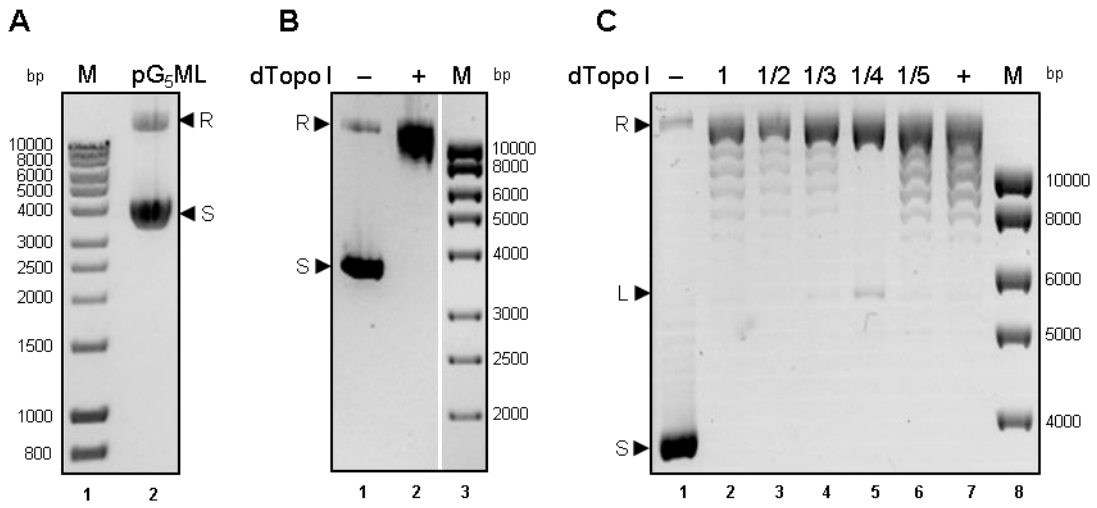



**Figure 2.31. Profile of supercoiled and relaxed form of $G_5$ML plasmid:** (A) Gel profile of purified $G_5$ML plasmid. The majority of the plasmid is in the supercoiled form (S) while a lesser amount remains in the relaxed form (R). (B) Profile of p$G_5$ML (200 ng) with (lane 2) or without (lane 1) treatment with the core catalytic domain of *Drosophila* Topoisomerase I (dTopo I) as analyzed in a 1% w/v agarose gel. Relevant lanes from the same gel have been cropped and aligned. (C) Profile of p$G_5$ML (200 ng) after treatment with different dilutions of dTopo I as indicated (lanes 2 – 6) in the template relaxation assay. Lane 1 denotes the reaction without the addition of dTopo I (negative control) while lane 7 shows the profile of a pre-relaxed $G_5$ML plasmid (positive control). Dilution '1/5' of dTopo was sufficient to relax 200 ng of the supercoiled template and hence was chosen as the standard dilution for use in the histone transfer assay. L: linear form of p$G_5$ML, M: DNA ladder.

### *2.11.6. In vitro histone transfer (plasmid supercoiling) assay*

200 ng of supercoiled $G_5$ML plasmid was relaxed by dTopoI (dilution as standardized) in 1X assembly buffer consisting of 5 mM Tris-HCl (Sigma) pH 8, 0.5 mM Na-EDTA (Sigma), 100 mM NaCl (Sigma), and 0.05 mg/ml BSA (Sigma) for 40 min at 37°C. In a parallel reaction, the histone chaperone protein was incubated with 350 ng (standardized dose) of rat liver core histones in a 1X assembly buffer in a total volume of 20 µl, at 37°C for 40 min for histone binding to the histone chaperone. This was followed by combining the two mixtures and incubating at 37°C for 40 min whereby histones are deposited onto the relaxed template by the chaperone generating one negative supercoil per nucleosome assembled. The reaction was terminated by incubating in Proteinase K stop buffer (200 mM NaCl (Sigma), 20 mM Na-EDTA (Sigma), 0.25 µg/µl glycogen (Sigma), 1% w/v SDS (Sigma), 0.125 µg/µl Proteinase K (Sigma)) at 37°C for 30 min. The reaction mixtures were then deproteinized and the DNA was extracted using phenol-chloroform-isoamyl alcohol (25:24:1, pH 8) (Sigma). The extracted DNA was air-dried and analyzed by agarose gel electrophoresis (1% w/v agarose gel in 1X TBE (111.4 mM Tris, 103.1 mM boric acid, 2

mM EDTA) at 50 V, RT for about 12 h. Figure 2.32A shows the schematic of the assay and Figure 2.32B shows the result representing the histone chaperone activity of NPM1 (Okuwaki et al. 2001b; Swaminathan et al. 2005) revealed by the histone transfer or plasmid supercoiling assay.

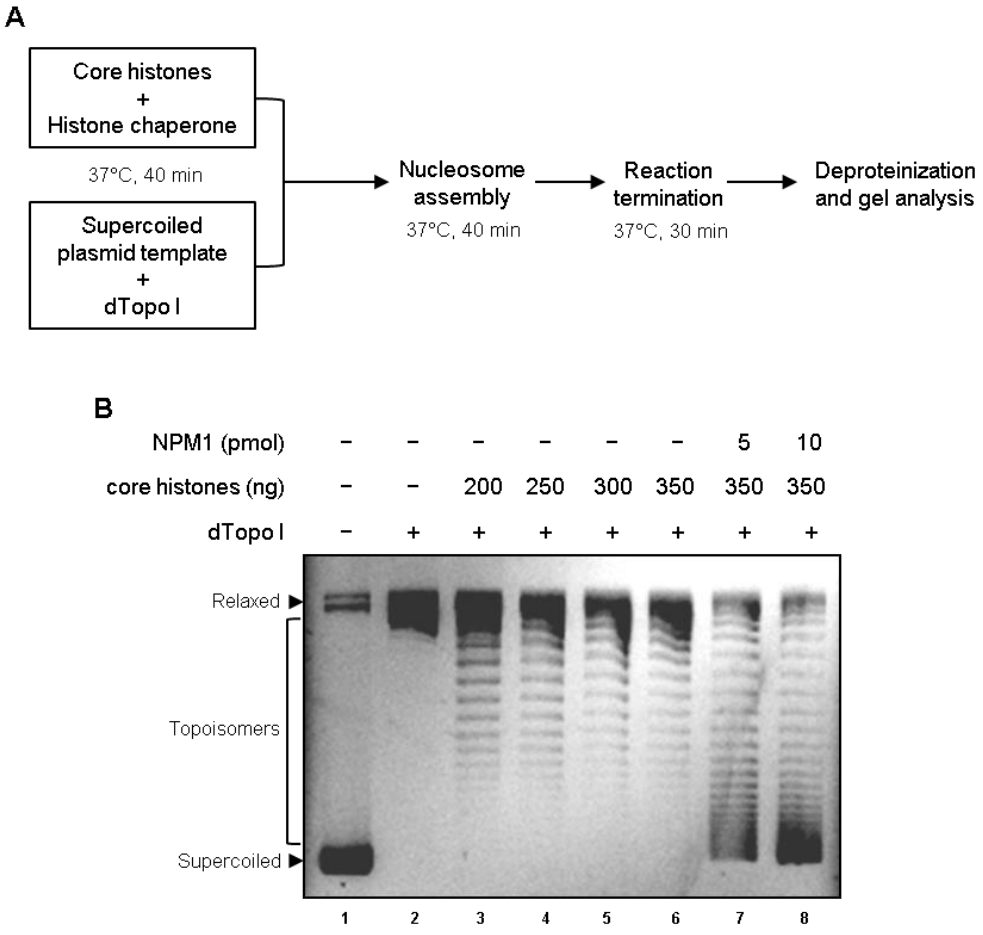



**Figure 2.32. *In vitro* histone transfer or plasmid supercoiling assay:** (A) Schematic representation of the *in vitro* histone transfer or plasmid supercoiling assay. (B) A representative result showing the effect of increasing doses of core histones as indicated (lanes 3 – 6) on the generation of topoisomers from a relaxed plasmid template in the supercoiling assay. Lane 1 and lane 2 indicates the reference positions for the supercoiled and relaxed forms of the plasmid. The addition of histone chaperone NPM1 at two different doses as indicated results in the recovery of a significant amount of supercoiled form of the plasmid from its relaxed conformation (lanes 7 – 8). dTopo I: core catalytic domain of *Drosophila* Topoisomerase I.

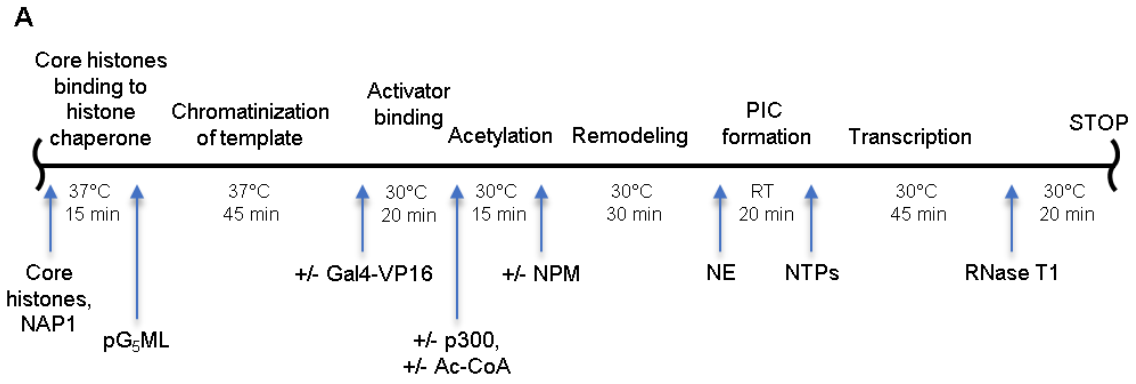
### *2.11.7. In vitro NAP1-mediated chromatin assembly and transcription assay*

In the first step of this assay, chromatin assembly was carried out using mouse NAP1. $His_6$-tagged mouse NAP1 was purified by expressing in *E. coli*. 1 µg of NAP1 was incubated with 500 ng of rat liver core histones in 22 µl of 1X transcription assembly buffer consisting

of 10 mM Tris-HCl (Sigma) pH 8, 1 mM EDTA (Sigma), 150 mM NaCl (Sigma), and 0.1 mg/ml BSA (Sigma) at 37°C for 15 min for histone binding to the chaperone. This was followed by the addition of 168 ng of relaxed $G_5$ML array template plasmid (Section 2.11.5). The total volume of this chromatin assembly reaction is made up to 24 µl in 1X transcription assembly buffer and incubated at 37°C for 45 min for nucleosome assembly.

Next, 28 ng of this chromatinized template was incubated with appropriately diluted transcription activator protein Gal4-VP16 (FLAG-tagged protein previously purified from *E. coli*) in 0.5X HAT buffer (2X HAT buffer composition: 20 mM HEPES (Sigma) pH 7.8, 100 mM KCl (Sigma), 10 mM DTT (Calbiochem), 1 mM PMSF (Sigma), 40 mM sodium butyrate (Sigma), 0.5 mg/ml BSA, and 10% v/v glycerol) at 30°C for 20 min. Appropriately diluted $His_6$-tagged p300 enzyme (purified from baculovirus-infected Sf21 cells) and 0.5 mM acetyl-CoA (Ac-CoA) were added and the reaction was incubated at 30°C for 15 min. This was followed by the optional addition of a specific amount of a remodeler protein (such as NPM1) to assess its transcription coactivation ability. The salt concentration in the reaction so far was balanced with buffer BC-100 (prepared by mixing equal volumes of buffer BC-0 and BC-200) and the reaction mixture was incubated at 30°C for 30 min. This is followed by addition of 24.5 µl of transcription master mix (Tmix) consisting of 2.5 µl of 20X buffer (0.4 M HEPES, pH 8.4, 100 mM $MgCl_2$), 5 µl of HeLa nuclear extract (8.0 mg/ml protein in BC-100) balanced for salt concentration with buffer BC-200 (20 mM Tris-HCl, pH 7.9, 20% v/v glycerol, 0.2 mM EDTA, pH 8.0, 200 mM KCl) and incubation at RT for 20 min, with an intermittent addition of 250 mM DTT in the reaction. Parallelly, a control DNA transcription reaction was carried out for loading control consisting of suitable amount of ML200 DNA, 3.75 µl of buffer BC-200, 1.25 µl of buffer BC-0, 1.25 µl of 20X buffer, 0.5 µl of RNase Out (Invitrogen), 10 µl of HeLa nuclear extract in a total volume of 22.75 µl including 1 µl of 250 mM DTT. After pre-initiation complex (PIC) formation in the previous step, 4.5 µl of ribonucleotide master mix (Rmix) consisting of 2.5 µl of 20X NTP mix (12 mM ATP, 12 mM CTP, 0.5 mM UTP, and 2 mM 3'O-methyl GTP (Amersham Biosciences)), 0.5 µl of RNase Out and 15 µCi of [α-$^{32}$P]-UTP (Board of Radiation and Isotope Technology, Dept. Of Atomic Energy, Govt. of India), was added. In the ML200 reaction, 1.25 µl of 20X NTP mix and 10 µCi of [α-$^{32}$P]-UTP were added and the reactions were incubated at 30°C for 30 min for transcription to occur. 1 µl of 20 U/µl RNaseT1 (Sigma) was then added and incubated at 30°C for 20 min for specific cleavage of the transcript past the incorporated 3'O-methyl GTP. The control ML200

sample was mixed with the test $pG_5ML$ sample and the reaction was finally stopped in 200 µl of transcription stop buffer (20 mM Tris-HCl, pH 8, 1 mM EDTA, pH 8, 100 mM NaCl, 1% w/v SDS (Sigma), 20 ng/µl tRNA) with 250 µl of phenol-chloroform (acid equilibrated to pH 4.7, Sigma). The samples were vortexed and centrifuged at 13000 rpm, RT for 15 min. 230 µl of the aqueous phase was extracted and RNA was precipitated by mixing it with 23 µl of 3M NaOAc (Sigma), pH 5.2 and 650 µl of absolute ethanol, with subsequent incubation at −80°C for at least 1 h. The precipitated RNA was recovered by centrifugation at 4°C, 13000 rpm for 30 min and drying the pellet in Speed Vac (RVC 2-18, Martin Christ Gefriertrocknungsanlagen GmbH, Osterode am Harz, Germany). The dried RNA pellet was dissolved in 10 µl of transcription loading dye (8.0 M urea in 1X TBE, xylene cyanol, 0.005% w/v bromophenol blue). RNA samples were electrophoresed in 7 M urea-5% polyacrylamide gel in 0.5X TBE for 2 – 2.5 h at 300 V, RT. The gel was rinsed in distilled water for 10 – 15 min to remove urea and the free [$^{32}$P]-UTP, vacuum-dried at 80°C for 1 h and subjected to autoradiography at −80°C for the desired duration of time. Figure 2.33A shows the schematic of the assay procedure and Figure 2.33B shows a representative result of chromatin transcription using the activator, p300, and Ac-CoA.



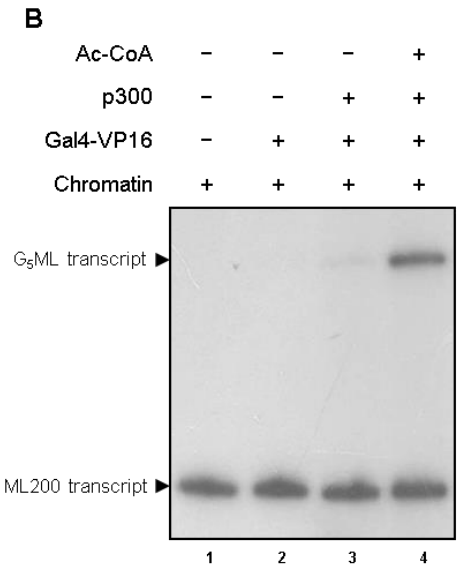



**Figure 2.33. *In vitro* chromatin transcription assay:** (A) Schematic representation of the *in vitro* chromatin transcription assay. Adapted from (Senapati et al. 2015). NE: nuclear extract, PIC: pre-initiation complex, NTPs: ribonucleotide triphosphates, RT: room temperature, Ac-CoA: acetyl CoA. (B) A representative result showing the effect of p300 in the presence of Ac-CoA (lane 4) on chromatin transcription. The lower ML200 transcript serves as the loading control.

### *2.11.8. In vitro nucleosome reconstitution*

The *in vitro* nucleosome reconstitution was done for use in a subsequent *in vitro* histone eviction assay (protocol to be adapted from (Luebben et al. 2010) to test the ability of NPM1 and AcNPM1 to evict histones from a nucleosomal template which potentially activates transcription. For this, the nucleosome was reconstituted using *Xenopus* core histones purified from *E. coli* (Section 2.6.3) and a biotinylated DNA template amplified from the $G_5$ML (Figure 2.34A) array having nucleosome positioning sequence for the sea urchin 5S rRNA. Following is the detailed methodology for this technique.

The template DNA for nucleosome reconstitution was amplified from the $G_5$ML plasmid using specific primers (FP (biotinylated): 5’ – [Btn Tg]GAATCTTTAAACTCGAGTGCATGC – 3’ and RP: 5’ – GCCTAGGGGAGTGGAATGA – 3’). This PCR amplicon (Figure 2.34A) was purified using the NucleoSpin Gel and PCR Clean-up kit (Macherey Nagel), aliquoted and stored at −20°C temporarily.

For histone octamer reconstitution, the recombinant untagged *Xenopus* core histones were normalized for the similar stoichiometry of H2A, H2B, H3, and H4 by SDS-PAGE

followed by Coomassie staining (Figure 2.34B). About 6.25 µg each of the core histones (in SAU-500 or SAU-600 buffers) were taken (~ 25 µg of core histones) and dialyzed in dialysis bags (3 kDa cut-off membranes (Sigma), activated for mononucleosome assembly by boiling in 25 mM EDTA (Sigma) for 1 h, followed by washing in water multiple times and storing in 25 mM EDTA containing 20% v/v ethanol) against 1 l of histone refolding buffer (2 M NaCl (Sigma), 10 mM Tris-HCl (Sigma), pH 7.5, 1 mM Na-EDTA, pH 8, and 5 mM β-mercaptoethanol) overnight (about 12 h) at 4°C. The dialyzed samples were mixed with 80% v/v glycerol saturated with NaCl, to a final concentration of 15% v/v. A part of this sample was resolved by SDS-PAGE followed by Coomassie blue staining to assess the equal stoichiometry of the core histones after the octamer reconstitution (Figure 2.34C).

For the nucleosome reconstitution, a 1:1 ratio of the template DNA and core histones in the octamer were taken. The mixture was set up in the following way: calculated volume of the histone refolding buffer was mixed with 5M NaCl (two-thirds volume of the template DNA); 3 µg of core histone octamer was added near the surface wall of the liquid in the tube, followed by addition of 3 µg of biotinylated template ($G_5$ML) DNA, which were then tap-mixed and spun down; finally, 1 µl of acetylated BSA (Sigma, Cat. No. B8894) and 1% v/v Nonidet P-40 (Sigma) were added and mixed. This mixture was sequentially dialyzed in 3 kDa cut-off dialysis bags (Sigma, activated for mononucleosome assembly) at 4°C against histone refolding buffers with decreasing NaCl concentration as follows: 1.6 M for 45 min, 1.2 M for 1.5 h, 1 M for 1.5 h, 0.8 M for 1.5 h, 0.6 M for 1.5 h, 0.3 M for 30 min, and finally 0.01 M for overnight. The samples were collected and kept at 4°C. The concentration of the DNA in the reconstituted nucleosome was estimated spectrophotometrically using NanoDrop 1000 (Thermo Scientific). A sample having 90 ng of this DNA was diluted in 0.01 M NaCl containing histone refolding buffer, mixed with 6X DNA-loading dye having 0.1% v/v Nonidet P-40 to a final concentration of 1X, and subjected to native-PAGE (5% gel containing 0.3X TBE, 3.75% w/v acrylamide-bis-acrylamide (Sigma) solution, 1% v/v glycerol, 0.1% w/v ammonium persulfate (Sigma) and 0.1% v/v TEMED (Sigma)) in 0.3X TBE at 100 V, RT for 1 h. The electrophoresis was performed in an SDS-free apparatus (pre-cleaned with methanol and water) and the gel was pre-run with the loading dye containing 0.1% v/v Nonidet P-40 before resolving the nucleosome samples. Finally, the gel was stained with SYBR Green dye (Invitrogen) in dark for 15 min on a reciprocal shaker and imaged using the VersaDoc imaging system (Bio-Rad) (Figure 2.34D). A fraction of the reconstituted nucleosome sample was also

analyzed by SDS-PAGE followed by Coomassie staining to check the stoichiometry of the core histones in the reconstituted nucleosome (Figure 2.34E).

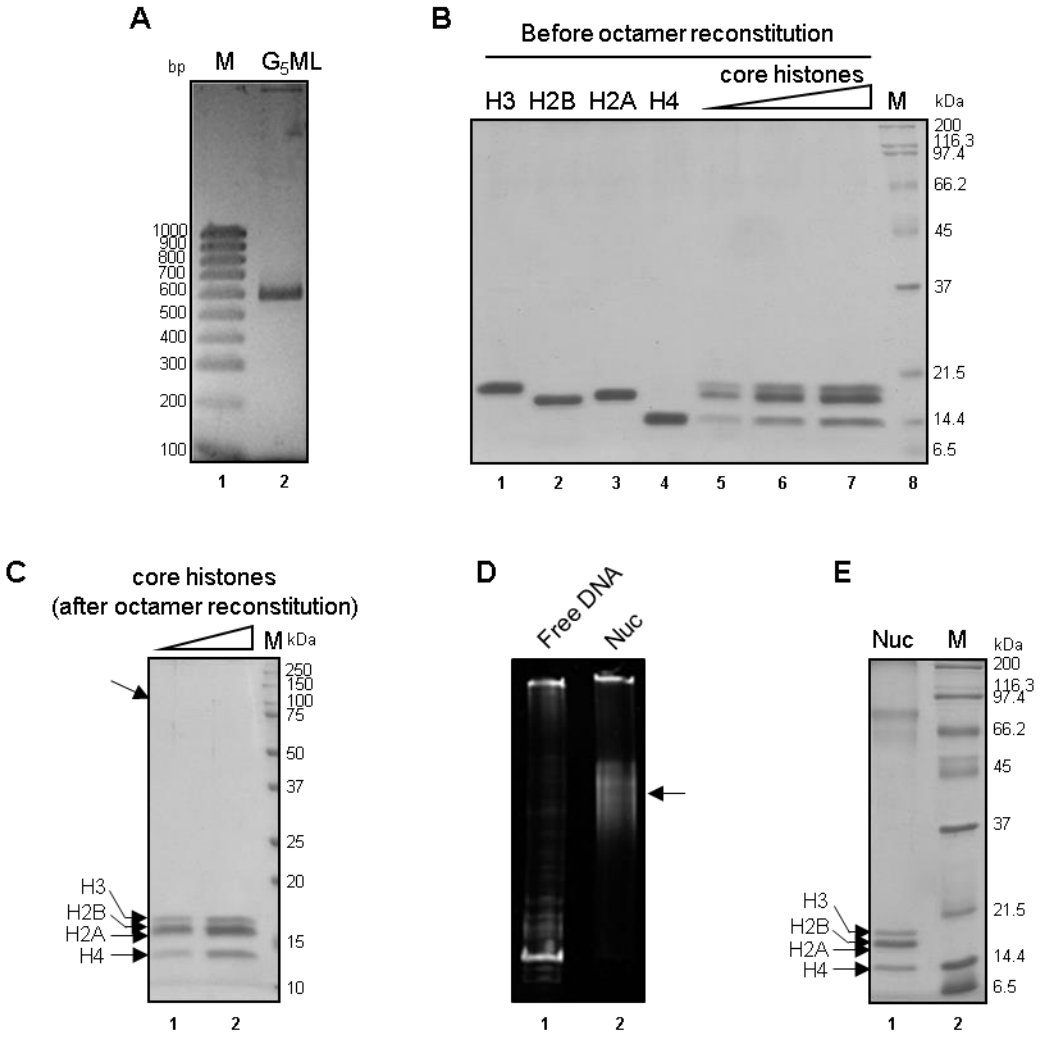
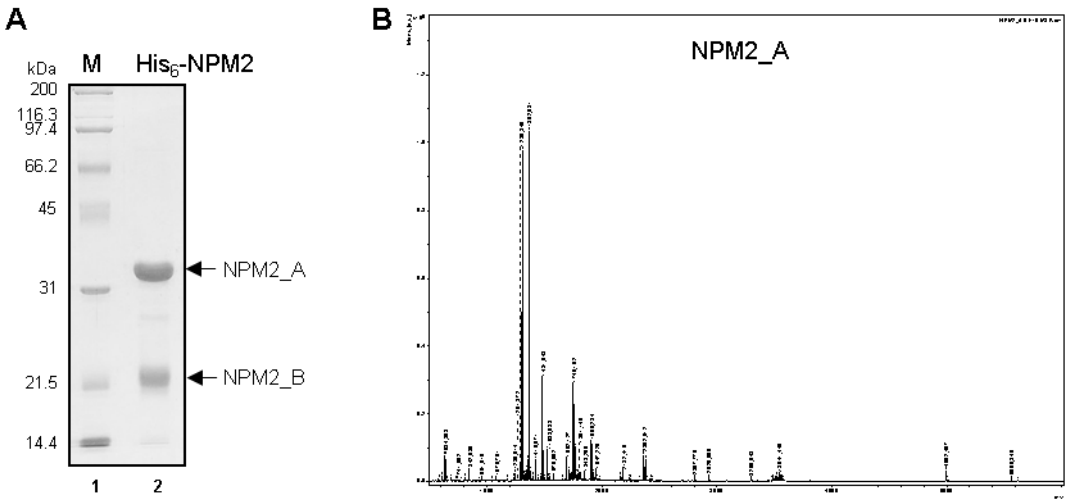



**Figure 2.34. *In vitro* nucleosome reconstitution:** (A) Gel profile of the DNA template PCR amplified from the $G_5$ML plasmid, for *in vitro* nucleosome reconstitution. (B) SDS-PAGE profile of normalized amounts of the recombinant untagged *Xenopus* core histones purified from *E. coli,* before octamer reconstitution. Similar Coomassie staining intensities for the different core histones indicate almost equal stoichiometry of the proteins (lanes 1 – 4). The bands of H2A and H2B merge together between upper (H3) and lower (H4) bands (lanes 5 – 7). (C) Coomassie-stained SDS-PAGE profile of core histones from the reconstituted octamer. (D) SYBR Green-stained native PAGE profile of 90 ng of free DNA template (lane 1) and reconstituted nucleosome (Nuc, lane 2, indicated by the arrow). (E) Coomassie-stained SDS-PAGE profile of core histones from the reconstituted nucleosome (Nuc, lane 1). M: protein molecular weight marker or DNA ladder.

## *2.12. Mass spectrometry*

### *2.12.1. Protein identification by mass spectrometry analysis*

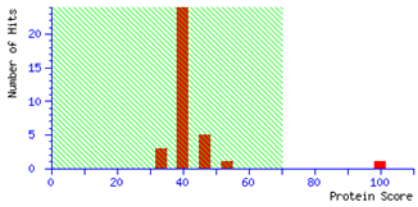
As mentioned in Section 2.6.1, the SDS-PAGE profile of recombinant $His_6$-NPM2 purified from *E. coli* detected a faster-migrating band near 21.5 kDa protein molecular weight marker position (Figure 2.19D). To confirm if this protein (named NPM2_B) was a degradation product of the full-length $His_6$-NPM2 (named NPM2_A) or some other protein co-purified along with $His_6$-NPM2 expressed in the *E.coli* cells, both the bands were excised from a Coomassie blue-stained SDS-PAGE gel (Figure 2.35A) and subjected to MALDI-TOF mass spectrometric analysis (Proteomics Facility at Molecular Biophysics Unit, Indian Institute of Science, Bangalore, India) for protein identification. Figure 2.35B and D show the raw mass spectra of 'NPM2_A' and 'NPM2_B'. The raw data were analyzed using the Mascot database (www.matrixscience.com). The results showed that the protein named 'NPM2_A' had a top score of 100 for human NPM2 (Figure 2.35C), and the protein named 'NPM2_B' had a top significant score of 82 for human NPM2 (Figure 2.35E). The other hits below score 70 ($P > 0.05$) were considered non-significant. The faster-migrating protein in SDS-PAGE was thus confirmed as a degradation product of $His_6$-NPM2.




C

**Mascot Score Histogram**

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 70 are significant (p<0.05).




**Concise Protein Summary Report**

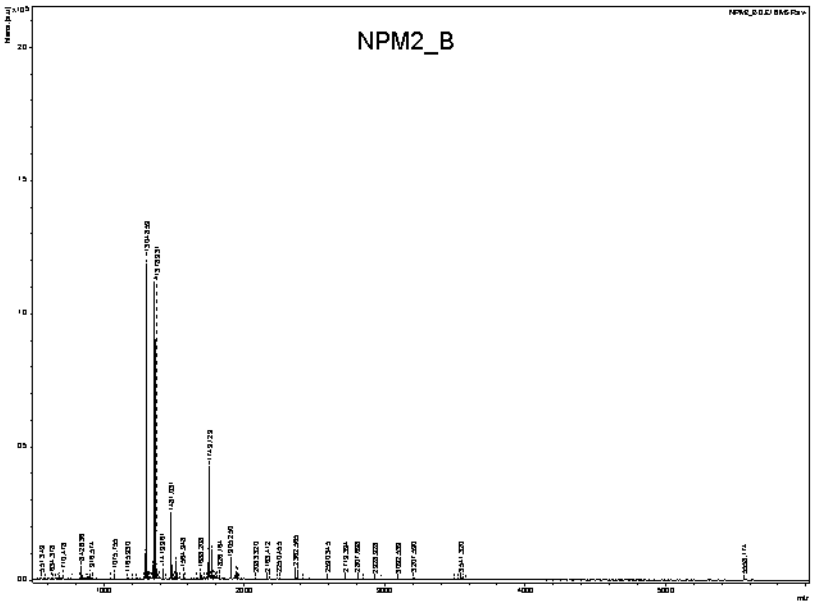
Format As: Concise Protein Summary — Help
Significance threshold p< 0.05 — Max. number of hits AUTO
Preferred taxonomy: All entries

Re-Search All | Search Unmatched

1. NPM2_HUMAN **Mass:** 24308 **Score:** 100 **Expect:** 5.9e-05 **Matches:** 17
Nucleoplasmin-2 OS=Homo sapiens OX=9606 GN=NPM2 PE=1 SV=1

2. HIR2_ARATH **Mass:** 31671 **Score:** 51 **Expect:** 4.6 **Matches:** 14
Hypersensitive-induced response protein 2 OS=Arabidopsis thaliana OX=3702 GN=HIR2 PE=1 SV=1

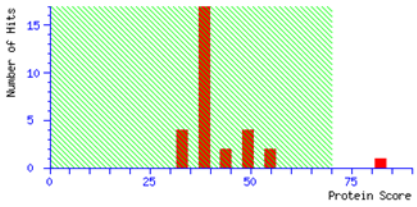
**D**




**E**

**Mascot Score Histogram**

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 70 are significant (p<0.05).




**Concise Protein Summary Report**

Format As: Concise Protein Summary — Help
Significance threshold p< 0.05 — Max. number of hits: AUTO
Preferred taxonomy: All entries
Re-Search All — Search Unmatched

1. NPM2_HUMAN **Mass:** 24308 **Score:** 82 **Expect:** 0.0035 **Matches:** 17
Nucleoplasmin-2 OS=Homo sapiens OX=9606 GN=NPM2 PE=1 SV=1
RL14_RHOPA **Mass:** 13538 **Score:** 34 **Expect:** 2.2e+02 **Matches:** 6
50S ribosomal protein L14 OS=Rhodopseudomonas palustris (strain ATCC BAA-98 / CGA009) OX=258594 GN=rplN PE=1 SV=1

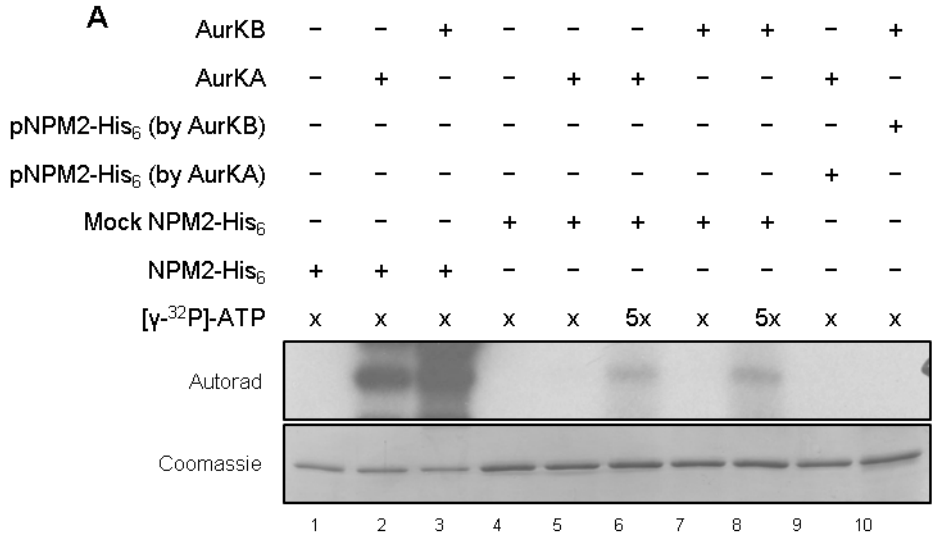
2. ARGR_LACDA **Mass:** 17489 **Score:** 53 **Expect:** 2.6 **Matches:** 9
Arginine repressor OS=Lactobacillus delbrueckii subsp. bulgaricus (strain ATCC 11842 / DSM 20081 / JCM 1002 / NBRC
ARGR_LACDB **Mass:** 17489 **Score:** 53 **Expect:** 2.6 **Matches:** 9
Arginine repressor OS=Lactobacillus delbrueckii subsp. bulgaricus (strain ATCC BAA-365) OX=321956 GN=argR PE=3 SV=1

**Figure 2.35. Mass spectrometry analysis for NPM2 protein identification:** (A) SDS-PAGE profile of $His_6$-NPM2 resolved in a 12% gel. The regions of the gel with the protein bands named 'NPM2_A' and 'NPM2_B' were excised and subjected to MALDI-TOF mass spectrometric analysis. (B, D) Mass spectra of (B) 'NPM2_A' and (D) 'NPM2_B'. (C, E) Mascot database search result snapshots showing significant (score > 70) and non-significant hits for the identification of (C) 'NPM2_A' and (E) 'NPM2_B' protein samples.

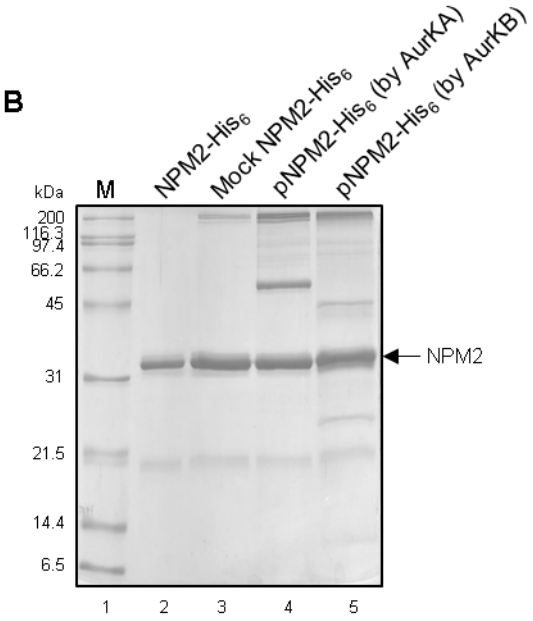
### *2.12.2. Identification of NPM2 residues phosphorylated by Aurora Kinases in vitro by mass spectrometry analysis*

For mass spectrometric identification of *in vitro* phosphorylated sites of recombinant human $His_6$-tagged NPM2 by Aurora Kinase A and B, 1 µg of the substrate protein was taken per reaction for mass phosphorylation as mentioned in Section 2.11.1. The reaction samples were prepared in triplicates. A 'mock phosphorylated' (without the kinase enzyme) sample was also kept as a negative control to negate any non-enzymatic incorporation of ATP in the NPM2 substrate. After the second replenishment with the kinase and ATP, the incubation was extended for 6 – 8 h at 30°C to ensure complete phosphorylation of the substrate. The samples of the same type were pooled together. About $1/3^{rd}$ sample equivalent to 1 µg substrate protein was used in a radioactive *in vitro* kinase assay as described in Section 2.11.1 to check for the incorporation of [$\gamma$-$^{32}$P]-ATP in the mass phosphorylated substrate, in which case it would indicate incomplete mass phosphorylation (Figure 2.36A). The rest of the samples were resolved on 12% gel by SDS-PAGE and stained with Coomassie blue (Figure 2.36B). The specific bands were excised, dried and subjected to mass spectrometric analysis for identification of phosphorylated sites (Proteomics facility at Tohoku University, Sendai, Japan). The results of the mass spectrometry analysis revealed Ser174 and Ser196 as the sites of NPM2 phosphorylated by Aurora Kinase A (Figure 2.36D) and Ser174 as the site of NPM2 phosphorylated by Aurora Kinase B *in vitro* (Figure 2.36E), while no phospho-peptides were detected in the 'mock phosphorylated' sample (Figure 2.36C).

**A**




**B**




**C**

Mock phosphorylation

Coverage – 75%

>sp|Q86SE8|NPM2_HUMAN Nucleoplasmin-2 OS=Homo sapiens GN=NPM2 PE=1 SV=1

MNLSSASSTEEKAVTTVLWGCELSQERRT
WTFRPQLEGKQSCRLLLHTICLGEKAKEEM
HRVEILPPANQEDKKMQPVTIASLQASVLPM
VSMVGVQLSPPVTFQLRAGSGPVFLSGQE
RYEASDLTWEEEEEEEGEEEEEEEDDED
EDADISLEEQSPVKQVKRLVPQKQASVAKK
KKLEKEEEEIRASVRDKSPVKKAKATARAKK
PGFKK

No phospho-peptides detected

**D**

Aurora A phosphorylation

Coverage – 75%

>sp|Q86SE8|NPM2_HUMAN Nucleoplasmin-2 OS=Homo sapiens GN=NPM2 PE=1 SV=1

MNLSSASSTEEKAVTTVLWGCELSQERRT
WTFRPQLEGKQSCRLLLHTICLGEKAKEEM
HRVEILPPANQEDKKMQPVTIASLQASVLPM
VSMVGVQLSPPVTFQLRAGSGPVFLSGQE
RYEASDLTWEEEEEEEGEEEEEEEDDED
EDADISLEEQSPVKQVKRLVPQKQA**S**VAKK
KKLEKEEEEIRASVRDK**S**PVKKAKATARAK
KPGFKK

S174, S196

**E**

Aurora B phosphorylation

Coverage – 75%

>sp|Q86SE8|NPM2_HUMAN Nucleoplasmin-2 OS=Homo sapiens GN=NPM2 PE=1 SV=1

MNLSSASSTEEKAVTTVLWGCELSQERRT
WTFRPQLEGKQSCRLLLHTICLGEKAKEEM
HRVEILPPANQEDKKMQPVTIASLQASVLPM
VSMVGVQLSPPVTFQLRAGSGPVFLSGQE
RYEASDLTWEEEEEEEGEEEEEEEDDED
EDADISLEEQSPVKQVKRLVPQKQA**S**VAKK
KKLEKEEEEIRASVRDKSPVKKAKATARAKK
PGFKK

S174

**Figure 2.36. Mass spectrometric identification of NPM2 residues phosphorylated by Aurora Kinase A and B *in vitro*:** (A) *In vitro* phosphorylation assay with 1 µg of unphosphorylated $His_6$-tagged NPM2 (lanes 1 – 3), mock (where kinase was absent) phosphorylated NPM2 (lanes 4 – 8) and NPM2 mass phosphorylated by Aurora Kinase A (AURKA) (lane 9) and Aurora Kinase B (AURKB) (lane 10) incubated with [$\gamma$-$^{32}$P]-ATP and the respective kinases as indicated. The unphosphorylated substrate was phosphorylated readily, indicated by the greater incorporation of the radiolabeled ATP (lanes 2 – 3). The mass phosphorylated NPM2 did not get further phosphorylated, indicated by the lack of signal (lanes 9 – 10). The mock phosphorylated NPM2 did not get incorporated with the radiolabeled ATP at lower concentrations due to the presence of higher amounts of cold ATP in the mixture which competed with the hot ATP (lanes 5 and 7). However, using 5 times higher amounts of hot ATP, the mock phosphorylated NPM2 could get incorporated with hot ATP (lanes 6 and 8), thereby validating the mass phosphorylation process. Reaction samples with no enzyme were kept as a negative control (lanes 1 and 4). (B) SDS-PAGE profile of unphosphorylated (lane 2), mock phosphorylated (lane 3), AURKA-mediated phosphorylated (lane 4) and AURKB-mediated phosphorylated (lane 5) $His_6$-tagged NPM2 resolved in a 12% gel. The regions of the gel with the protein bands corresponding to NPM2 were excised, dried and subjected to mass spectrometric analysis for phospho-site identification. (C – E) Sites phosphorylated in $His_6$-tagged NPM2 by (C) buffer ('no enzyme' negative control or 'mock phosphorylation'), (D) Aurora Kinase A or (E) Aurora Kinase B *in vitro* respectively as identified by mass spectrometric analysis at 75% coverage.

## *2.13. Bioinformatics and software-based analyses*

### *2.13.1. ChIP-seq analysis*

AcNPM1 ChIP-seq was performed in HeLa S3 cells (Senapati P, Ph.D. thesis, 2014). The library preparation and sequencing were performed at GATC Biotech (Konstanz, Germany) using a HiSeq 2500 Illumina sequencer (Illumina Inc., San Diego, CA, USA). The sequencing reads from each library were adapter-trimmed using TrimGalore (https://github.com/FelixKrueger/TrimGalore) which were then aligned to the human hg19 assembly using Bowtie 2 (Langmead and Salzberg 2012) using the parameters --phred33 --local -N 1. Alignment rates of more than 89% were obtained and the alignment statistics are listed in Table 2.4. Peak calling was performed using the tool MACS2/2.1.1.20160309 (Zhang et al. 2008) using --broad option with default parameters. 24660 AcNPM1 peaks were identified which were conserved between the two replicates after filtering out the hg19 blacklisted regions (Amemiya et al. 2019). The peaks were assigned to the nearest RefSeq TSS using annotatePeaks.pl from the homer package (Heinz et al. 2010). Random peaks were obtained by using shuffle from the BEDTools suite (Quinlan and Hall 2010). The bigwig files for visualization as well as the aggregate plots and heatmaps were generated using deepTools (Ramirez et al. 2014). The broadpeak and bigwig files for ENCODE histone modification ChIP-seq data for HeLa S3 cells were downloaded from

GEO (GSE29611). The Jaccard indexes were calculated using the jaccard function in BEDtools. Combined segmentation bed file for HeLa S3 was downloaded from the UCSC genome browser (http://hgdownload.cse.ucsc.edu/goldenpath/hg19/encodeDCC/wgEncodeAwgSegmentation/wgEncodeAwgSegmentationCombinedHelas3.bed). The motif enrichment analyses were performed using homer. BED files for DNase I hypersensitivity sites for HeLa S3 were downloaded from the UCSC genome browser (http://hgdownload.cse.ucsc.edu/goldenPath/hg19/encodeDCC/wgEncodeAwgDnaseUniform/wgEncodeAwgDnaseUwdukeHelas3UniPk.narrowPeak.gz). The BED files for transcription factor ChIP-seq data were downloaded from GEO (GSE33213, GSE32465, GSE3147).

| Sl. No. | Sample | Number of reads | Sequencing method | Read length | Alignment rate |
|---|---|---|---|---|---|
| 1 | AcNPM1 (replicate – 1) | 122476375 | Single-end | 50 | 98.46% |
| 2 | AcNPM1 (replicate – 2) | 90379345 | Single-end | 50 | 97.69% |
| 3 | Input | 82242857 | Single-end | 50 | 89.38% |

**Table 2.4. Read number information related to AcNPM1 ChIP-seq analysis.**

### *2.13.2. RNA-seq analysis*

The sample preparation for RNA-seq and its quality control analyses have been described in Section 2.4.7.3. The RNA integrity was measured by Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA) and listed in Table 2.2. The RNA-seq libraries were prepared using the KAPA Stranded mRNA-Seq Kit (KAPA Biosystems, Cat. No. KK8421) from 250 ng of total RNA. The library quality was assessed using the Agilent Bioanalyzer 2100 and quantified using a Qubit 3.0 Fluorometer (Life Technologies). The sequencing was performed using an Illumina HiSeq 4000 (Illumina Inc.) at Quick Biology Inc. (Pasadena, CA, USA; via Science Exchange, Palo Alto, CA, USA) in paired-end mode. About 24 – 32 million paired-end reads of length 150 were obtained from each library (Table 2.5). The raw sequences were aligned to the hg19 reference genome using HISAT2 2.1.0 (Kim et al. 2015) using its default parameters. The transcripts were assembled using

the tool Stringtie 1.3.4 (Pertea et al. 2015) with its default parameters and the RefSeq annotation. The assembled transcripts from all libraries were further merged using --merge option in Stringtie. Merged transcript abundances were measured using bedtools coverage and DESeq2 package (Love et al. 2014) was used to normalize the counts and identify the differentially expressed genes (log2 fold change ≥ 0.5 and padj < 0.1). Gene Set Enrichment Analysis (GSEA) (Subramanian et al. 2005) was used to determine the significantly altered gene ontology and pathways.

| Sl. No. | Sample | Number of reads | Sequencing method | Read length | Alignment rate |
|---|---|---|---|---|---|
| 1 | Untreated (UT, replicate – 1) | 26344055 | Paired-end | 150 | 91.70% |
| 2 | Doxycycline-treated (Dox, replicate – 1) | 27165653 | Paired-end | 150 | 91.03% |
| 3 | Untreated (UT, replicate – 2) | 24340029 | Paired-end | 150 | 91.52% |
| 4 | Doxycycline-treated (Dox, replicate – 2) | 32580313 | Paired-end | 150 | 90.68% |

**Table 2.5. Read number information related to RNA-seq analysis.**

### *2.13.3. The Cancer Genome Atlas (TCGA) database analysis*

RSEM normalized transcript expression data (Li and Dewey 2011) and the DNA methylation data (HM450) (Bibikova et al. 2011) of NPM1 and NPM2 genes were downloaded from www.cbioportal.org. The TCGA RNA-seq and the methylation data corresponding to 32 different types of cancers (Appendix Table A.9) were plotted as box-whisker plots using R software (R Core Team 2018).

For studying the correlation between transcript expressions of NPM1 and c-fos, and NPM1 and YY1, the transcript expression data of c-fos was downloaded from www.cbioportal.org while that of YY1 was previously obtained (Behera et al. 2019). The correlations between the expressions of the two genes were determined after performing principal component analysis (PCA) (Leroy 2016; Peterson and Carl 2014) using two different R packages.

### *2.13.4. Image analysis and processing*

All confocal fluorescence images were visualized, analyzed and processed (as per acceptable guidelines) using the Zeiss LSM Image Brower (version 3.5.0.359). The phase-contrast bright-field and fluorescence images were visualized, analyzed and processed (following acceptable guidelines) using the Olympus CellSens Standard 1.17 software (build 16030) and the Zeiss AxioVision software (release 4.8.2). Real-time images were analyzed using the Zeiss AxioVision software (release 4.8.2). Images of gels were visualized, analyzed and processed using the Quantity One version 4.6.7 (build 012) and Image Lab version 2.4.0.03 software (Bio-Rad). Densitometric quantification of bands in images and specific image processing was performed using the ImageJ software (https://imagej.nih.gov/ij/).

### *2.13.5. Statistical analysis and figure representation*

Most of the data in this study (unless stated otherwise) are represented as the mean and standard error of the mean (mean + SEM). All the statistical analyses were performed using GraphPad Prism 8.2.1 software (GraphPad Software Inc., San Diego, CA, USA). Comparison between two means was done by unpaired two-tailed Student's *t*-test, and that between three or more groups were assessed using one-way analysis of variance (ANOVA) followed by Tukey's post-hoc test. A $P$ value of $< 0.05$ was considered statistically significant. All the figures have been generated or compiled using Microsoft PowerPoint software, Office 365 (Microsoft Corporation, Redmond, WA, USA). The models have been partially created using illustration templates from somersault18:24 (https://www.somersault1824.com) available under the Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) license.

# *Chapter 3: Regulation of NPM1 Expression: Implications in Cancer*

## *3.1. Introduction*

As discussed earlier in Section 1.4.1.4, NPM1 is involved in the complex network of pathways operating in cancer cells. The functions of NPM1 implicated as tumor-promoting or tumor-suppressive are often found to be context-dependent and hence difficult to generalize. However, in most of the cases, the tumor-promoting properties of NPM1 are generally manifested due to its overexpression in the cancer cells. NPM1 gets frequently overexpressed in various cancers such as gastric (Tanaka et al. 1992), pancreas (Zhu et al. 2015), colon (Nozawa et al. 1996; Liu et al. 2012c; Wong et al. 2013; Kim et al. 2014), breast (Skaar et al. 1998), skin (Veija et al. 2017; Cecconi et al. 2018), salivary gland (Li et al. 2017b), nasopharynx (Cai et al. 2015), lung (Sekhar et al. 2014; He et al. 2016), ovary (Shields et al. 1997; Kalra and Bapat 2013; Londero et al. 2014), endometrium (Zhou et al. 2014; Zhou et al. 2018), bladder (Tsui et al. 2004), thyroid (Pianta et al. 2010), brain (Yokota et al. 2006; Gimenez et al. 2010; Holmberg Olausson et al. 2015; Phi et al. 2019), liver (Yun et al. 2007; Liu et al. 2012b), prostate (Subong et al. 1999; Leotoing et al. 2008), multiple sclerosis (Vavougios et al. 2018), multiple myeloma (Weinhold et al. 2010) as well as oral cancer which was first reported from our laboratory (Shandilya et al. 2009) and again confirmed through other studies (Coutinho-Camillo et al. 2010; Shandilya et al. 2014a; Peng et al. 2019) (Senapati P, Ph.D. thesis, 2014). Such overexpression of NPM1 in cancer cells presents various advantages to them due to the general pro-survival and the growth and proliferation-promoting functions of NPM1 (Grisendi et al. 2006). In accordance with this fact, it is worthwhile to expect that inhibiting or reducing the overexpression of NPM1 in cancer cells by one or more means, should result in a reduction in their tumorigenic properties such as cell survival, proliferation, migration, and invasion. Indeed that is the case in several cancers such as breast (Chen et al. 2013; Lin et al. 2017; Zeng et al. 2019), brain (Gimenez et al. 2012), endometrium (Lin et al. 2016), colon (Liu et al. 2012c; Wong et al. 2013), prostate (Loubeau et al. 2014), skin (Li et al. 2017a), salivary gland (Li et al. 2017b), nasopharynx (Cai et al. 2015), melanoma (Cecconi et al.

2018), colon (Tang et al. 2018), leukemia (Balusu et al. 2011; Qin et al. 2011; Pozzo et al. 2017; Lin et al. 2019b), and so on.

However, there are also few reports about the downregulated expression of NPM1 in certain cancers. For example, in a cohort of gastric cancer patients, NPM1 expression was found to be heterogeneous in the tumor samples but with a greater trend of downregulation in the tumors compared to matched normal tissues. The downregulation of NPM1 expression was more evident at the protein level. Although *NPM1* mRNA was also significantly reduced in a certain percentage of the tumor samples, its level was increased in some other samples but showed a negative correlation with the protein level expression. Also, NPM1 was found to be less associated with metastasis in that particular cohort of patients (Leal et al. 2014). In another study with a large number of breast cancer patient samples, low NPM1 protein levels were associated with poor prognosis. In cultured cells, high levels of NPM1 was observed in luminal epithelial cells derived from the histologically normal breast tissue, and NPM1, when overexpressed in the invasive MDA-MB-231 cells, showed growth inhibition (Karhemo et al. 2011). In such contexts, NPM1 seems to have tumor-suppressor like properties.

Hence, it is difficult to generalize the expression patterns and functions of NPM1 across all types of cancers. Here lies the importance of regulation of NPM1 expression which can be expected to be governed by various factors and mechanisms, either generally or in a context-dependent manner, given the heterogeneity in the molecular networks operating in different types of cancers and the multitude of molecular functions played by NPM1 under various cellular stages and stresses. For this reason, we were interested to study the regulation of NPM1 expression, especially in oral cancer where we have observed overexpression of NPM1 with increasing grades of the tumors (Shandilya et al. 2009).

We have focussed on transcription factor (TF)-mediated regulation of NPM1 expression in cancer. Several studies have found the induction of NPM1 expression in response to specific growth-promoting stimuli and factors. For example, some of the initial studies related to NPM1 had shown that its levels are induced in highly proliferating cells upon mitogenic signals (Feuerstein and Mond 1987; Feuerstein et al. 1988a; Feuerstein et al. 1988b). Further, several cellular factors regulate the expression of NPM1 under different contexts as revealed by the alteration of NPM1 levels upon perturbation of such factors. For instance, NPM1 expression was found to be regulated by the human pituitary tumor-

transforming gene 1 (hPTTG1) in the breast cancer cell line MDA-MB-231 (Khazaei et al. 2019). However, only a few studies to date have addressed the mechanisms of transcriptional regulation of NPM1 expression. One of the first studies on NPM1 promoter regulation showed the presence of a YY1-binding site in the *NPM1* gene promoter (Chan et al. 1997). YY1 is a transcription factor that exhibits both transcription-activating as well as repressive functions. While specific regions of the YY1 protein have been found to be partially responsible for these opposing functions (Figure 3.1.1A), the context-dependent associations with partner proteins also contribute to YY1-mediated transcriptional regulation (Gordon et al. 2006). Regarding *NPM1* gene regulation in the context of Hepatitis C-virus (HCV) infection, it was found that the YY1-binding site on the NPM1 promoter is bound by YY1-HDAC repressor complex that represses NPM1 expression. Upon infection by HCV, the YY1-HDAC repressor complex is replaced by the YY1-p300-NPM1-HCV core transcription activation complex, which relieves the repression and induces NPM1 expression during HCV infection (Mai et al. 2006) (Figure 3.1.1B). NPM1 was also identified as a myc-responsive gene by a subtractive hybridization screen (Zeller et al. 2001) and recruitment of c-myc at the NPM1 promoter induced by Ras signaling resulted in the transcriptional activation of *NPM1* gene in U1 bladder cancer cells (Yeh et al. 2006). Conversely, retinoic acid (RA)-mediated differentiation of human leukemic HL-60 cells into mature granulocytic cells has been associated with a reduction of total c-myc levels including that at the NPM1 promoter, and a consequent NPM1 downregulation (Yung 2004). In IFN-α-stimulated Jurkat cells (immortalized T lymphocyte cells), phosphorylated STAT3 could transcriptionally upregulate NPM1 expression. In this network, NPM1 itself was responsible for the nuclear translocation of phospho-STAT3 and the concomitant activation of the STAT3 pathway in cancer cells (Ren et al. 2015). mTOR, HIF-1, and AP-2α are other factors that have been shown to transcriptionally activate NPM1 expression under different physiological contexts (Li et al. 2004; Liu et al. 2007a; Boudra et al. 2016).

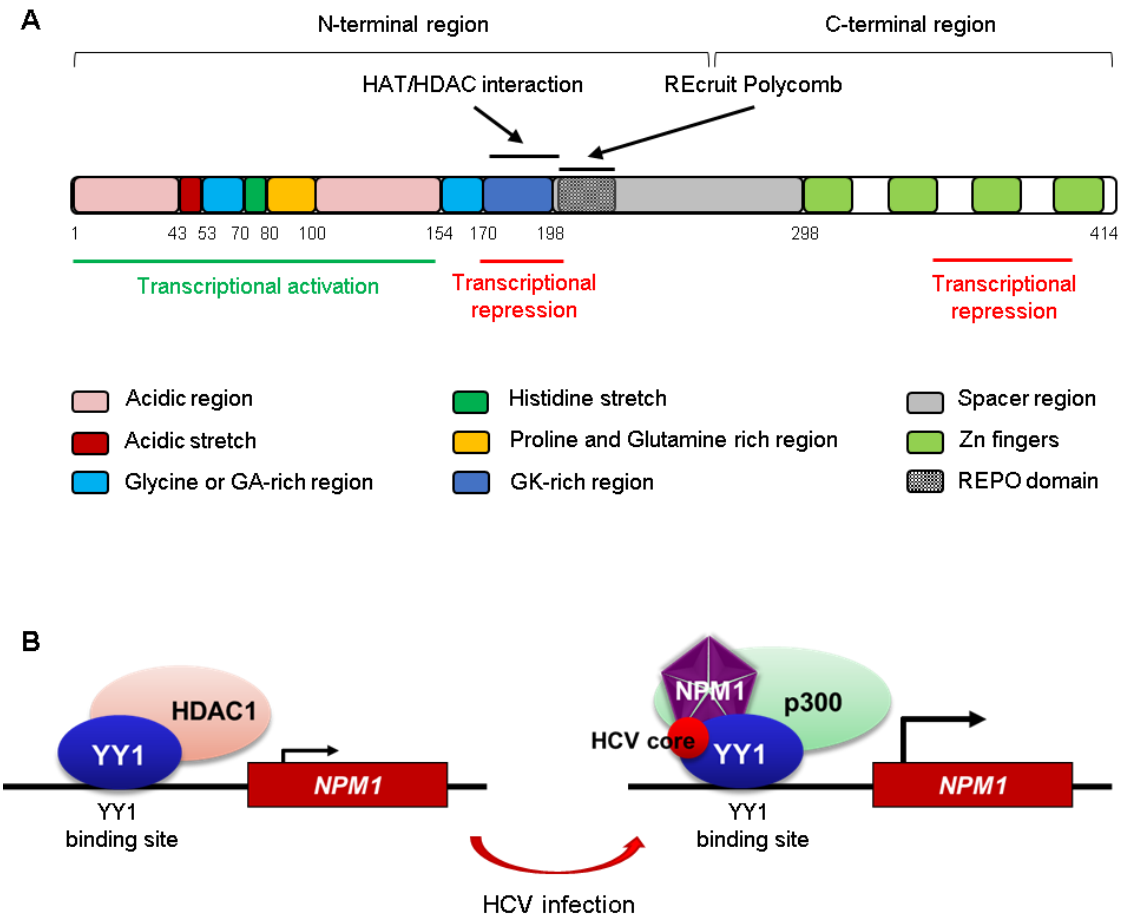



**Figure 3.1.1. Regulation of NPM1 expression by YY1:** (A) Domain organization of transcription factor YY1. The different domains and motifs are color-coded as mentioned. The domain responsible for interactions with histone (lysine) acetyltransferases (HAT) or histone (lysine) deacetylases (HDAC), and that, for interaction with and recruitment of the Polycomb repressor, are indicated. The regions which are generally associated with transcriptional activation or repression by YY1, are also indicated in green and red respectively. The figure has been adapted from (Atchison 2014; Zhang et al. 2016). (B) Model depicting the mechanism of YY1-mediated regulation of NPM1 expression during Hepatitis C virus (HCV) infection (Mai et al. 2006). The YY1-binding site on the NPM1 promoter is normally occupied with the YY1-HDAC1 repressor complex that represses NPM1 expression (indicated by the smaller black arrow). Upon HCV infection, the repressor complex is replaced by the YY1-p300-NPM1-HCV core transcription activation complex, that relieves the repression and induces NPM1 expression (indicated by the bigger black arrow).

These studies shed some light on the regulation of NPM1 expression. However, a comprehensive and systematic analysis of other oncogenic factors regulating NPM1 expression was lacking. Hence, our study was initiated with an unbiased approach by first characterizing the regulatory elements of the *NPM1* gene and identification of putative TFs regulating NPM1 promoter activity (Senapati P, Ph.D. thesis, 2014).

## 3.2. Results

### 3.2.1. Characterization of human NPM1 promoter

Using the information available from the hg19 assembly of the human genome (2012) about the genomic location of the *NPM1* gene and its transcription start site (TSS), different NPM1 promoter constructs in pGL3 basic vector were generated having varying lengths of sequences upstream and downstream of the TSS of *NPM1* gene cloned upstream to the Luciferase reporter gene (Figure 3.2.1A). When Luciferase reporter assays were carried out after transfecting these constructs in HEK-293 cells, three of the constructs, Luc 2, 3 and 5 had shown higher relative Luciferase activity (Figure 3.2.1A). This data indicated that the 1 – 2 kb sequence upstream to the TSS has high promoter activity and may harbor binding motifs for activating transcription factors. In support of this hypothesis, we had observed relatively lower Luciferase activities for the constructs containing more of the region downstream than upstream of the TSS, especially the intron 1 present in NPM1 Luc 1 and NPM1 Luc 4 constructs (Figure 3.2.1A). A careful examination of the sequence upstream and downstream of the NPM1 TSS revealed the core promoter elements, such as $BRE_u$ (upstream transcription factor II B (TFIIB) recognition element), $BRE_d$ (downstream TFIIB recognition element), TATA and INR (initiator motif) (Venters and Pugh 2013), to be present downstream to the TSS (Figure 3.2.1B). To test any significant contribution of these core promoter elements to the promoter activity of the *NPM1* gene, deletion constructs of the NPM1 promoter were generated which lacked the 246 nucleotides downstream to the TSS and thereby these core promoter elements (Figure 3.2.1C). When Luciferase assays were performed using these constructs, a nearly 100-fold decrease in the promoter activity of the deletion constructs was observed compared to the respective full-length constructs (Figure 3.2.1C).

This result showed that the elements present in the sequence downstream to the NPM1 TSS (+1 to +264) have an appreciable effect on the NPM1 promoter activity. However, the deletion of this region did not result in a drastic reduction of the NPM1 promoter activity which indicates that the upstream sequences may be a part of the upstream regulatory region of NPM1, containing binding motifs for several transcription factors which are contributing to the activation of *NPM1* gene expression.

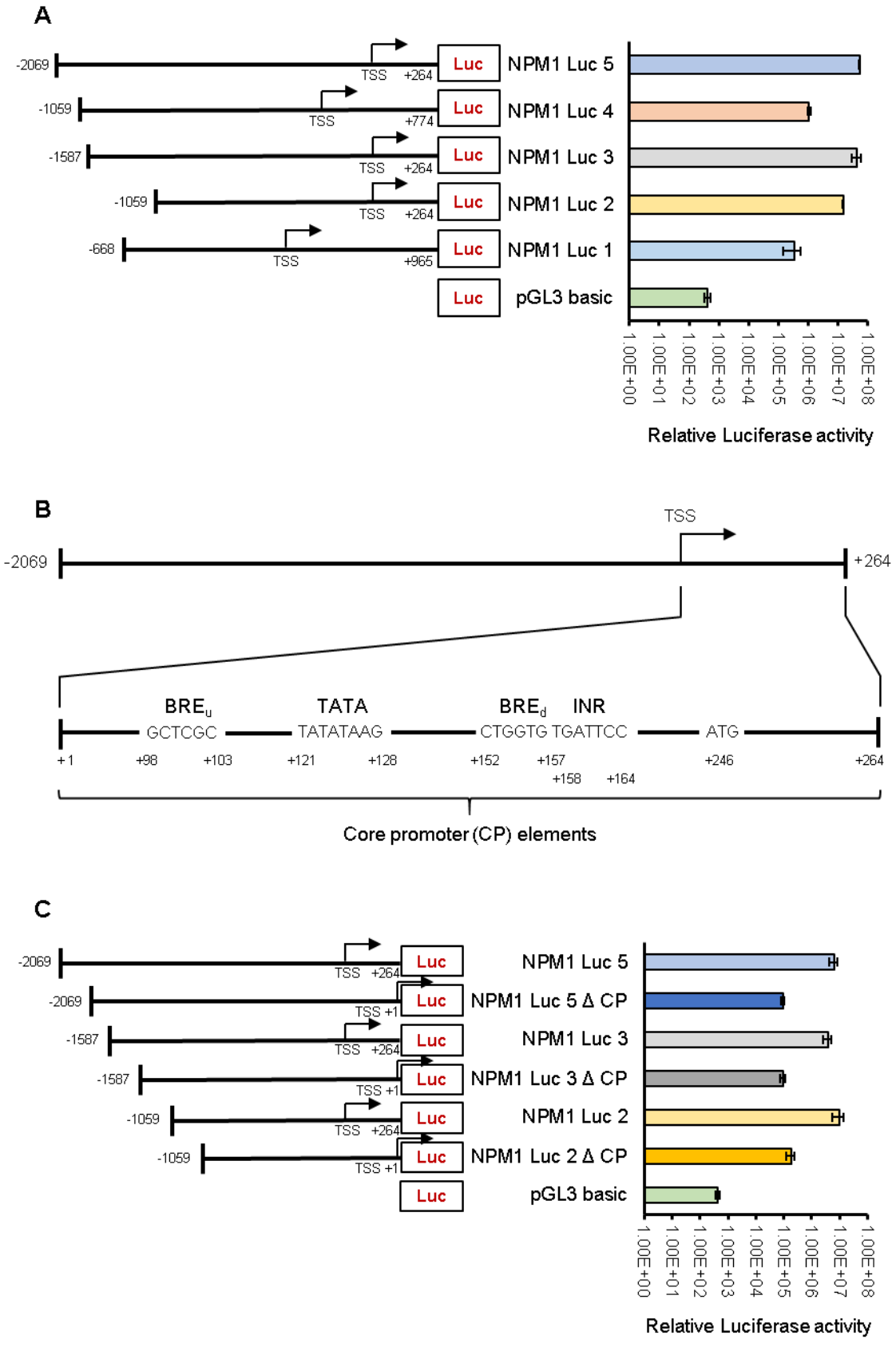



**Figure 3.2.1. Identification of the regulatory elements of NPM1 promoter:** (A, C) Bars represent relative Luciferase activity observed after transfecting 50 ng of the different NPM1 promoter constructs as indicated, for 24 h in HEK-293 cells. Data were normalized to internal transfection

control β-galactosidase. Values are mean ± SEM from two independent experiments. Schematic representations of the constructs used in the experiment are shown in the left. pGL3 basic denotes the empty vector and numbers on the schematics of the constructs denote the positions of the nucleotides indicating the length of the respective promoter fragments. Luc: Luciferase reporter gene, TSS: transcription start site, Δ: deletion, CP: core promoter. (B) Schematic representation of the putative NPM1 promoter region tested: Core promoter (CP) elements are indicated. $BRE_u$: upstream transcription factor II B (TFIIB) recognition element, $BRE_d$: downstream TFIIB recognition element, TATA: TATA box, INR: initiator motif, TSS: transcription start site. ATG denotes the translation start codon. Numbers denote the positions of the nucleotides. The data was obtained from Senapati P, Ph.D. thesis, 2014.

### *3.2.2. Identification of c-fos as a potential transcription factor regulating NPM1 promoter activity*

To identify potential transcription factors regulating NPM1 expression, the upstream 1 kb sequence of the NPM1 promoter was used to predict the various transcription factor (TF)-binding motifs, using the TRANSFAC and Consite databases (Wasserman and Sandelin 2004). Predicted binding sites for several TFs were obtained, out of which the ones with higher confidence scores are listed in Appendix Tables A.10 and A.11. Figure 3.2.2 shows a pictorial representation of the presence of binding motifs of some of the TFs in the NPM1 promoter. Some of these TFs have been previously reported to regulate NPM1 expression such as myc. Among the novel factors, we found c-fos as an interesting candidate to study further, regarding its role in the regulation of NPM1 expression (Figure 3.2.2). c-fos and NPM1 share several common functions and expression patterns. c-fos is an oncogene that is known to regulate genes involved in cell proliferation. The c-fos protein forms a functional heterodimer with members of the JUN family of proteins such as c-jun, forming the transcriptional activator protein complex AP-1 (Activator Protein-1) (Curran et al. 1985; Chiu et al. 1988; Halazonetis et al. 1988). These transcription factors, c-fos and c-jun, are also the first responders to extracellular signals such as growth factors, mitogens, and various stresses, and hence considered as immediate early genes. Their immediate but short-acting responses to such signals lead to different cellular responses, some of which are implicated in differentiation, metabolism, and proliferation (Herschman 1991). Like NPM1, c-fos is also highly expressed in different cancers such as breast (Bland et al. 1995), endometrial (Bamberger et al. 2001), pancreatic (Wakita et al. 1992), and liver (Yuen et al. 2001) cancer. Further, c-fos and c-jun are reported to be overexpressed in oral cancers (Turatti et al. 2005) including the Indian patient cohorts (Sachdev et al. 2008). AP-1 majorly manifests its oncogenic function by regulating genes involved in tumorigenic

processes such as invasion, metastasis, angiogenesis, hypoxia, among others (Karin et al. 1997; Healy et al. 2013). Due to these attributes, the c-fos expression is often correlated with cancer prognosis, which is observed with NPM1 as well. Similar to the induction of c-fos expression, NPM1 expression also gets activated in response to serum that is a source of growth factors (Feuerstein et al. 1988b). Hence, it can be presumed that the regulations of expressions of NPM1 and c-fos are related and could be occurring in the same pathway. The overexpression of NPM1 (Shandilya et al. 2009) and c-fos (Turatti et al. 2005; Sachdev et al. 2008) individually in oral cancer further supported the above-mentioned hypothesis and encouraged us to study the role of c-fos/AP-1 in the regulation of NPM1 expression in cancer, especially in oral squamous cell carcinoma (OSCC).

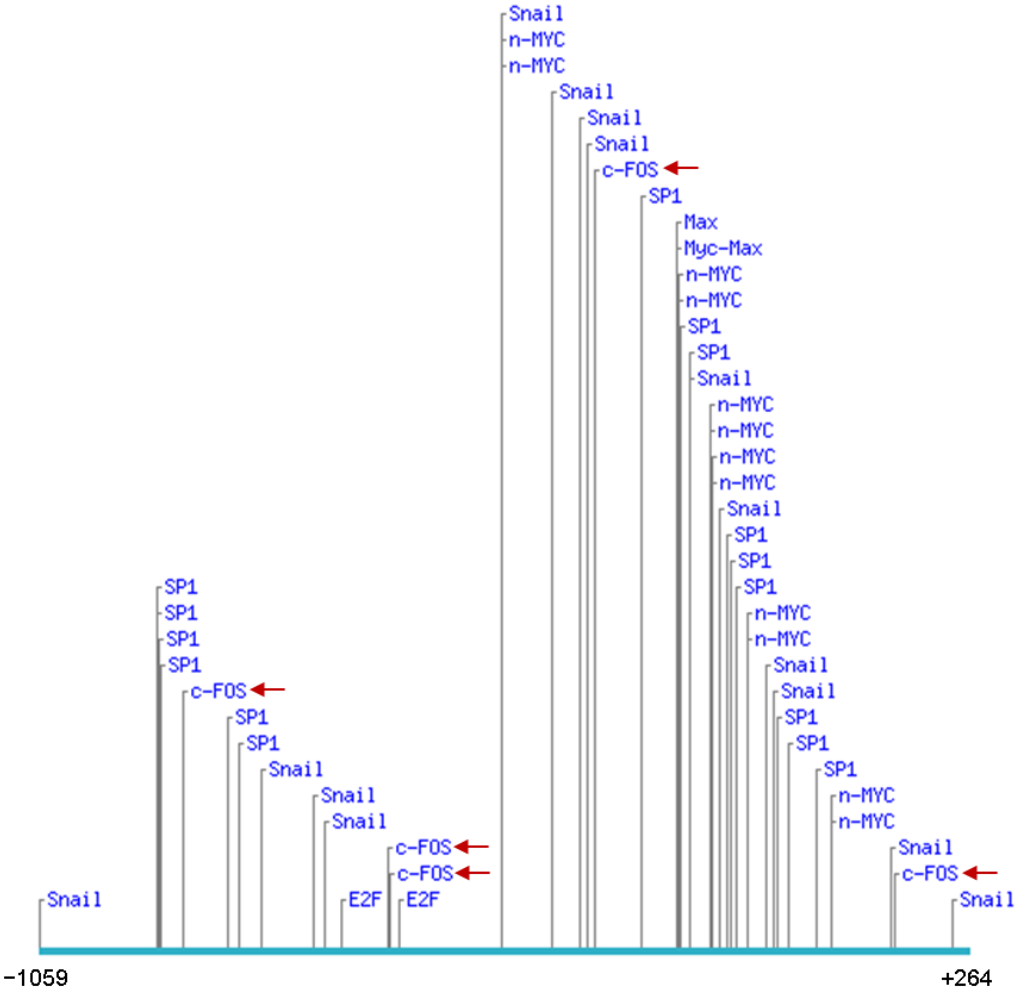



**Figure 3.2.2. Transcription factors having binding sites in the NPM1 promoter:** Screenshot from Consite database showing some of the numerous TFs having binding sites in the human *NPM1* gene promoter DNA (−1059/+264) at 80% TF score cut-off. The c-fos binding sites are indicated with red arrows.

To test if transcription factor c-fos has any effect on the NPM1 promoter activity, Luciferase reporter assay was performed after transfection of the NPM1 promoter construct NPM1 Luc 5 (−2069 to +264), with or without c-fos in HEK-293 cells. Overexpression of c-fos resulted in a significantly higher promoter activity of NPM1 (Figure 3.2.3A). There was a similar extent of significantly increased NPM1 promoter activity when the experiment was done using the core promoter deleted construct of NPM1 Luc 5 (Figure 3.2.3B) which indicated that the c-fos-mediated regulation of NPM1 promoter was possibly occurring through the region upstream to the NPM1 TSS and the downstream sequence did not influence the c-fos-induced increase in the NPM1 promoter activity.

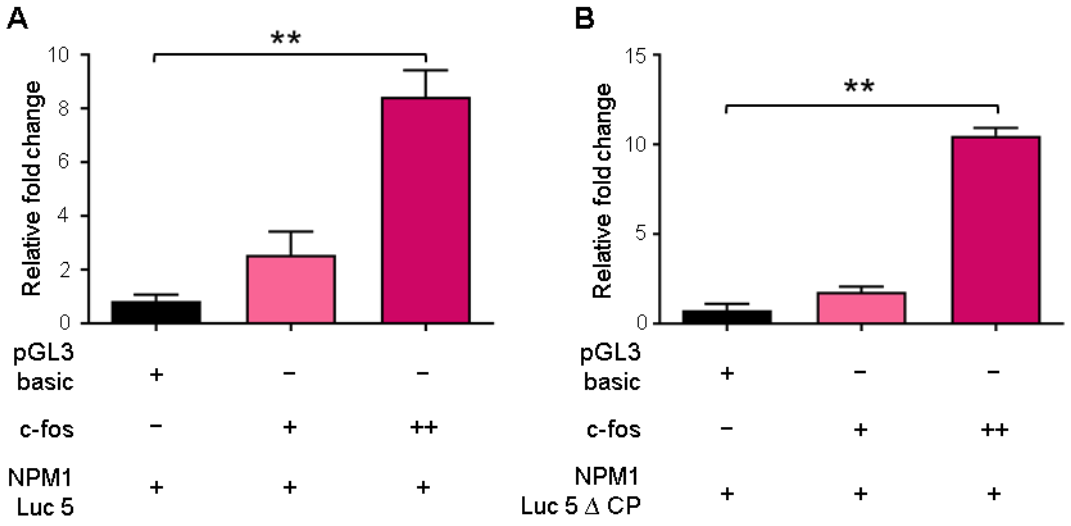



**Figure 3.2.3. Activation of NPM1 promoter activity by transcription factor c-fos:** (A – B) Bars represent fold change in relative Luciferase activity after transfection of HEK-293 cells with 25 ng each of NPM1 promoter construct (A) 'NPM1 Luc 5' (−2069/+264) and (B) 'NPM1 Luc 5 Δ CP' (−2069/+1), FLAG-c-fos (150 ng or 200 ng) or empty vector (EV) (200 ng) as indicated, for 24 h. Data were normalized to internal transfection control β-galactosidase. Values are mean + SEM from two independent experiments having two technical replicates per experiment. Statistical significance was calculated using Student's *t*-test. ***P* < 0.01. The data was obtained from Senapati P, Ph.D. thesis, 2014.

### *3.2.3. Regulation of NPM1 promoter activity by transcription factor c-fos/AP-1*

To reproduce our initial observations made in HEK-293 cells, we performed similar Luciferase reporter assays in the human non-small cell lung adenocarcinoma cell line NCI-H1299. We found that the co-transfection of FLAG-c-fos and the NPM1 promoter construct NPM1 Luc 2 (−1059/+264) increased the NPM1 promoter activity in a dose-dependent fashion (Figure 3.2.4A). Since c-fos is known to heterodimerize with c-jun forming the

functional complex AP-1 to bind to its cognate sites in the DNA (Chiu et al. 1988), we tested the effect of c-jun on the promoter activity of NPM1. We co-transfected H1299 cells with the NPM1 promoter construct NPM1 Luc 2 (−1059/+264), along with FLAG-c-fos and c-jun and performed the Luciferase reporter assay. We observed that co-expression of c-fos and c-jun significantly enhanced the promoter activity of NPM1 by several folds, in comparison to the ectopic expression of c-fos or c-jun alone (Figure 3.2.4B). This shows that AP-1 (c-fos/c-jun heterodimer) positively regulates NPM1 promoter activity.

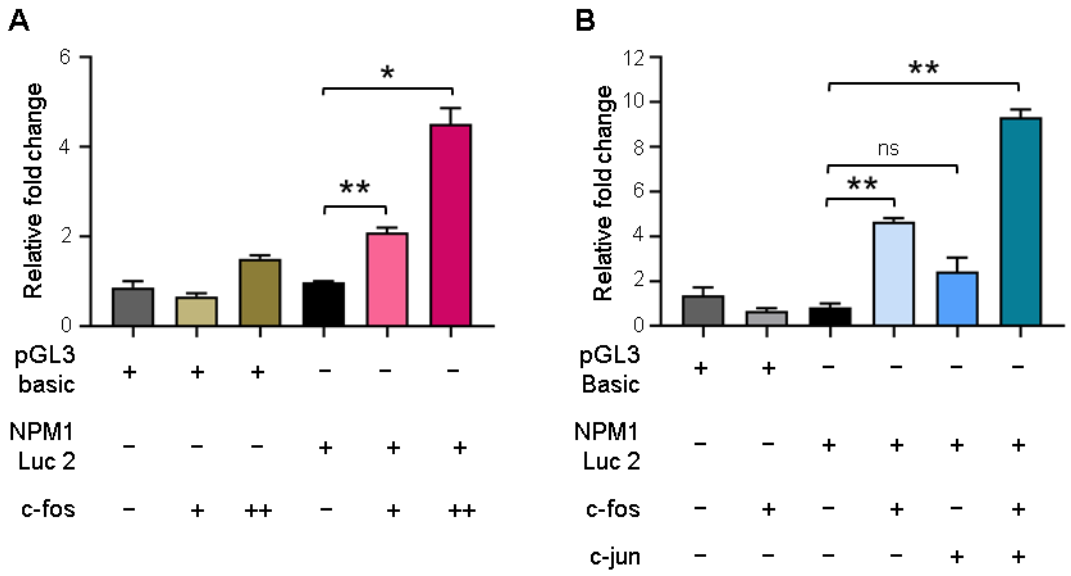



**Figure 3.2.4. Activation of NPM1 promoter activity by transcription factor c-fos/AP-1:** (A) Bars represent fold change in relative Luciferase activity after transfecting 200 ng of empty vector pGL3 basic or NPM1 promoter construct 'NPM1 Luc 2' (−1059/+264), with or without FLAG-c-fos (200 ng or 400 ng) as indicated, in H1299 cells for 24 h. (B) Bars represent fold change in relative Luciferase activity after transfecting 200 ng of empty vector pGL3 basic or NPM1 promoter construct 'NPM1 Luc 2' (−1059/+264), without or with FLAG-c-fos (200 ng) or c-jun (200 ng) or both as indicated, in H1299 cells for 24 h. (A – B) Fold change is relative to pGL3 basic or NPM1 Luc 2. Data were normalized to internal transfection control β-galactosidase. Values are mean + SEM from two independent experiments having two technical replicates per experiment. Statistical significance was calculated using Student's *t*-test. **P* < 0.05, ***P* < 0.01, ns: non-significant.

### *3.2.4. Regulation of NPM1 promoter activity by direct binding of transcription factor c-fos to the NPM1 promoter*

The next step in our study was to test if c-fos directly bound to the NPM1 promoter. Using the Consite database, we obtained the predicted c-fos-binding sites in the NPM1 promoter sequence. At 80% TF score cut-off, we obtained four high-scoring putative binding sites of c-fos/AP-1 in the 1 kb NPM1 promoter sequence upstream of the TSS, which were

predicted based on the consensus binding motif. These sites have been referred to as AP-1-binding sites (AP1BS) in our study, numbered based on their score on the Consite database (Figure 3.2.5A and Table 3.2.1). To check if c-fos could bind to these AP-1-binding sites on the NPM1 promoter, we mutated these four c-fos-binding sites using site-directed mutagenesis as depicted in Figure 3.2.5B. We performed the Luciferase reporter assay after co-transfection of the wild-type (WT) or mutant (mut) NPM1 promoter construct NPM1 Luc 2 and c-fos in H1299 cells. From the experiment, we observed that there was a significant reduction in the transactivation of the NPM1 mutant promoter compared to the WT promoter after the ectopic expression of c-fos (Figure 3.2.5C). However, there was no complete abrogation of the c-fos-mediated activation of NPM1 mutant promoter, indicating that other intact low-scoring binding sites of c-fos might be contributing to its binding and subsequent transactivation of the NPM1 promoter.

| Sl. No. | Binding site | Sequence | From | To | Score | Strand |
|---|---|---|---|---|---|---|
| 1 | AP1BS3 | TTCCTCAC | −853 | −846 | 6.720 | − |
| 2 | AP1BS1 | CTGATTCA | −562 | −555 | 8.378 | + |
| 3 | AP1BS1 | TGATTCAG | −561 | −554 | 7.410 | − |
| 4 | AP1BS2 | TGGCTCAT | −269 | −262 | 7.546 | − |
| 5 | AP1BS4 | GTGATTCC | +158 | +165 | 6.331 | + |

**Table 3.2.1: c-fos/AP-1 binding sites considered in this study.** List of sequences of high-scoring c-fos-binding sites on NPM1 promoter (−1059/+264) as predicted by Consite database at 80% TF score cut-off. Positions with respect to TSS are mentioned.

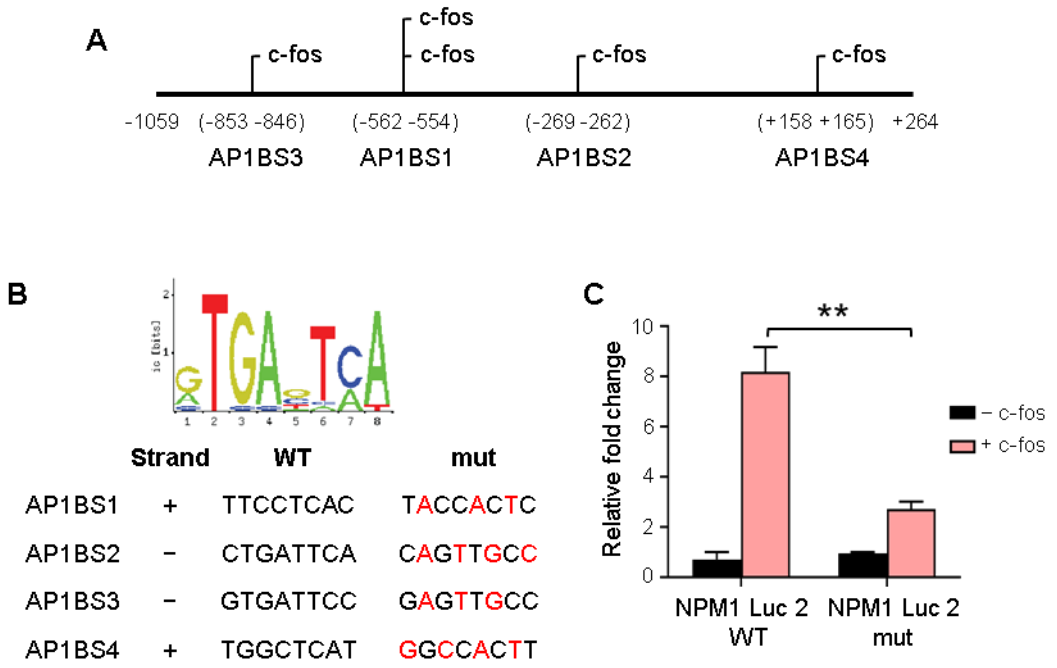



**Figure 3.2.5. Activation of NPM1 promoter activity by direct binding of transcription factor c-fos at its cognate sites on the NPM1 promoter sequence:** (A) Schematic representation of the human NPM1 promoter region (−1059/+264) denoting the positions of high-confidence c-fos-binding sites identified through the Consite database (Table 3.2.1). Numbers denote the positions of the nucleotides. (B) c-fos transcription factor-binding matrix from Consite. The table at the bottom shows the sequences of the AP-1-binding site (AP1BS) mutations. Red letters denote the mutated residues. (C) Bars represent fold change in relative Luciferase activity after transfection of 200 ng of wild-type (NPM1 Luc 2 WT) (−1059/+264) or mutant (NPM1 Luc 2 mut) NPM1 promoter with or without c-fos (400 ng) in H1299 cells for 24 h. Fold change is relative to respective −c-fos control. Data were normalized to internal transfection control β-galactosidase. Values are mean + SEM from two independent experiments having two technical replicates per experiment. Statistical significance was calculated using one-way ANOVA, Tukey's multiple comparisons test. ***P* < 0.01.

We also tested the direct binding of c-fos to the NPM1 promoter the other way around. We mutated several critical residues in the DNA-binding domain of c-fos to generate point mutations namely, the basic residues K153Q, R155Q, R157Q, R158Q, and R159Q that are important for its DNA binding activity. In addition, we mutated several other residues in its leucine zipper region, namely L179V, L186A, and L193V, that are required for c-fos to heterodimerize with c-jun (Kouzarides and Ziff 1988; Neuberg et al. 1989). When we performed the Luciferase reporter assay using the WT or mutant (mut) c-fos by co-transfecting them along with the NPM1 promoter construct NPM1 Luc 2 in H1299 cells, we found that the mutant c-fos could not transactivate NPM1 promoter activity as compared to the WT c-fos (Figure 3.2.6). This proves that c-fos activates NPM1 promoter activity by directly binding to its promoter.

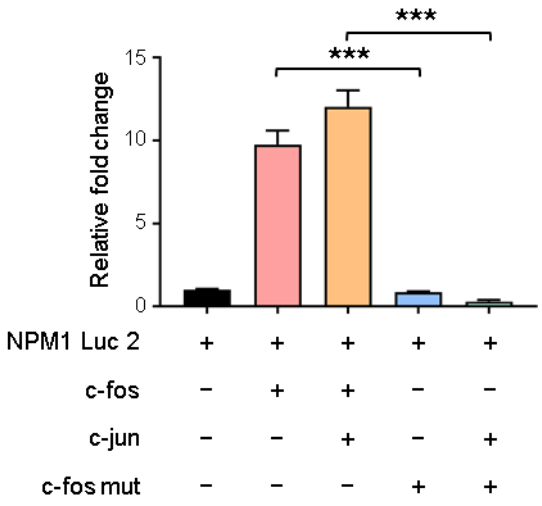



**Figure 3.2.6. Activation of NPM1 promoter activity by direct binding of transcription factor c-fos at the NPM1 promoter:** (A) Bars represent fold change in relative Luciferase activity after transfection of 200 ng of NPM1 promoter construct (NPM1 Luc 2) (−1059/+264) with or without 200 ng of c-fos, c-jun or the DNA-binding deficient mutant of c-fos (c-fos mut) in H1299 cells for 24 h. Fold change is relative to −c-fos control (lane 1). Data were normalized to internal transfection control β-galactosidase. Values are mean + SEM from two independent experiments having two technical replicates per experiment. Statistical significance was calculated using one-way ANOVA, Tukey's multiple comparisons test. ***$P < 0.001$.

### *3.2.5. Role of c-fos in the regulation of endogenous NPM1 expression*

We checked if c-fos/AP-1 could regulate the expression of endogenous NPM1. Upon transfection of FLAG-c-fos in H1299 cells, there was a moderate but significant increase in NPM1 mRNA (Figure 3.2.7A) as well as protein levels (Figure 3.2.7B). Likewise, silencing or knockdown of c-fos by transfection of the specific si-RNA (si-c-fos) or the scrambled negative control si-RNA (si-scr) in H1299 cells, led to a reduction in the NPM1 protein levels compared to the control (Figure 3.2.7C). Similarly, the co-transfection of FLAG-c-fos and c-jun (AP-1) in H1299 cells resulted in a significant increase in NPM1 mRNA (Figure 3.2.7D) which was evident at the protein levels as well (Figure 3.2.7E). The increase in NPM1 protein amounts was observed after transfection of higher doses of c-fos or AP-1 (1.5 µg) compared to the dose required to observe upregulation of its mRNA (500 ng of c-fos or AP-1), which could possibly contribute to the varying folds of upregulation of NPM1 at the transcript and protein levels. However, we do not rule out the possibility of stabilization of NPM1 protein, which could be brought about by various mechanisms and factors after its gene transcription (Zhang et al. 2012).

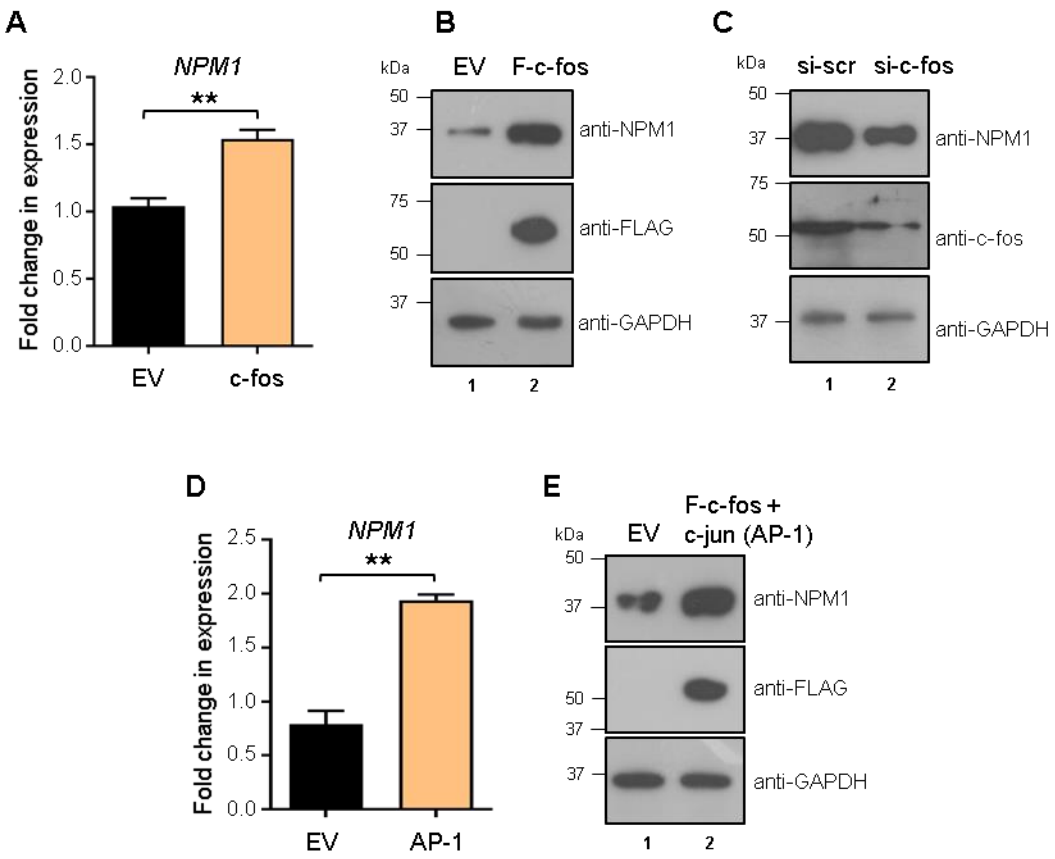



**Figure 3.2.7. Effect of c-fos on the endogenous expression of NPM1:** (A, D) Bars represent fold change in NPM1 mRNA levels as analyzed by RT-qPCR after transfecting 500 ng of empty vector (EV) or (A) FLAG-c-fos and (D) AP-1 (FLAG-c-fos and c-jun, 250 ng each) as indicated, in H1299 cells for 24 h. Internal normalization was done with housekeeping gene β-actin levels. Values are mean + SEM from three independent experiments. Statistical significance was calculated using Student's *t*-test. ***P* < 0.01. (B, E) Western blot analysis after transfection of H1299 cells with 1.5 µg of empty vector (EV) or (B) FLAG-c-fos (F-c-fos) or (E) AP-1 (FLAG-c-fos and c-jun, 750 ng each) as indicated, for 24 h. The upper panel shows western blot with anti-NPM1, the middle panel with anti-FLAG and bottom panel with anti-GAPDH antibody. (C) Western blot analysis after transfection of H1299 cells with 30 nM scrambled negative control si-RNA (si-scr) or c-fos-specific si-RNA (si-c-fos) for 48 h. The upper panel shows western blot with anti-NPM1, the middle panel with anti-c-fos and bottom panel with anti-GAPDH antibody.

### *3.2.6. Occupancy of c-fos at the endogenous NPM1 gene promoter*

To check if c-fos activates NPM1 transcription through direct binding at the endogenous NPM1 promoter, we performed chromatin immunoprecipitation (ChIP) assays in H1299 cells after transfecting them with FLAG-c-fos or the empty vector (EV). We checked the enrichment of c-fos at the sites identified in the *in-silico* prediction analysis in our study, as well as three other sites, deduced from the publicly available c-fos ChIP-seq data in human cancer cell lines from the ENCODE consortium (2012) (Figure 3.2.8A). We

observed significantly increased enrichment of c-fos after transfection, at the sites AP1BS1, AP1BS3, and AP1BS4, identified in our study, as well as the other sites selected from the ENCODE data (Figure 3.2.8B – G). The endogenous region corresponding to the site AP1BS2 did not get enriched in our ChIP assay, probably due to its shearing during the ChIP sonication step. c-fos enrichment was not observed at a negative control region where there was no significant change after c-fos transfection (Figure 3.2.8H). These results indicate that c-fos can bind to its cognate binding sites on the NPM1 promoter bringing about the activation of NPM1 expression in cells.

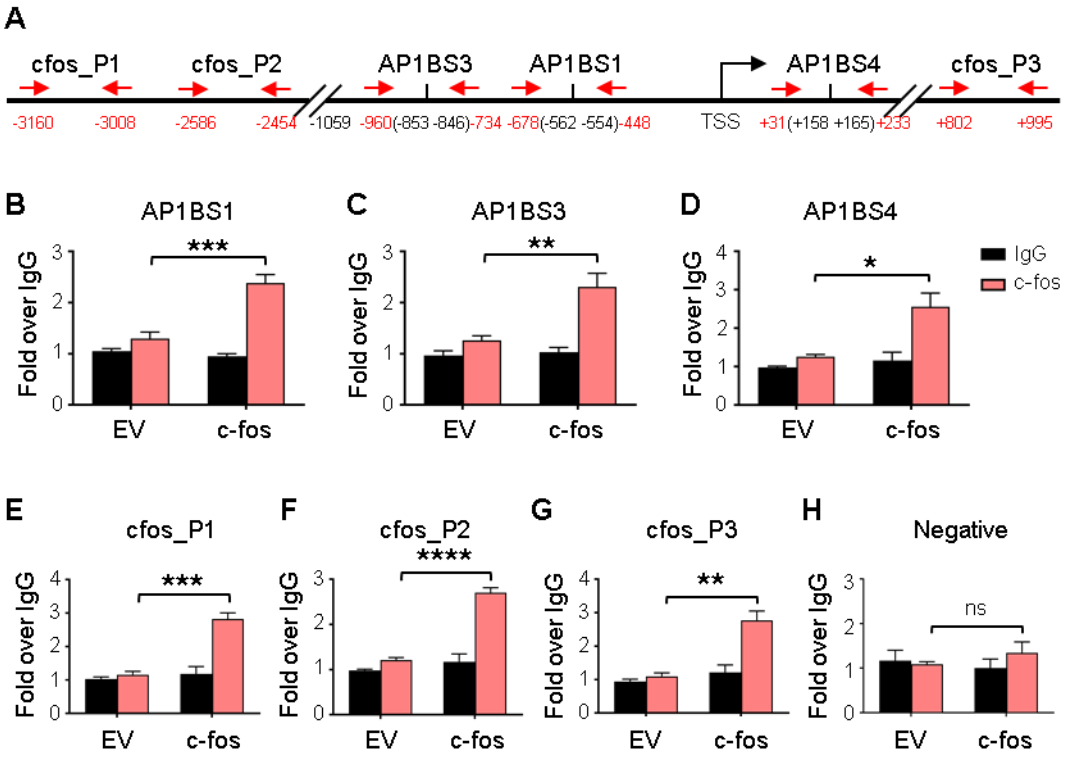



**Figure 3.2.8. Occupancy of c-fos at the endogenous NPM1 promoter:** (A) Schematic representation of the NPM1 promoter region tested, showing the positions of the c-fos/AP-1-binding sites and the primers used for the ChIP experiments. Primer positions are indicated in red. The figure is not drawn to scale. (B – G) c-fos occupancy analyzed by ChIP-qPCR at various c-fos/AP-1-binding sites on the NPM1 promoter as indicated. (H) c-fos occupancy analyzed by ChIP-qPCR at a negative control region, showing no significant c-fos binding. Bars represent the fold enrichment over IgG pull-down in empty vector (EV) or c-fos transfected cells. Values are mean + SEM from three independent experiments. Statistical significance was calculated using one-way ANOVA, Tukey's multiple comparisons test. **P* < 0.05, ***P* < 0.01, ****P* < 0.001, *****P* < 0.0001, ns: non-significant.

### *3.2.7. Role of c-fos/AP-1 knockdown on endogenous NPM1 expression in oral squamous cell carcinoma (OSCC)*

Previous studies conducted from our group and others have shown that NPM1 (Shandilya et al. 2009; Shandilya et al. 2014a) and c-fos (Turatti et al. 2005; Sachdev et al. 2008) individually, are overexpressed in human oral tumor samples. We were interested to investigate any pathophysiological relevance of the c-fos/AP-1-mediated regulation of NPM1 expression in OSCC which could potentially serve as a prognostic marker for oral cancer manifestation and progression, especially in the Indian patient cohorts. We initially performed experiments in human oral tumor-derived cell lines to validate the above-described findings using an approach alternative to overexpression. We screened few human oral cancer cell lines and selected the line UPCI:SCC-29B having high endogenous c-fos expression, to downregulate c-fos/AP-1. We transfected c-fos specific si-RNA (si-c-fos) or non-specific (negative control) scrambled si-RNA (si-scr) in SCC-29B cells and analyzed NPM1 mRNA and protein levels. The c-fos si-RNA could bring about ~75% knockdown of c-fos at the transcript level (Figure 3.2.9A) as compared to scrambled si-RNA control. We observed significant downregulation of NPM1 mRNA (Figure 3.2.9A) and protein levels (Figure 3.2.9C) after c-fos knockdown. Similarly, we also tested the effect of both c-fos and c-jun (AP-1) knockdown on the expression of NPM1 and observed significant downregulation of NPM1 mRNA (Figure 3.2.9B) and protein levels (Figure 3.2.9D) after knockdown of c-fos and c-jun.

As positive and negative controls for the c-fos and AP-1 knockdown, based on literature, we selected and checked the mRNA expressions of few AP-1 target (such as *CCNA2* (Sylvester et al. 1998; Katabami et al. 2005), *CCND1* (Wisdom et al. 1999; Bakiri et al. 2000), *CCND2* (Vanden Bush and Bishop 2011), *CDK1* (Wisdom et al. 1999), *CD44* (Raivich et al. 2004; Nateri et al. 2005) and *MMP1* (Hu et al. 1994; Park et al. 1999)) (Eferl and Wagner 2003) and non-target (such as *VEGFA*, *PDGFA* and *PDGFB* (Orlandini et al. 1996)) genes. The expressions of the AP-1 target genes *CCNA2* and *CDK1* were decreased significantly upon c-fos knockdown whereas those of other targets like *CCND1*, *CCND2*, *CD44*, and *MMP1* did not show significant downregulation in our experimental system (Figure 3.2.9A). However, upon AP-1 (c-fos and c-jun) knockdown, the expressions of most of these genes except for *MMP1*, decreased significantly as compared to that of scrambled si-RNA control knockdown. The expressions of the reported AP-1 non-target genes such as *PDGFA*, *PDGFB* and *VEGFA*, remained mostly unaltered. A modest by

statistically significant downregulation in the mRNA level of *PDGFB* gene was observed upon AP-1 knockdown under the present experimental conditions (Figure 3.2.9B). This could be due to conducting our experiments in a system and conditions, different from what was used in the previously published reports. These results indicate that the transcription factor c-fos/AP-1 is involved in the regulation of NPM1 expression in oral squamous cell carcinoma.

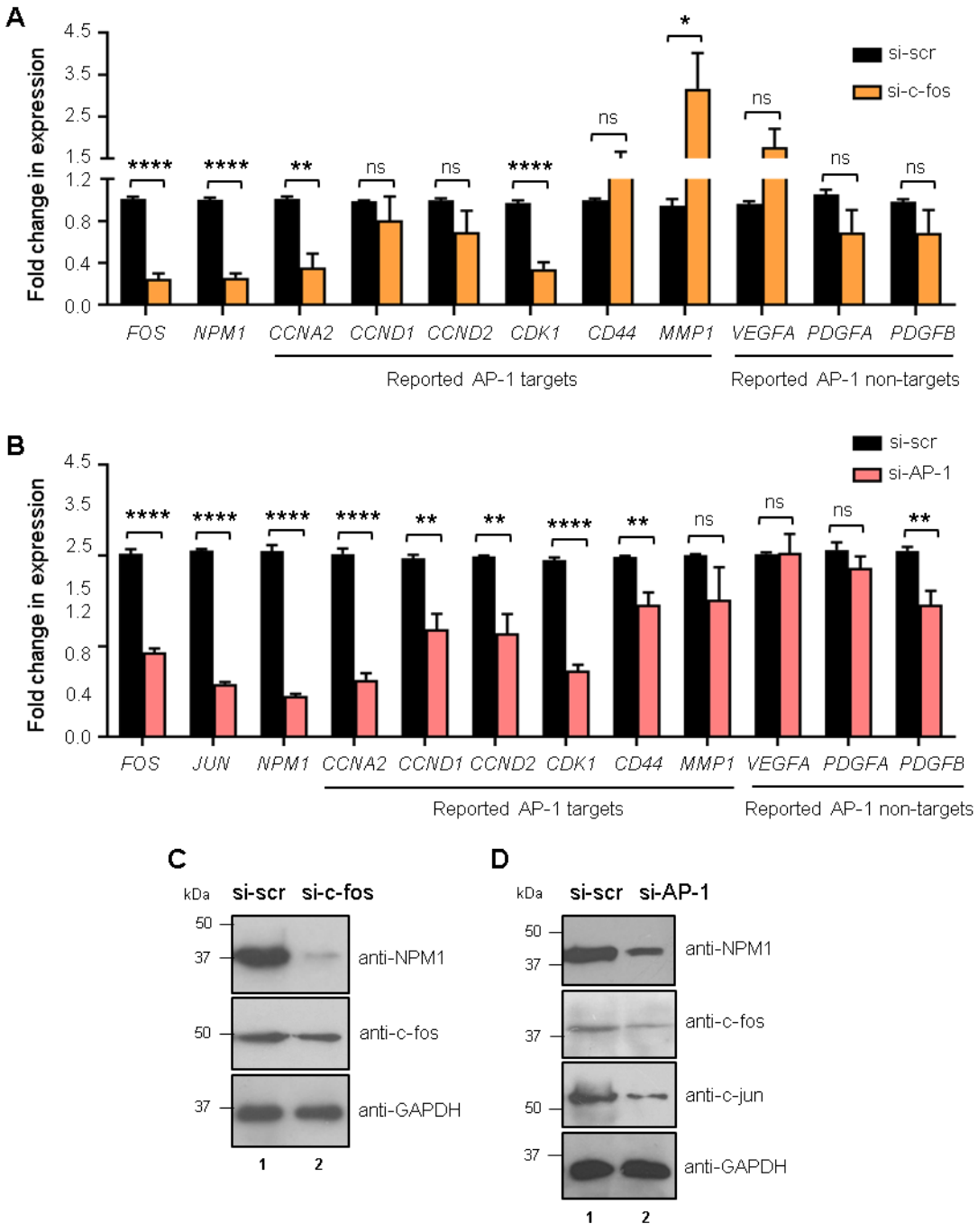
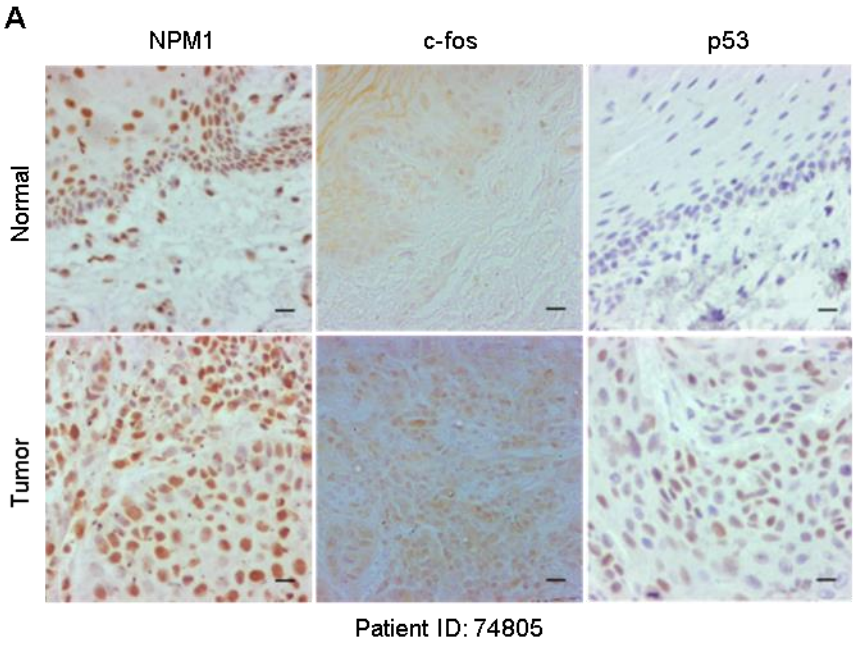



**Figure 3.2.9. Effect of c-fos/AP-1 knockdown on the endogenous expression of NPM1 in OSCC cells:** (A – B) Bars represent fold change in mRNA expression of genes as indicated,

analyzed by RT-qPCR upon transfecting (A) 30 nM si-RNA against human c-fos (si-c-fos) or scrambled negative control (si-scr), or (B) 30 nM each of si-RNA against human c-fos and human c-jun (si-AP-1) or 60 nM of si-scr, for 48 h in UPCI:SCC-29B cells. Internal normalization was done with *18S* rRNA levels. Values represent mean + SEM from four independent experiments. Statistical significance was calculated using Student's *t*-test. **P* < 0.05, ***P* < 0.01, *****P* < 0.0001, ns: non-significant. (C – D) Western blot analysis after transfection of UPCI:SCC-29B cells with (C) 30 nM si-c-fos or si-scr, or (D) 30 nM each of si-RNA against human c-fos and human c-jun (si-AP-1) or 60 nM of si-scr, for 48 h. Upper panels show western blot with anti-NPM1, middle panels with anti-c-fos and anti-c-jun as indicated, and bottom panels, with anti-GAPDH antibody respectively.

### *3.2.8. Expression analysis of NPM1, c-fos, and p53 in human OSCC tissue samples*

We checked the expression status of NPM1 in human oral tumor tissue samples showing c-fos overexpression. Using immunohistochemistry (IHC) analysis of human oral tumor tissue array prepared using the samples collected from local hospitals, we found that both NPM1 and c-fos are overexpressed in the tumor samples compared to the adjacent normal tissues (Figure 3.2.10A – C and Appendix Tables A.12 and A.13), which re-confirmed the previous reports that NPM1 and c-fos are overexpressed in human oral cancer. In addition, we found a significant positive correlation ($r^2 = 0.49$) between c-fos and NPM1 expressions in these oral cancer patient-derived tissue samples (Figure 3.2.10D). This result further supported our findings regarding the regulation of NPM1 expression by c-fos in cancer.



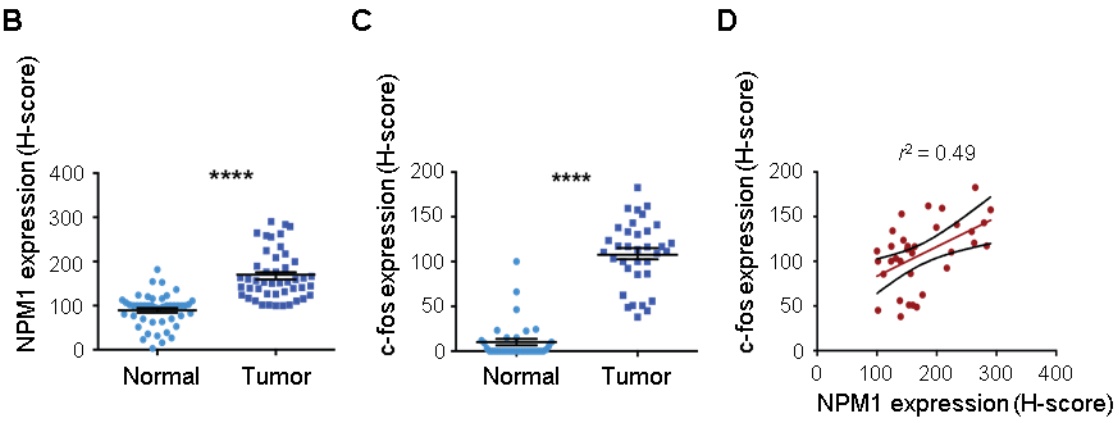



**Figure 3.2.10. Expression status of NPM1, c-fos, and p53 in human OSCC patient-derived tissue samples:** (A) Representative immunohistochemical images showing expressions of NPM1, c-fos, and p53 in matched normal (upper panel) and oral tumor (lower panel) tissue samples. Scale bar is 50 µm. (B) Analysis of NPM1 expression (H-score) between matched normal and oral tumor samples. n = 46. (C) Analysis of c-fos expression (H-score) between matched normal and oral tumor samples n = 35. (B – C) Statistical significance was calculated using the Mann-Whitney test, ****$P$ < 0.0001. (D) Correlation analysis of NPM1 and c-fos expressions in oral cancer samples (Spearman's correlation coefficient $r^2$ = 0.49, **$P$ = 0.003, n = 35).

Cancer is a very heterogenous disease in which various molecular networks are changed in comparison to the normal physiological condition. Variations in the underlying molecular pathways are found even among different types of cancers. However, the enormous amount of scientific research conducted over decades have confirmed some general causal aspects of cancer. One of the most important causes is the abnormal activation of oncogenes (c-fos being one of the most well-known oncogenes) and suppression of functions of tumor suppressors (the most classical one being p53). Often we find the occurrence of both these phenomena together in the cancer cells, and their molecular networks are found to overlap. In the context of inhibition of p53 function in cancer, in about 50% of human cancers, mutations in the *TP53* gene or its loss have been reported. In several instances, the expression and mutational status of p53 substantially contribute to the manifestation and prognosis of human cancers. While the levels of tumor suppressor wild-type p53 are generally kept low by ubiquitination-mediated degradation and tightly regulated through its other post-translational modifications such as phosphorylation and acetylation, mutant p53 often lose this regulation by the ubiquitin ligase Mouse Double Minute 2 (MDM2), which results in its overexpression in tumor cells (Peng et al. 2001; Lukashchuk and Vousden 2007). Several of these mutants have unique functions distinct from the WT protein which contributes to tumor growth and hence known as gain-of-function mutants

(Freed-Pastor and Prives 2012). In this regard, we were interested to check the expression status of p53 in the oral tumor tissue samples where we have found upregulation of NPM1 and c-fos levels. From our IHC analysis, we found a distinct positive staining pattern of p53 only in a few cells in the sections of the tumor samples (Figure 3.2.10A). This could imply the presence of overexpressed, mutant form(s) of p53 in those cells as has been the scenario of mutant p53 expression in cancers observed by other groups (Levine and Oren 2009; Peltonen et al. 2010; Freed-Pastor and Prives 2012).

Expression analysis of p53 and NPM1 in the oral tumor tissue samples showed that tumor samples showing positive staining for p53 expression (High p53) had higher levels of NPM1 (Figure 3.2.11A – B) in comparison to those with negative or lesser staining for p53 expression (Low p53). Their expressions also showed a modest but significant positive correlation in tumor tissues from these patient samples (Figure 3.2.11B). This suggested that NPM1 and p53 could be involved in a regulatory pathway in cancer.

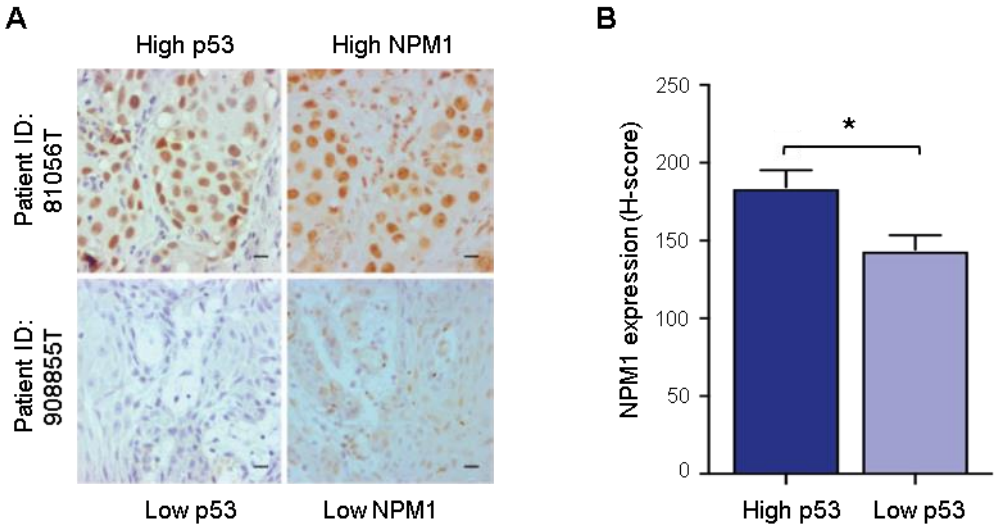



**Figure 3.2.11. Expression status of NPM1 and p53 in human OSCC patient-derived tissue samples:** (A) Representative immunohistochemical images showing NPM1 and p53 expressions in representative oral tumor tissue samples having positive or more intense p53 staining (High p53) and negative or less intense p53 staining (Low p53). Scale bar is 50 µm. (B) Bars represent relative expression (H-score) of NPM1 in oral tumor tissue samples showing High (n = 27) and Low p53 (n = 18) expression. Statistical significance was calculated using the Mann-Whitney test. **P* = 0.01.

There have been few studies which showed that interaction between p53 and NPM1 leads to various context-dependent consequences, such as enhancement of p53-mediated transcriptional activation, stabilization of p53 (Kurki et al. 2004b), or prevention of

premature activation of p53 in response to UV-induced DNA damage (Maiguel et al. 2004). While several such studies have dealt with the association of NPM1 in the p53 pathway (Rubbi and Milner 2003; Brady et al. 2004; Li et al. 2005; Soussi and Wiman 2015), there was no study regarding the transcriptional regulation of NPM1 expression by either wild-type (WT) or mutant p53. These facts together with our observation regarding the positive correlation between NPM1 and p53 expression in the oral tumor tissue samples, motivated us to dig deeper into the role of p53 and its mutants in transcriptional regulation of NPM1 expression. We initiated the investigation by checking if WT p53 had any effect on NPM1 expression.

### *3.2.9. Effect of wild-type (WT) p53 on endogenous NPM1 expression*

When we analyzed the NPM1 promoter sequence using the Consite prediction tool, we found several low-scoring p53 binding sites (11 sites at 65% TF score cut-off) in the 7 kb sequence (−6 kb/+1 kb) of the NPM1 promoter region tested (Figure 3.2.12A). To check if there is any effect of wild-type (WT) p53 on *NPM1* gene transcription, we overexpressed WT p53 in H1299 p53 null cells and observed that there was no significant alteration in NPM1 transcript levels (Figure 3.2.12B) although p53 protein was expressed after the transfection (Figure 3.2.12C).

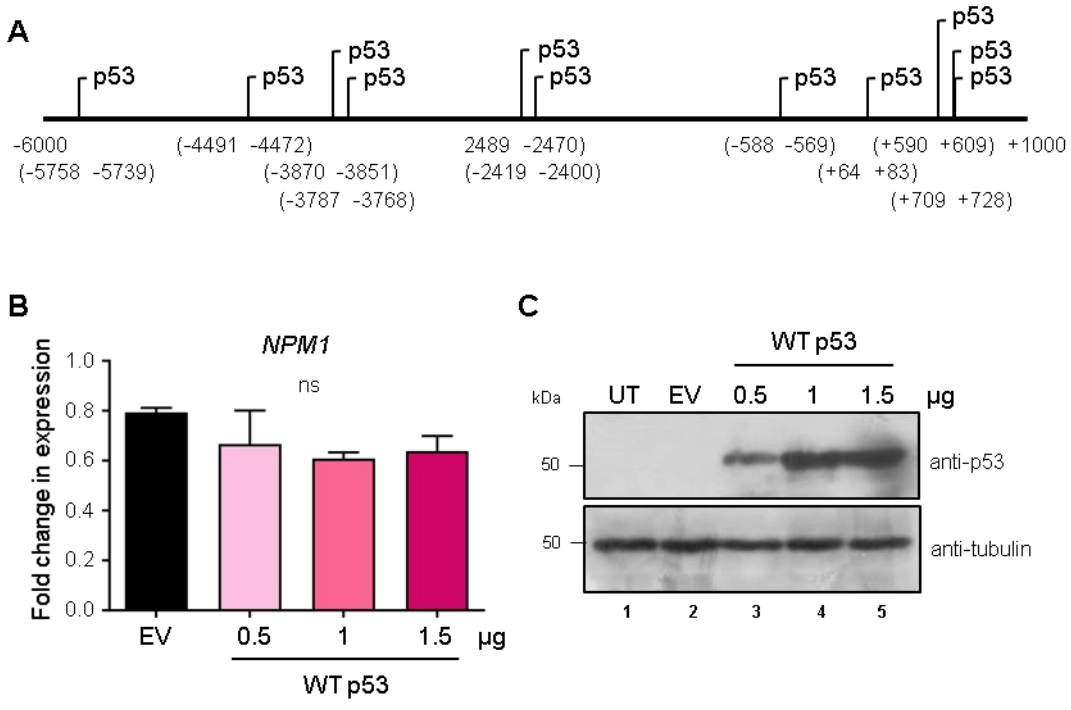
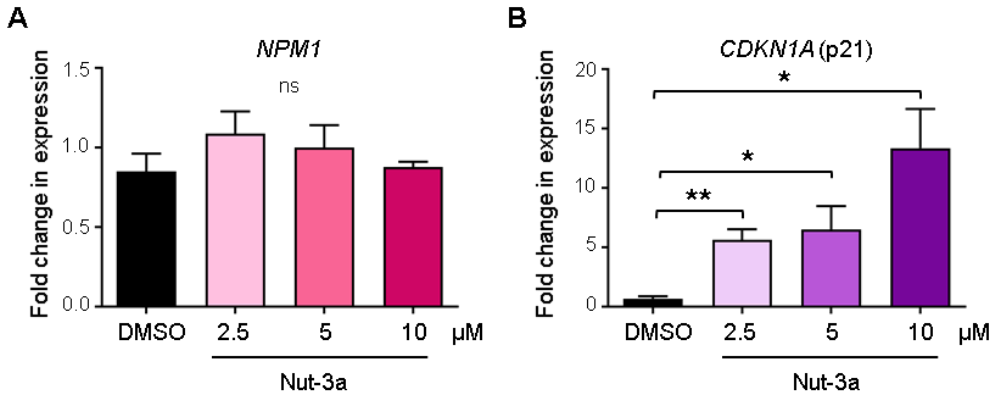


**Figure 3.2.12. Effect of overexpression of WT p53 on *NPM1* gene transcription:** (A) Schematic representation of the human NPM1 promoter region (−6000/+1000) denoting the positions of p53-binding sites (at 65% TF score cut-off) identified through the Consite database. Numbers denote the positions of the nucleotides. Numbers within parentheses denote the positions and lengths of the sites. (B) Bars represent fold change in NPM1 mRNA levels as analyzed by RT-qPCR after transfection of H1299 $p53^{-/-}$ cells with different doses of WT p53 as indicated, for 24 h. Internal normalization was done with housekeeping gene β-actin levels. Values are mean + SEM from two independent experiments and three technical replicates from each experiment. Statistical significance was calculated using Student's *t*-test. ns: non-significant, EV: empty vector. (C) Western blot analysis after transfection of H1299 cells with different doses of WT p53 as indicated, for 24 h. The upper panel shows western blot with anti-p53 and the lower panel shows western blot with anti-tubulin antibody.

We also tested the effect of WT p53 on NPM1 expression through an alternate approach, where we stabilized and enhanced the endogenous levels of WT p53 protein in the human colorectal carcinoma cell line, HCT 116 $p53^{+/+}$ using the small molecule Nutlin-3a, an MDM2 inhibitor (Vassilev et al. 2004). Treatment of HCT 116 cells with Nutlin-3a did not significantly alter the NPM1 transcript levels (Figure 3.2.13A) although the expression of the known p53 target gene, cyclin-dependent kinase inhibitor 1A (*CDKN1A*) or p21 was highly upregulated (Figure 3.2.13B) under the same experimental conditions. These results show that WT p53 has no appreciable effect on the transcriptional regulation of NPM1 expression. This observation was also in accordance with the absence of high-scoring p53-binding sites in the NPM1 promoter sequence as analyzed *in-silico* and mentioned earlier. At the protein level also, treatment of HCT 116 cells with Nutlin-3a in increasing doses for two different time points, namely 6 h and 12 h, did not significantly alter NPM1 expression (Figure 3.2.13C, compare lanes 3 – 6 versus lanes 1 and 2, and 3.2.13D) although the compound treatment resulted in highly enhanced levels of endogenous p53 protein in the cells (Figure 3.2.13C and E).



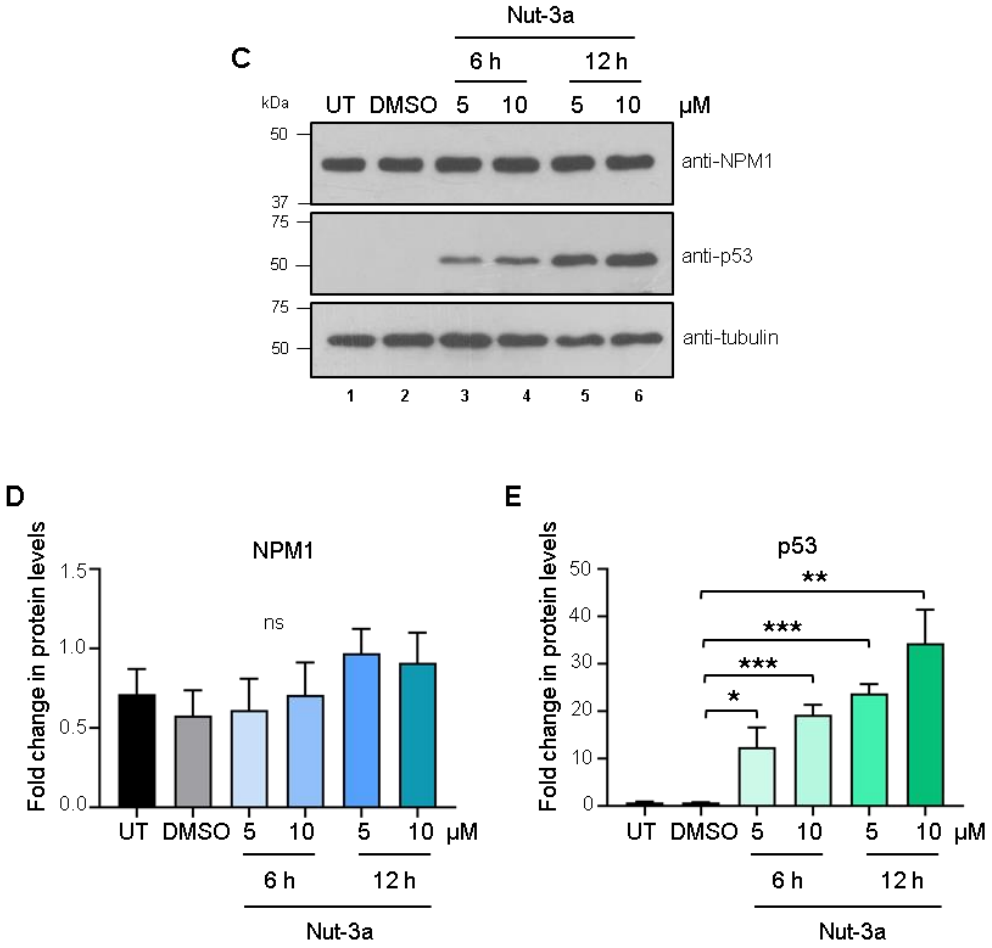



**Figure 3.2.13. Effect of stabilization of endogenous WT p53 on NPM1 expression:** (A – B) Bars represent fold change in mRNA levels of (A) NPM1 and (B) p21 as analyzed by RT-qPCR after treatment of HCT 116 $p53^{+/+}$ cells with different doses of Nutlin-3a (Nut-3a) as indicated, for 24 h. Internal normalization was done with housekeeping gene β-actin levels. (C) Western blot analysis after treatment of HCT 116 $p53^{+/+}$ cells with 5 μM (lanes 3 and 5) and 10 μM (lanes 4 and 6) of Nutlin-3a (Nut-3a) for 6 (lanes 3 and 4) and 12 h (lanes 5 and 6) as indicated. The upper panel shows western blot with anti-NPM1, the middle panel with anti-p53 and bottom panel with anti-tubulin antibody. (D – E) Densitometric quantification of (D) NPM1 and (E) p53 protein levels, normalized to tubulin. Bars represent fold change in protein levels with respect to the untreated (UT) control. (A – B and D – E) Values are mean + SEM from three independent experiments. Statistical significance was calculated using Student's *t*-test. **P* < 0.05, ***P* < 0.01, ****P* < 0.001, ns: non-significant.

### *3.2.10. Role of mutant p53 (R175H) in the regulation of NPM1 expression*

Even though in our studies we did not find any significant effect of WT p53 in the regulation of NPM1 expression, this does not rule out the possibility of the role of p53 mutants in this process. As mentioned earlier, mutations in the *TP53* gene which get selected in cancer, often result in mutant forms of p53 protein that cannot be targeted for ubiquitination-mediated degradation like the wild-type protein. This leads to the

overexpression of stable forms of such p53 mutants in the cancer cells. Since most of these p53 proteins are a result of point mutations, they are detected in cancer cells and tissues with the same antibody that recognizes the wild-type protein. Indeed the initial studies that reported p53 as an oncogene by observing its overexpression in cancer tissue samples, were actually about the mutant forms of p53 in the cancer patient samples (Freed-Pastor and Prives 2012). We presumed that there is a similar scenario of the presence of mutant p53 in our oral cancer patient samples whose expression correlated positively with NPM1 (Figure 3.2.11A – B). The role of mutant p53 in oral cancer pathogenesis has been reported previously in other studies (Li and Zhang 2015; Zhou et al. 2016; Patel et al. 2018; Ragos et al. 2018; Singh et al. 2019). In some instances, the expressions of certain genes such as the oncogenes stathmin and FOXM1, in OSCC or HNSCC, were found to be transcriptionally upregulated by mutant p53 but not by wild-type p53 (Ma et al. 2017; Tanaka et al. 2018). The so-called 'hot-spot' p53 mutants which are found in higher frequencies in the cancer patients, exhibit gain-of-function properties in regulating specific sets of genes by their unique mechanisms which are different from that of WT p53 (Di Agostino et al. 2006; Freed-Pastor and Prives 2012). Based on this premise, we hypothesized that mutant p53 could be regulating NPM1 expression in cancer.

Among the six hot-spot residues in p53 that are most commonly mutated in all types of cancers, R175H is the most frequent mutation, which is also the fourth frequent mutation out of all gene mutations across all cancer types (Di Agostino et al. 2006). R175H belongs to the gain-of-function conformational class of p53 mutants, which is incapable of binding to the DNA without the help of accessory factors. We first tested the effect of R175H p53 overexpression on NPM1 promoter activity. We performed Luciferase reporter assay after co-transfection of H1299 $p53^{-/-}$ cells with NPM1 promoter construct NPM1 Luc 2 (−1059/+264) and R175H p53 expression plasmid. We found a significant increase in the NPM1 promoter activity in the presence of R175H p53 (Figure 3.2.14A). R175H p53 overexpression in H1299 cells also resulted in a significant increase in NPM1 mRNA (Figure 3.2.14B) and protein levels (Figure 3.2.14C). As controls for R175H p53-mediated transactivation of genes, we checked the expression of some of the known targets of R175H p53 such as the cell cycle genes *CCNA2*, *CCNB2*, *CDC25C*, and *CDK1* (Di Agostino et al. 2006), and c-myc (Frazier et al. 1998). Expressions of all these genes were significantly enhanced upon R175H overexpression in the cells (Figure 3.2.14D – H) as reported earlier. It can be presumed that the enhanced expressions of all these genes along with NPM1

contribute collectively to the process of tumorigenesis. This validates our system and shows that R175H p53 can positively regulate NPM1 expression.

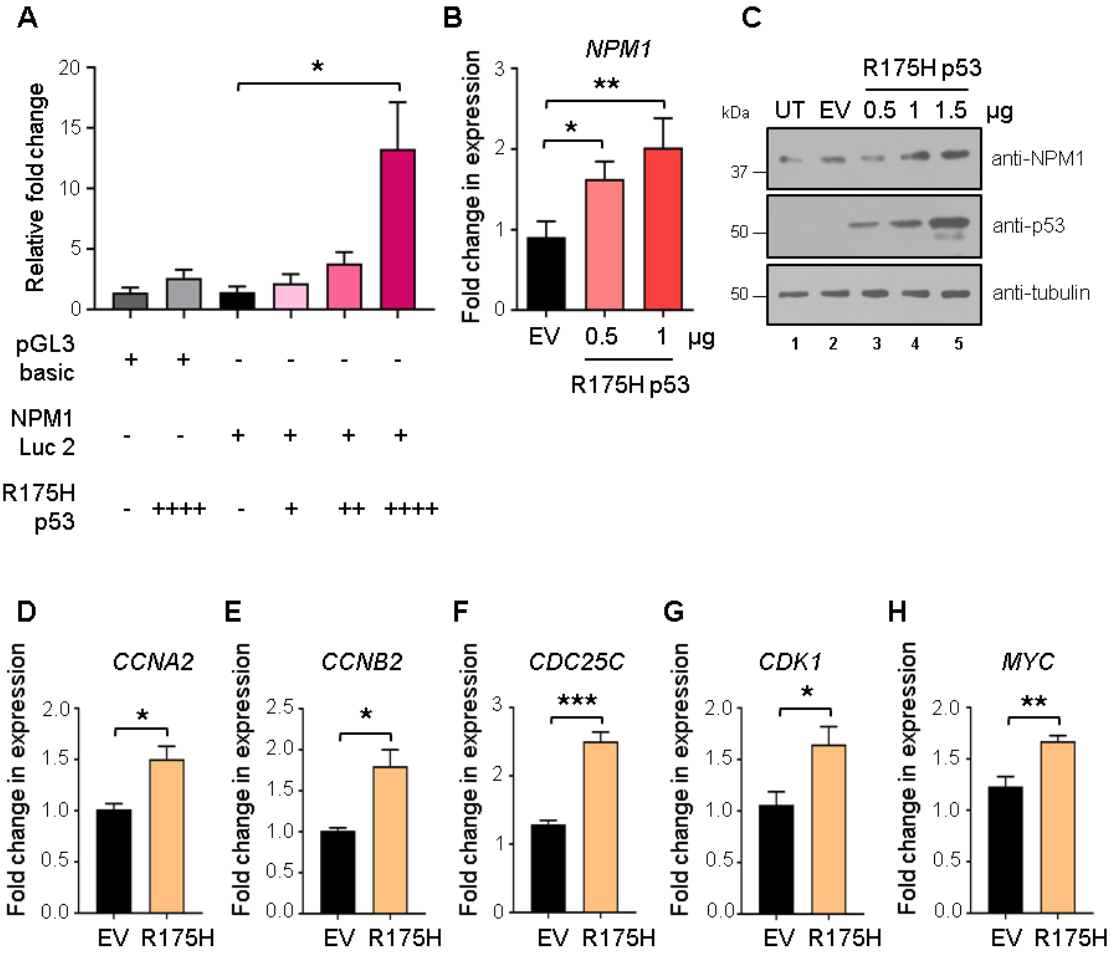
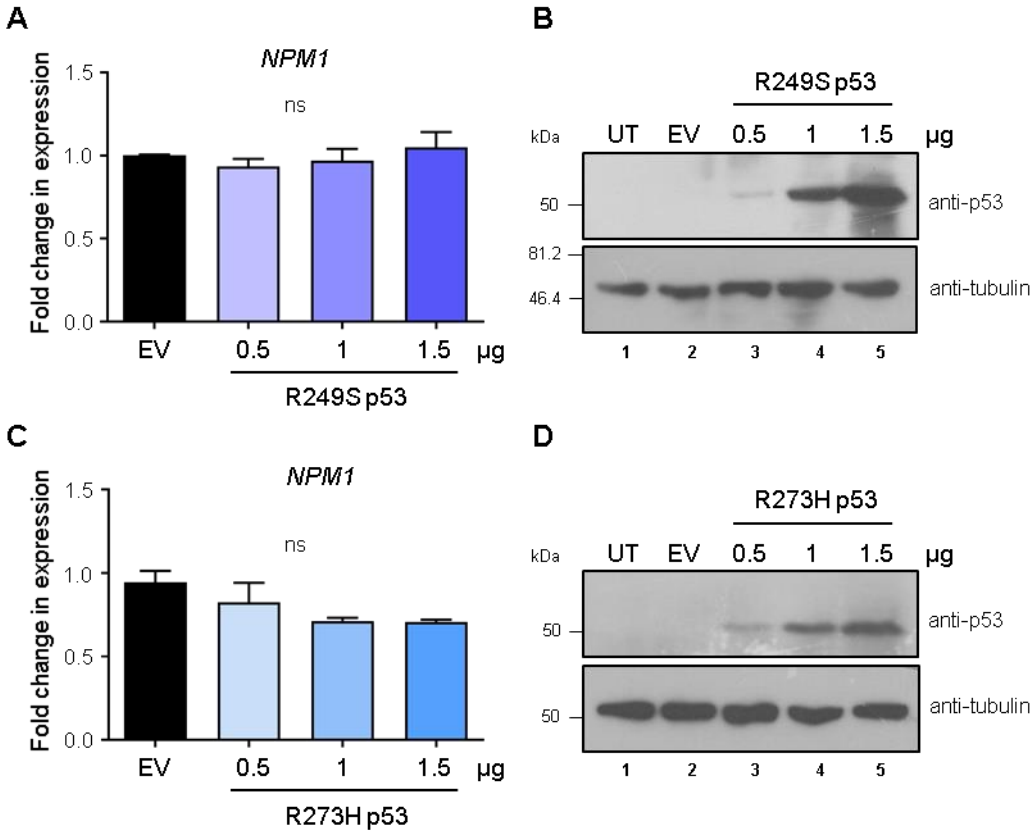



**Figure 3.2.14. Effect of R175H p53 overexpression in the regulation of NPM1 expression:** (A) Bars represent fold change in relative Luciferase activity after transfection of H1299 $p53^{-/-}$ cells with 200 ng of empty vector pGL3 basic or NPM1 promoter construct 'NPM1 Luc 2' (−1059/+264) with or without R175H p53 as indicated, for 24 h. Doses of R175H p53 transfected are 100 ng (lane 4), 200 ng (lane 5) and 400 ng (lanes 2 and 6). Data were normalized to internal transfection control β-galactosidase. Values are mean + SEM from two experiments and two technical replicates in each experiment. Statistical significance was calculated using one-way ANOVA, Tukey's multiple comparisons test. (B) Bars represent fold change in NPM1 mRNA levels as analyzed by RT-qPCR after transfection of H1299 cells with empty vector (EV) or R175H p53 as indicated, for 24 h. Internal normalization was done with housekeeping gene β-actin levels. Values are mean + SEM from three independent experiments. (C) Western blot analysis after transfection of H1299 cells with empty vector (EV) or different doses of R175H p53 as indicated, for 24 h. The upper panel shows western blot with anti-NPM1, the middle panel with anti-p53 and bottom panel with anti-tubulin antibody. UT: untransfected. (D – H) Bars represent fold change in mRNA levels of (D) CCNA2 (E) CCNB2 (F) CDC25C (G) CDK1 and (H) c-myc as analyzed by RT-qPCR after transfection of H1299 cells with 500 ng of empty vector (EV) or R175H p53 as indicated, for 24 h. Internal normalization was done with housekeeping gene β-actin levels. Values are mean + SEM from two independent experiments and three technical replicates from each experiment. Statistical significance was calculated using Student's *t*-test. (A – B and D – H) **P* < 0.05, ***P* < 0.01, ****P* < 0.001.

### *3.2.11. Role of mutant p53 (R249S, R273H) overexpression in the regulation of NPM1 expression*

Since we observed a positive effect of R175H p53 in the regulation of NPM1 expression, it was intriguing for us to test the role of other similar hot-spot gain-of-function mutants of p53 in the regulation of NPM1 expression. We screened a few such hot-spot mutants of p53 such as R249S and R273H p53. We observed that none of these mutants could induce *NPM1* gene transcription (Figure 3.2.15 A and C) under the tested experimental conditions, although the respective transfections led to high expressions of the p53 proteins (Figure 3.2.15B and D) and these specific mutants had similar extent of stability and expressions in the cells (Figure 3.2.15E). These results imply that among these p53 mutants tested, the effects of R175H p53 on NPM1 expression were specific to this mutant. Since we were particularly interested in studying the transcriptional regulation of NPM1 by mutant p53, and these mutants were not found to enhance NPM1 transcript levels, we did not further check the protein levels of NPM1. Further, the stability and upregulation of NPM1 protein could be brought about by other mechanisms that need not result in an outcome directly proportional to its transcript levels. Hence, to study this aspect, it would require further investigations, which were beyond the scope of the present study.



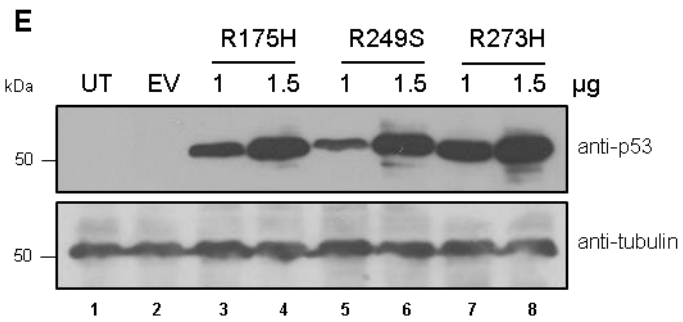



**Figure 3.2.15. Effect of R249S and R273H p53 overexpression in the regulation of *NPM1* gene transcription:** (A, C) Bars represent fold change in NPM1 mRNA levels as analyzed by RT-qPCR upon transfection of H1299 $p53^{-/-}$ cells with empty vector (EV) or different doses of (A) R249S and (C) R273H p53 as indicated, for 24 h. Internal normalization was done with housekeeping gene β-actin levels. Values are mean + SEM from two independent experiments and three technical replicates from each experiment. Statistical significance was calculated using Student's *t*-test. ns: non-significant. (B, D – E) Western blot analysis after transfection of H1299 cells with different doses of R249S, R273H, and R175H p53 as indicated, for 24 h. The upper panel shows western blot with anti-p53 and the bottom panel shows western blot with anti-tubulin antibody respectively. UT: untransfected.

### *3.2.12. Role of mutant p53 (R175H) overexpression in the regulation of NPM1 expression in head and neck squamous cell carcinoma (HNSCC)*

Since our initial observations of the positive correlation of NPM1 and p53 expressions were made in human oral tumor tissue samples, we wanted to validate our finding of the positive effect of R175H p53 on NPM1 expression in the oral cancer background. The most appropriate cell line available to us in this regard was the head and neck squamous cell carcinoma (HNSCC) cell line UM-SCC-1 which is p53 null. We generated a stable cell line for the constitutive expression of the empty vector or 3xFLAG-tagged R175H p53 in the UM-SCC-1 background and checked the expression of NPM1 in these cells. We observed that the mRNA levels of NPM1, as well as the other reported R175H target genes, were significantly high in the R175H expressing cells compared to the empty vector control (Figure 3.2.16A). The levels of NPM1 protein was also significantly high in the R175H expressing cells compared to vehicle control (Figure 3.2.16B – C). This validated our findings in oral cancer as well.

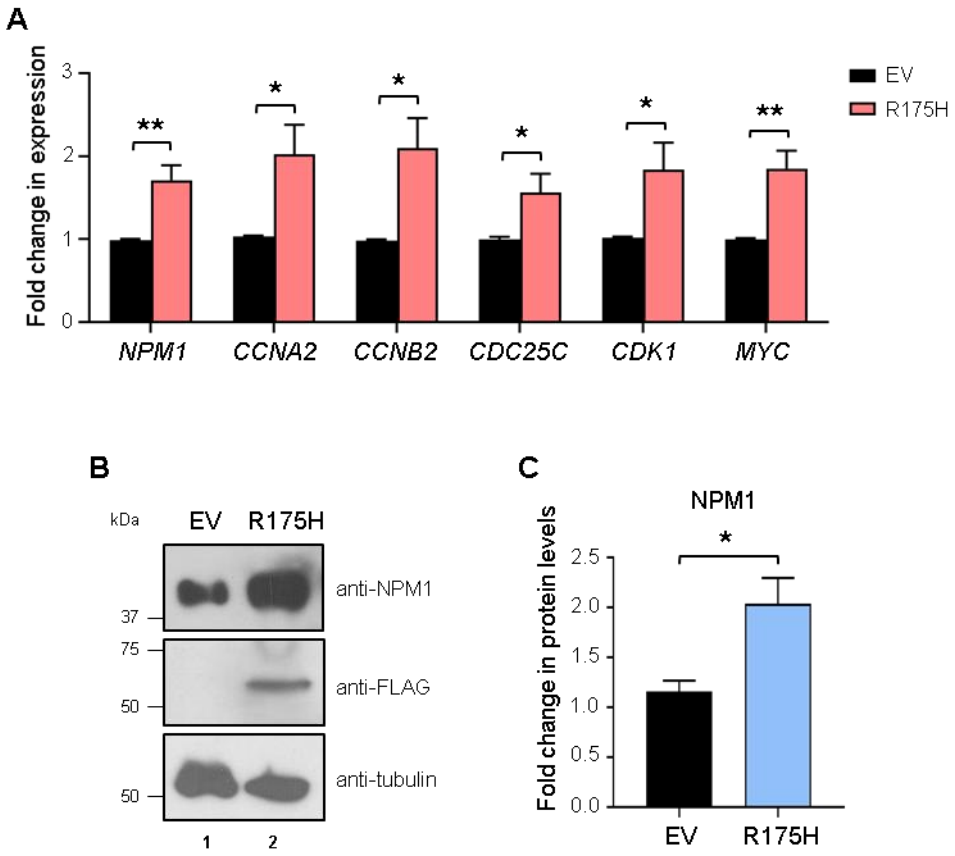



**Figure 3.2.16. Effect of R175H p53 overexpression in the regulation of NPM1 expression in HNSCC:** (A) Bars represent fold change in mRNA levels of NPM1 and other R175H target genes as indicated, analyzed by RT-qPCR upon stable constitutive overexpression of empty vector (EV) or 3xFLAG-tagged R175H p53 in UM-SCC-1 $p53^{-/-}$ cells. Internal normalization was done with housekeeping gene β-actin levels. (B) Western blot analysis after stable constitutive overexpression of empty vector (EV) or 3xFLAG-tagged R175H p53 in UM-SCC-1 cells. The upper panel shows western blot with anti-NPM1, the middle panel with anti-FLAG and bottom panel with anti-tubulin antibody respectively. (C) Densitometric quantification of NPM1 protein levels, normalized to tubulin. Bars represent fold change in protein levels with respect to the empty vector control. (A, C) Values are mean + SEM from four independent experiments. Statistical significance was calculated using Student's *t*-test. **P* < 0.05, ***P* < 0.01.

### *3.2.13. Mechanism of R175H-mediated activation of NPM1 expression: Interaction of c-fos and mutant p53*

From our study so far, we have found that c-fos and R175H p53 individually can activate *NPM1* gene transcription. While we demonstrated the occupancy of c-fos at the c-fos/AP-1 binding sites on the NPM1 promoter by ChIP assays (Figure 3.2.8B – G), R175H itself is known to be incapable of binding to the DNA to cause gene transactivation, without the help of accessory factors. This mechanism is employed by several gain-of-function p53 mutants for bringing about gene transactivation. Some of the reported accessory factors are

nuclear transcription factor Y (NF-Y) (Di Agostino et al. 2006), DNA topoisomerase II binding protein 1 (TopBP1) (Liu et al. 2011), Sp1 (Gualberto and Baldwin 1995; Chicas et al. 2000; Vogiatzi et al. 2016), Vitamin D Receptor (VDR) (Stambolsky et al. 2010), Sterol regulatory element-binding protein (SREBP) (Freed-Pastor et al. 2012), Ets-1 (Sampath et al. 2001), to name a few (Freed-Pastor and Prives 2012).

To test if a similar mechanism could operate in this scenario, we first checked if c-fos and R175H could interact, which could potentially lead to the recruitment of this mutant p53 to the NPM1 promoter. We performed co-immunoprecipitation experiments using the anti-FLAG antibody after overexpressing FLAG-tagged c-fos and untagged R175H p53 together in H1299 cells. The results showed that R175H p53 was pulled down with FLAG-c-fos (Figure 3.2.17A), which indicated that these two proteins interact. c-fos and R175H p53 also colocalized in the nucleus of the cells co-transfected with these expression plasmids, as observed by co-immunofluorescence analysis (Figure 3.2.17B).

To check if this interaction was specific to R175H p53 and indicating its absolute requirement for R175H p53-mediated *NPM1* gene transactivation, we performed similar co-immunoprecipitation experiments for R249S and R273H p53 mutants which did not show any effect on *NPM1* gene expression and observed that these mutants individually could also interact with FLAG-c-fos (Figure 3.2.17C – D). To test if there is any preferential interaction of c-fos with any of these mutants, we performed the co-immunoprecipitation experiment to detect the interaction of FLAG-c-fos and R175H, R249S or R273H, at the same time and under the same experimental conditions. The results showed that all these three p53 mutants could interact with c-fos almost with similar affinity (Figure 3.2.17E). However, this experiment is majorly qualitative and in order to conclude on the specific affinity of the interactions, further biochemical and biophysical investigations would be needed. Nevertheless, these results imply that the interaction between R175H p53 and c-fos per se is not the sole determining factor behind the R175H p53-mediated regulation of NPM1 expression, which leaves scope for further study.

To test if the interaction between R175H p53 and c-fos is physiologically true and not just an artifact of overexpression, we performed an immunoprecipitation experiment in the breast cancer cell line AU565 having naturally mutated R175H p53. The results showed that endogenous R175H p53 was pulled down with immunoprecipitated endogenous c-fos in these cells (Figure 3.2.17F). This proves that R175H p53 can indeed interact with c-fos.

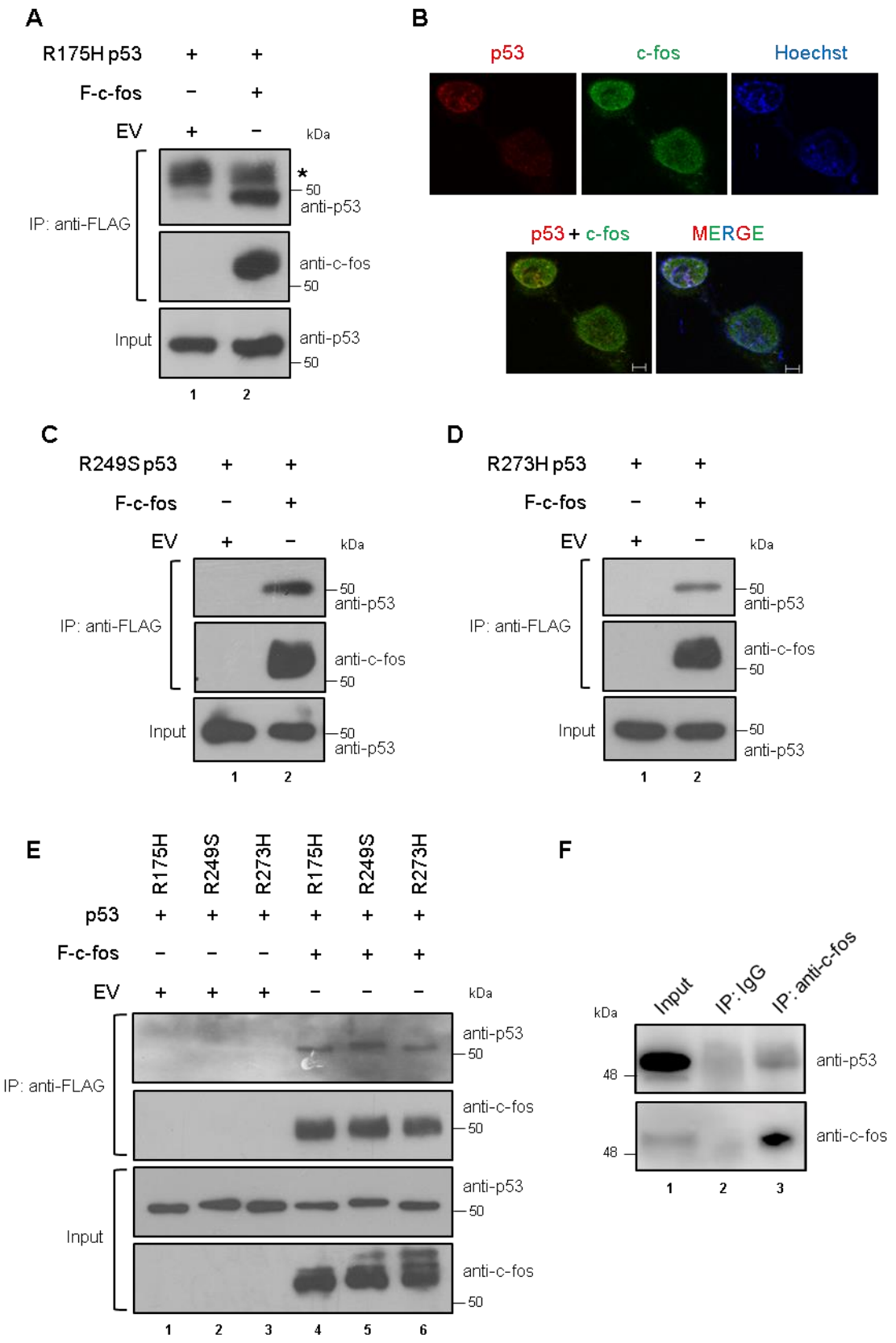



**Figure 3.2.17. Interaction of c-fos with mutant p53:** (A) Western blot analysis after anti-FLAG immunoprecipitation (IP) in H1299 p53$^{-/-}$ cells co-transfected with 5 µg of R175H p53 and 5 µg of empty vector (EV) or FLAG-c-fos as indicated, for 24 h. The upper panel shows western blot with anti-p53 and middle panel with anti-c-fos antibody in IP fraction, while the bottom panel shows

western blot with anti-p53 antibody in input lysates. Asterisk (*) indicates the antibody heavy chain band. The input was 1% of the total lysate used for immunoprecipitation. (B) Co-immunofluorescence analysis of cellular localization of R175H p53 (in red) and c-fos (in green) after transfection of H1299 cells with 700 ng each of R175H p53 and FLAG-c-fos for 24 h. Nuclei were stained with Hoechst. Yellow pixels in the merged panel (p53 + c-fos) indicate colocalization of R175H p53 and c-fos in the nuclei of the cells. Magnification is 63X and the scale bar is 5 µm. (C – D) Western blot analyses after anti-FLAG immunoprecipitation (IP) in H1299 cells co-transfected with 5 µg of (C) R249S p53 or (D) R273H p53 and 5 µg of empty vector (EV) or FLAG-c-fos as indicated, for 24 h. Upper panel shows western blot with anti-p53 and middle panel with anti-c-fos antibody in IP eluates, while the bottom panel shows western blot with anti-p53 antibody in input lysates. The input was 0.25% of the total lysates used for the respective immunoprecipitations. (E) Western blot analysis after anti-FLAG immunoprecipitation in H1299 cells co-transfected separately with R175H, R249S and R273H p53 (2 µg each) with 2 µg of empty vector (EV) or FLAG-c-fos as indicated, for 24 h. Upper panel and third panel from top show western blot with anti-p53, bottom panel and second panel from top show western blot with anti-c-fos antibody. Inputs were 1% and 3% of the lysates used for immunoprecipitation for western blot with anti-p53 and anti-c-fos, respectively. (F) Western blot analysis after immunoprecipitation of c-fos from AU565 $p53^{R175H/R175H}$ cell lysates. The upper panel shows western blot with anti-p53 and lower panel with anti-c-fos antibodies respectively. Inputs were 1% and 5% of the lysates used for immunoprecipitation, for western blotting with anti-p53 and anti-c-fos antibodies, respectively.

### *3.2.14. Mechanism of R175H-mediated activation of NPM1 expression: Synergistic effect of c-fos and R175H p53 on NPM1 expression*

We have established that c-fos and R175H p53 themselves can positively regulate NPM1 expression. To check if they had any synergistic effect on the expression of NPM1, we conducted the subsequent experiments in a system where we could conveniently induce the expressions of these factors and check their downstream effect. We generated a stable cell line in H1299 $p53^{-/-}$ background with doxycycline-inducible expression of R175H p53. We confirmed the doxycycline-induced expression of FLAG-tagged R175H p53 using immunofluorescence and western blot analyses (Chapter 2, Figure 2.8C – D)

The induction of FLAG-tagged R175H p53 expression in this cell line by doxycycline treatment (Dox) resulted in an appreciable enhancement of the NPM1 protein (Figure 3.2.18A) and mRNA (Figure 3.2.18B) levels, while it did not have a significant effect on the transcription of c-fos and c-jun (Figure 3.2.18B) even though c-fos is a reported target of wild-type p53 (Elkeles et al. 1999). This point would be discussed again in the subsequent sections.

To probe into a possible synergism among c-fos and R175H p53 on NPM1 expression, we overexpressed c-fos or the empty vector under R175H p53 induced and uninduced conditions in these cells and checked the levels of NPM1 by western blotting analysis.

Results showed that NPM1 expression was upregulated upon c-fos overexpression (Figure 3.2.18C, lane 2 versus 1) or induction of R175H p53 expression by doxycycline treatment (Dox) (Figure 3.2.18C, lane 3 versus 1), but the extent of NPM1 upregulation was noticeably greater when both these factors were overexpressed together (Figure 3.2.18C, lane 4 versus 2 and 3). This indicated that there could be a synergistic effect of c-fos and R175H p53 on NPM1 expression. It can be presumed that these proteins through their interaction, were probably getting co-recruited at the NPM1 promoter. Our subsequent experiments were carried out to test this hypothesis.

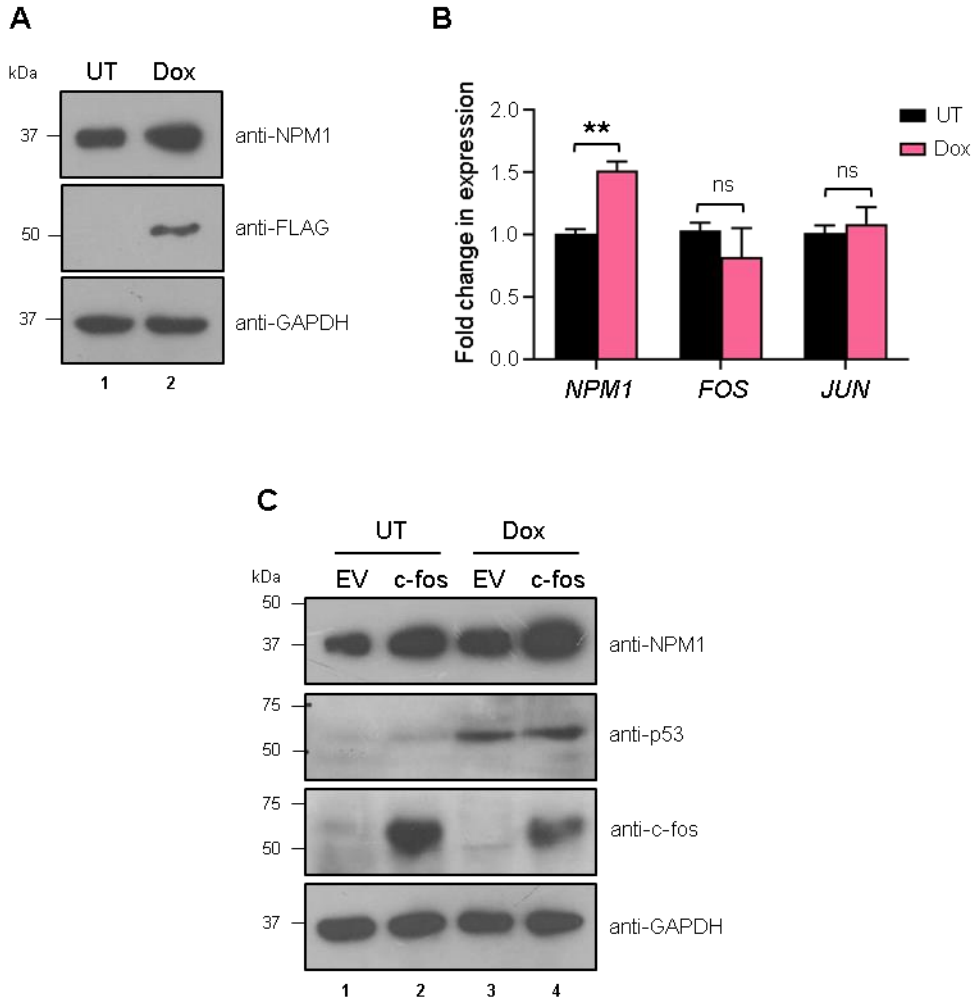



**Figure 3.2.18. Synergistic effect of c-fos and R175H p53 on NPM1 expression:** (A) Western blot analysis after induction of R175H p53 expression by doxycycline treatment (Dox) (1 µg/ml) to the H1299 $p53^{-/-}$ cells stably harboring the Tet-On plasmid for 3xFLAG-tagged R175H p53. The upper panel shows western blot with anti-NPM1, the middle panel with anti-FLAG and bottom panel with anti-GAPDH antibody respectively. (B) Bars represent fold change in mRNA levels of NPM1, c-fos, and c-jun as analyzed by RT-qPCR after induction of R175H p53 expression by doxycycline treatment (Dox). Values are mean + SEM from three independent experiments. Statistical significance was calculated using Student's *t*-test. $^{**}P < 0.01$, ns: non-significant. (C) Western blot analysis after overexpression of 1.5 µg of c-fos or empty vector (EV) with or without R175H p53

induction by doxycycline treatment (Dox) as indicated, for 24 h in the H1299 cells stably harboring the Tet-On plasmid for 3xFLAG-tagged R175H p53. Upper panel shows western blot with anti-NPM1, second panel from the top, with anti-p53, third panel from the top, with anti-c-fos and bottom panel with anti-GAPDH antibody, respectively.

### *3.2.15. Mechanism of R175H-mediated activation of NPM1 expression: Occupancy of c-fos and R175H p53 at the NPM1 promoter upon induction of R175H p53 expression*

One way to support our hypothesis of a synergistic role of c-fos and R175H p53 in regulating NPM1 expression was to check the occupancy of c-fos at the NPM1 promoter in the presence and absence of R175H p53. Towards this objective, we carried out ChIP assays with c-fos antibody under R175H p53 uninduced and induced conditions. We found that induction of R175H p53 expression enhanced the occupancy of endogenous c-fos on NPM1 promoter at the c-fos/AP-1-binding sites (Figure 3.2.19A – F). Since R175H p53 did not directly affect the expression of c-fos/AP-1 in these cells (Figure 3.2.18B), it appears that in this scenario, c-fos, from same total pool of c-fos protein as in the R175H p53 uninduced condition, gets more enriched at the NPM1 promoter in the presence of R175H p53, which is presumably brought about through the protein-protein interaction between c-fos and R175H.

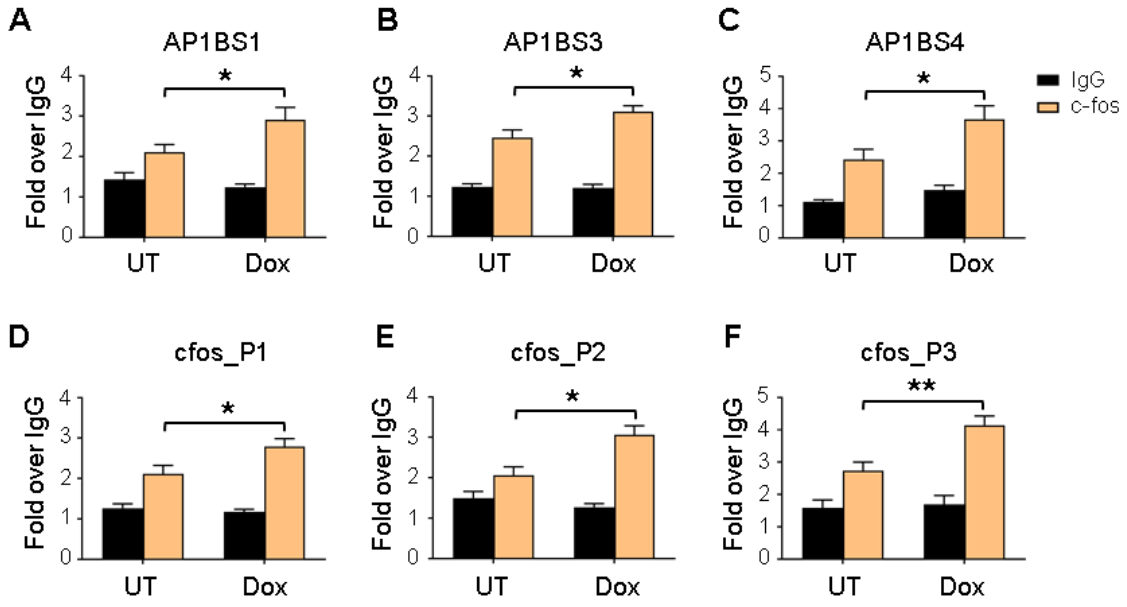



**Figure 3.2.19. Occupancy of c-fos at the NPM1 promoter upon induction of R175H p53 expression:** (A – F) c-fos occupancy as analyzed by ChIP-qPCR at various c-fos/AP-1-binding sites on the NPM1 promoter after induction of R175H p53 expression by doxycycline treatment

(Dox) (1 µg/ml) to the H1299 p53$^{-/-}$ cells stably harboring the Tet-On plasmid for 3xFLAG-tagged R175H p53. Bars represent the fold enrichment over IgG pull-down in untreated (UT) and doxycycline-treated (Dox) cells. Values are mean + SEM from two independent experiments and three technical replicates from each experiment. Statistical significance was calculated using one-way ANOVA, Tukey's multiple comparisons test. **P* < 0.05, ***P* < 0.01.

Under similar experimental conditions, interestingly we observed an enrichment of R175H itself at the AP-1-binding sites upon its doxycycline-mediated induction (Figure 3.2.20A – B). This adds further evidence towards our hypothesis that possibly there is a recruitment of R175H p53 and c-fos at the NPM1 promoter brought about by the interaction of c-fos and R175H p53, leading to the activation of NPM1 expression. As a validation of the R175H p53-mediated gene transactivation, we checked by ChIP assay and observed the enrichment of FLAG-tagged R175H p53 at known target promoters of the cell cycle genes *CCNA2*, *CDC25C*, and *CDK1* (Figure 3.2.20C – E).

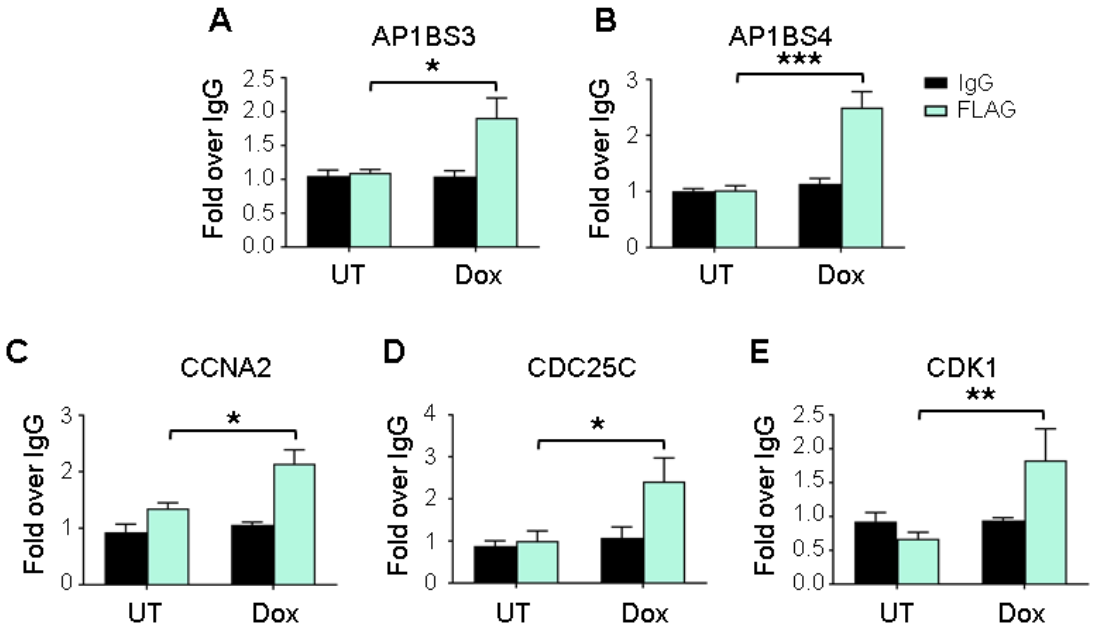



**Figure 3.2.20. Occupancy of FLAG-tagged R175H p53 at NPM1 and other target promoters upon induction of R175H p53 expression:** (A – E) FLAG-tagged R175H p53 occupancy analyzed by ChIP-qPCR after induction of R175H p53 expression by doxycycline treatment (Dox) (1 µg/ml) to the H1299 p53$^{-/-}$ cells stably harboring the Tet-On plasmid for 3xFLAG-tagged R175H p53, at (A – B) c-fos/AP-1-binding sites on the NPM1 promoter, and (C – E) on known target promoters of genes *CCNA2*, *CDC25*, and *CDK1* as indicated. Bars represent the fold enrichment over IgG pull-down in untreated (UT) and doxycycline-treated (Dox) cells. Values are mean + SEM from two independent experiments and three technical replicates from each experiment. Statistical significance was calculated using one-way ANOVA, Tukey's multiple comparisons test. **P* < 0.05, ***P* < 0.01, ****P* < 0.001.

The data presented so far proves conclusively that oncogene c-fos and mutant R175H p53 positively regulates NPM1 expression in cancer. Our observations also imply a possible co-recruitment of c-fos and R175H p53 at least at the c-fos/AP-1 binding sites on the NPM1 promoter which could synergistically enhance NPM1 expression. However, we speculate the existence of additional mechanism(s) other than the interaction of c-fos and R175H p53, which could be responsible for the specificity of the R175H p53-mediated transactivation of the *NPM1* gene. Our *in silico* transcription factor (TF) motif prediction analysis has revealed that the NPM1 promoter also contains binding motifs of other TFs which are known to interact with mutant p53 such as NF-Y, c-myc, among others. It is possible that R175H p53 could be recruited at the NPM1 promoter by one or more of such factors, which provides scope for further investigations.

### *3.2.16. Regulation of NPM1 promoter activity by transcription factor Yin Yang 1 (YY1)*

The regulation of NPM1 expression can be expected to be complicated by a complex interplay of factors including TFs which have their binding sites in the NPM1 promoter (Figure 3.2.2 and Appendix Tables A.10 and A.11). The regulations could also be sometimes context-dependent since NPM1 has been implicated in both tumor promotion as well as suppression in different conditions. Hence, an optimum level of expression of NPM1 in the cell, that is advantageous to the system at any given time, could be brought about by the interplay of transcriptional activators and repressors in regulating NPM1 expression. An interesting protein exhibiting both transcription activating and repressive properties under specific contexts is Yin Yang 1 (YY1) (Figure 3.1.1A), which is reported to have its binding site in the NPM1 promoter (Chan et al. 1997) and also regulate its expression in the context of Hepatitis C virus (HCV) infection (Mai et al. 2006) (Section 3.1, Figure 3.1.1B). YY1 is important for embryonic development and lineage differentiation, and are often dysregulated in different cancers (Gordon et al. 2006; Zhu et al. 2011; Atchison 2014). A recent report from our group has shown YY1 to be overexpressed in oral cancer which promoted the oncogenic properties in the cells and regulated the expression of genes implicated in cancer-related pathways (Behera et al. 2019). Although in this study, one of the mediating proteins in the proposed network was the coactivator associated arginine methyltransferase 1 (CARM1), we hypothesized the

presence of additional mechanisms related to YY1-mediated oncogenesis, operating in oral cancer. Our objective for further investigation of the regulation of NPM1 expression was to study the role of the dual activity transcription factor YY1, and any interplay among multiple other factors, on NPM1 expression in cancer.

We performed a Luciferase reporter assay to check if YY1 could activate NPM1 promoter activity as we have observed with the transcription factor c-fos. We observed that YY1 could significantly enhance NPM1 promoter activity after overexpression of the constructs in HEK-293 cells (Figure 3.2.21) which was in accordance with the previous report (Chan et al. 1997).

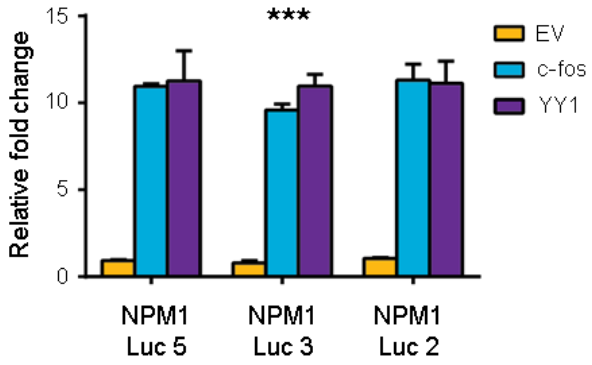



**Figure 3.2.21. Activation of NPM1 promoter activity by transcription factor YY1:** Bars represent fold change in relative Luciferase activity after transfection of HEK-293 cells with 25 ng of the NPM1 promoter constructs 'NPM1 Luc 5' (-2069/+264), 'NPM1 Luc 3' (-1587/+264) or 'NPM1 Luc 2' (-1059/+264) with or without c-fos (200 ng) or YY1 (100 ng) as indicated, for 24 h. EV: empty vector. Data were normalized to internal transfection control β-galactosidase. Values are mean + SEM from two independent experiments with two technical replicates per experiment. Statistical significance was calculated using Student's *t*-test. ****P* < 0.001.

### *3.2.17. Expression analysis of NPM1 and YY1 in human OSCC tissue samples and other cancers*

Our findings, published previously and in this study, have demonstrated that NPM1 is upregulated in human oral tumor tissue samples. So is the case with the transcription factor YY1 which has been shown to get overexpressed in oral cancer (Behera et al. 2019). We speculated that the expressions of YY1 and NPM1 might positively correlate in these oral cancer samples. To confirm that, we used the data sets for NPM1 and YY1 protein expressions in the oral cancer patient samples and performed a correlation analysis. The

results showed that the expressions of YY1 and NPM1 indeed correlated positively in oral cancer (Figure 3.2.22).

We also performed a pan-cancer expression analysis of YY1 and NPM1 transcript levels using the transcriptome data available for various human cancers from The Cancer Genome Atlas Program (TCGA Research Network: https://www.cancer.gov/tcga). The analyses showed that the mRNA expressions of YY1 and NPM1 correlate positively in 18 out of 32 types of human cancers tested (Appendix, Table A.14). This indicates that the transcriptional regulation of NPM1 expression by YY1 could be significantly relevant in the context of cancer having the potential of being used as a prognostic marker for cancer manifestation and progression. This encouraged us to investigate further in this angle of YY1-mediated regulation of NPM1 expression, focussing on oral cancer.

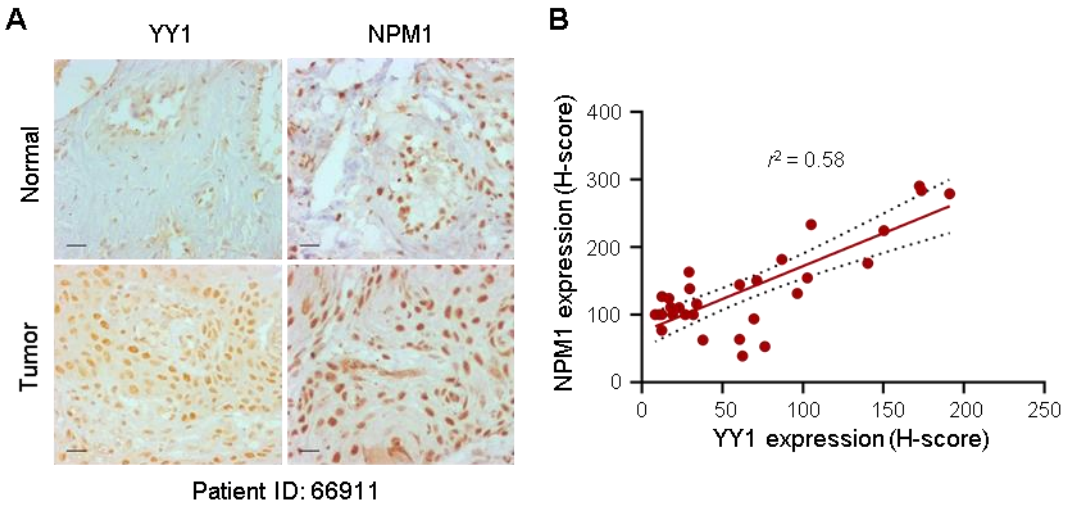
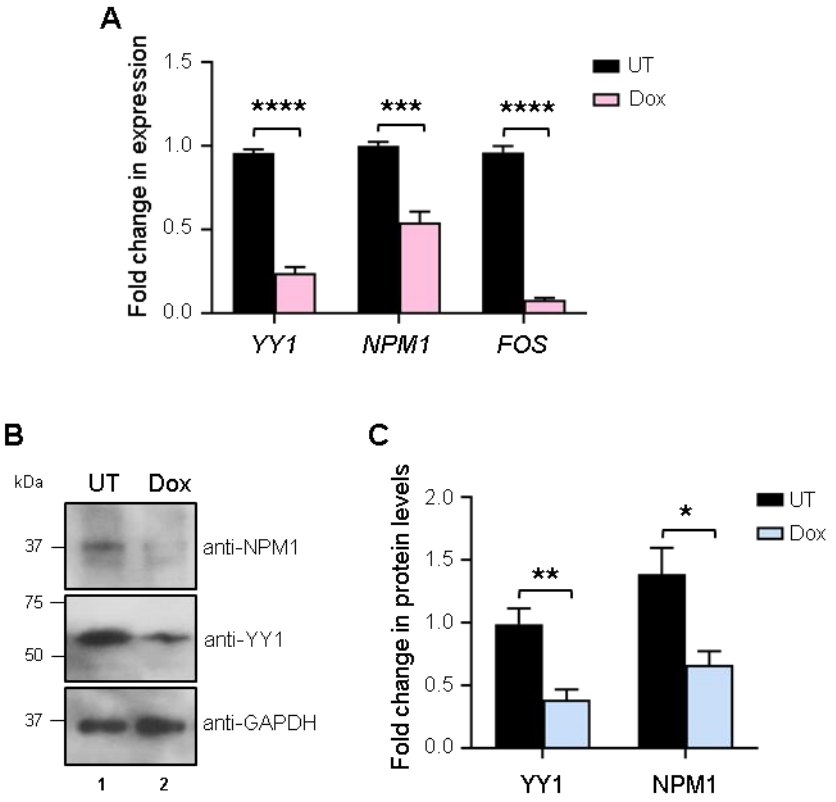



**Figure 3.2.22. Expression analysis of NPM1 and YY1 in human oral tumor tissue samples:** (A) Representative immunohistochemical images showing expressions of YY1 and NPM1 in matched normal (upper panel) and oral tumor (lower panel) tissue samples. The scale bar is 100 µm. (B) Correlation analysis of NPM1 and YY1 expressions in oral cancer patient samples (Spearman's correlation coefficient $r^2$ = 0.58, ***$P$ = 0.0007, n = 30).

### *3.2.18. Role of YY1 knockdown on endogenous NPM1 expression in oral squamous cell carcinoma (OSCC)*

To study the role of YY1 overexpression in oral cancer, our group had previously generated a stable cell line in the oral cancer background of the cell line AW8507 for doxycycline-inducible expression of the shRNA against YY1 (AW8507-shYY1) (Behera et al. 2019). We knocked down endogenous YY1 in these cells by doxycycline treatment and checked

the levels of NPM1 mRNA and protein. We found that treatment of the cells with doxycycline resulted in a significant downregulation of YY1 and consequently NPM1, both at the mRNA (Figure 3.2.23A) and protein levels (Figure 3.2.23B – C). The silencing of YY1 also resulted in significant downregulation of c-fos mRNA levels (Figure 3.2.23A), a reported target gene of YY1 itself (Natesan and Gilman 1993; Natesan and Gilman 1995; Zhou et al. 1995). As reported previously, YY1 generally represses c-fos promoter activity. However, the scenario is not always quite simple and straightforward as YY1 could also bring about the opposite effects on the c-fos promoter activity depending on the context such as association with other protein factors, specific sites for binding of the YY1 on the gene promoter, and other indirect modes of regulation. However, in this study, we found a positive correlation between the expressions of YY1 and c-fos in the oral cancer cell line as well as in the oral cancer patient tissue samples (Figure 3.2.23D – E). The implications and underlying mechanisms of YY1-mediated regulations of c-fos and NPM1 expressions in oral cancer need further exploration and would be discussed subsequently in Section 3.3.



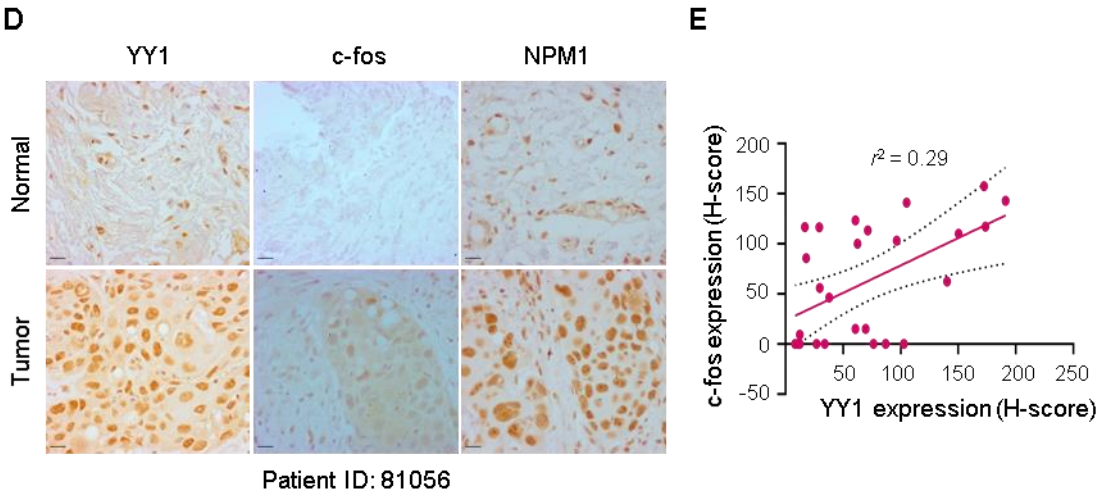



**Figure 3.2.23. Role of YY1 in the regulation of NPM1 and c-fos expressions in OSCC:** (A) Bars represent fold change in mRNA levels of YY1, NPM1 and c-fos as measured by RT-qPCR, after doxycycline treatment (Dox) (2 µg/ml every 24 h for 5 days) to AW8507-shYY1 cells. Internal normalization was done with housekeeping gene β-actin levels. (B) Western blot analysis after doxycycline treatment (Dox) to AW8507-shYY1 cells. Upper panel shows western blot with anti-NPM1, middle panel with anti-YY1, and bottom panel with anti-GAPDH antibody respectively. (C) Densitometric quantification of YY1 and NPM1 protein levels, normalized to tubulin. Bars represent fold change in protein levels with respect to the untreated (UT) control. (A, C) Values are mean + SEM from four independent experiments. Statistical significance was calculated using unpaired two-tailed Student's *t*-test. **P* < 0.05, ***P* <0.01, ****P* <0.001, *****P* < 0.0001. (D) Representative immunohistochemical images showing expressions of YY1, c-fos and NPM1 in matched normal (upper panel) and oral tumor (lower panel) tissue samples. The scale bar is 100 µm. (E) Correlation analysis of YY1 and c-fos expressions in oral cancer patient samples (Spearman's correlation coefficient $r^2$ = 0.29, ***P* = 0.0049, n = 26).

## *3.3. Discussion*

Nucleophosmin (NPM1) is a multifunctional protein exhibiting tumor-promoting as well as tumor-suppressive functions. However, it is generally found to be overexpressed in various types of cancers where its oncogenic properties are predominantly manifested. The expression of NPM1 in tumor cells is often found to increase with increasing grades of the tumors. This could be advantageous to the cancer cells since the growth-promoting functions of NPM1 are presumably enhanced due to its overexpression in the tumor cells. Due to this phenomenon, overexpressed NPM1 in the cells is often considered as a prognostic marker for cancer. However, at the molecular level, the mechanisms behind such augmented expression of NPM1 in cancer are largely unknown. Understanding the fundamental mechanisms of regulation of NPM1 expression could prove helpful in

targeting specific oncogenic pathways to diagnose cancer and reduce or inhibit its progression.

In this study, through systematic analysis, we have shown that the oncogenic transcription factor c-fos and mutant R175H p53 activate NPM1 gene transcription with a concomitant upregulation of its protein levels. The activation of NPM1 expression was brought about by the direct binding of c-fos (along with c-jun forming the AP-1 complex) at the c-fos/AP-1 binding motifs on the NPM1 promoter. Our results also imply an additional synergistic effect of c-fos and R175H p53 in the activation of NPM1 expression. We have uncovered this mechanism of c-fos- and mutant p53-mediated regulation of NPM1 expression through our biochemical and molecular biological approaches using human cancer cell lines of different origins. However, we also observed positive correlations in their expression levels in human oral cancer patient tissue samples which suggest that c-fos- and mutant-p53 mediated overexpression of NPM1 could presumably be involved in the process of oncogenesis. In fact, the overexpression of c-fos and mutant p53 has been observed in melanoma (Kroumpouzos et al. 1994), which supports this view.

R175H is one of the most common somatic mutations of p53 found in cancer (Leroy et al. 2014). This mutation adversely affects the zinc coordination sphere in the DNA binding domain (DBD) of p53 which results in its drastically reduced zinc-binding affinity and a global alteration in its three-dimensional structure (Joerger and Fersht 2007b; Joerger and Fersht 2016). Such a conformational change renders R175H p53, unlike wild-type p53, incapable of making contacts with the DNA and binding to its cognate site on its own (Joerger and Fersht 2007a; Joerger and Fersht 2008). However, through its interaction with and aid of other accessory factors such as NF-Y, c-myc, Sp1, PML, VDR, YAP1, among others, R175H p53 has been shown to bind to gene promoters bringing about the transactivation of the specific genes and a consequent increase in tumorigenic properties of the cells (Freed-Pastor and Prives 2012; Kim and Lozano 2018). While most p53 mutants are generally considered to be equivalent to each other, having similar functional outcome, the underlying mechanisms employed by the different mutants could sometimes be distinct with respect to loss-of-function, dominant-negative, and gain-of-function activities. In the case of two important hot spot mutants, namely R175H and R273H, comparative studies using mutant p53 knock-in mice revealed different tumor spectra, suggesting that the gain-of-function activities of different p53 mutants could vary (Zhou et al. 2016). Nonetheless, R175H p53 has been shown to contribute to oral cancer pathogenesis and regulate gene

expression in oral cancer (Acin et al. 2011; Grugan et al. 2013). The data presented in this report suggest that R175H transactivates the *NPM1* gene by getting recruited at the NPM1 promoter through its interaction with c-fos. Since c-fos was also found to interact with other p53 gain-of-function mutants such as R273H and R249S, we speculate the existence of additional mechanisms besides its interaction with c-fos, that might govern the specificity of R175H p53-mediated activation of NPM1 expression. As mentioned earlier, R175H is known to interact with TFs such as NF-Y, c-myc, and Sp1 (Freed-Pastor and Prives 2012), which have binding motifs on NPM1 promoter sequence as well (Figure 3.2.2 and Appendix Tables A.10 and A.11). Hence, potentially these TFs could also recruit R175H p53 to the NPM1 promoter at their cognate sites. Additionally, R175H p53, which has been shown to induce the expressions of several other TFs such as c-myc, early growth response 1 (EGR1), and nuclear factor kappa B subunit 2 (NF-κB2) (Freed-Pastor and Prives 2012), can indirectly activate *NPM1* gene transcription through these reported and potential regulators of NPM1 expression. Furthermore, R175H p53 was found to interact with co-activator CREB-binding protein (CBP)/p300 as well (Teufel et al. 2007; Freed-Pastor and Prives 2012), through its transactivation domain, which could potentially bring about the general coactivation of a multitude of genes in the cell under various contexts. These studies suggest that there could be multiple possible mechanisms of regulation of NPM1 expression by R175H p53, and further investigations could help to determine which of these mechanisms or a combination of them operate in specific cellular contexts. The events downstream to the activation of NPM1 expression could also contribute to this molecular network in cancer cells. NPM1 has been found to enhance the stability of R175H and R248W p53 mutants (Peng et al. 2001) which suggests that there could be a potential positive feedback loop in this pathway that can promote tumor progression. It should be noted that although in our studies we did not find a positive effect of the two hot-spot gain-of-function p53 mutants namely R249S and R273H on the transcriptional regulation of NPM1 expression, we do not rule out the role of other p53 mutants in this process, which have not been tested by us and might be relevant in the context of cancer.

In our study of the mechanisms of regulation of NPM1 expression, preliminary results obtained so far show that NPM1 expression is also positively regulated by the transcription factor YY1. In one of the first reports regarding the characterization of NPM1 promoter and regulation of its promoter activity, the NPM1 promoter was found to have a YY1 binding site at the −371/−344 nucleotide position upstream to its TSS (Chan et al. 1997).

In a later study, YY1 was found to regulate the expression of NPM1 in the context of Hepatitis C virus (HCV) infection (Mai et al. 2006). Mechanistically, it was shown that YY1 associated with histone deacetylase HDAC1 at the YY1 response element in the NPM1 promoter and repressed its expression. Upon infection by HCV, in the presence of the HCV core protein, the HDAC1 was replaced with the histone acetyltransferase, p300 along with NPM1 itself, which then activated its expression (Figure 3.1.1B). In this mechanism, depending on the associated factors, YY1 functions as a repressor or an activator of NPM1 expression. We also observed a positive regulation between YY1 and c-fos expression in oral cancer. YY1 is known to be a regulator of c-fos expression but depending on the context, the mechanisms of YY1-mediated regulation of c-fos expression vary. Through biochemical approaches, it was demonstrated that the manifestation of the repressive or activating functions of YY1 depended on the structure of the promoter region that is bound by YY1. The orientation of the YY1-binding site, the adjoining regions of its cognate site which could be bound by other protein factors, and the bending of the DNA by YY1 binding, could all contribute to YY1-mediated c-fos promoter regulation under specific contexts (Natesan and Gilman 1993). It was further found that YY1 could repress c-fos promoter activity indirectly, which involved the association of YY1 with the activating transcription factor/cAMP-responsive element-binding (ATF/CREB) proteins at the cyclic AMP (cAMP) response element located immediately upstream of the YY1-binding (−54 position) site on the c-fos promoter. In this way, YY1 could bind both to its own cognate site in the c-fos promoter as well as to adjacent promoter-bound ATF/CREB proteins to bring about effective repression of the c-fos promoter (Zhou et al. 1995). However, in another study by a different scientific group, it was shown that YY1 could bind to the serum response element (SRE) in the c-fos promoter and enhance the binding of the serum response factor (SRF). The transient complex of YY1 and SRF could bring about structural changes in the promoter DNA thereby potentially causing specific effects on the c-fos promoter activity (Natesan and Gilman 1995). As we have seen and discussed so far, NPM1 is also a serum-responsive factor (Feuerstein et al. 1988b) as well as regulated by c-fos. These findings suggest that the regulation of NPM1 expression in cancer by YY1 could be mediated by one or more direct or indirect mechanisms. Further in-depth investigations would be required to elucidate the precise mechanism behind this phenomenon. The interplay of c-fos/AP-1 and YY1 in the regulation of transcription has been described in the context of the miR-206 expression. It was found that AP-1 (c-fos and c-jun) enhanced miR-206 promoter activity and expression, while YY1 repressed it,

through their direct binding to the miR-206 promoter. The expression of YY1 itself was positively and negatively regulated by the proteins nuclear receptor estrogen-related receptor gamma (ERRgamma) and small heterodimer partner (SHP) respectively. Collectively in this specific scenario, SHP-mediated inhibition of ERRgamma resulted in a decrease in YY1 expression which could then de-repress AP-1 activity, ultimately causing the activation of miR-206 (Song and Wang 2009). We speculate the existence of such interconnected networks involving multiple factors operating to regulate NPM1 expression. The presence of both c-fos and YY1 binding motifs in close vicinity in the NPM1 promoter (Figure 3.3.1), the dual nature of YY1 function, and the general opposing effects brought about by YY1 and c-fos individually, support this hypothesis.

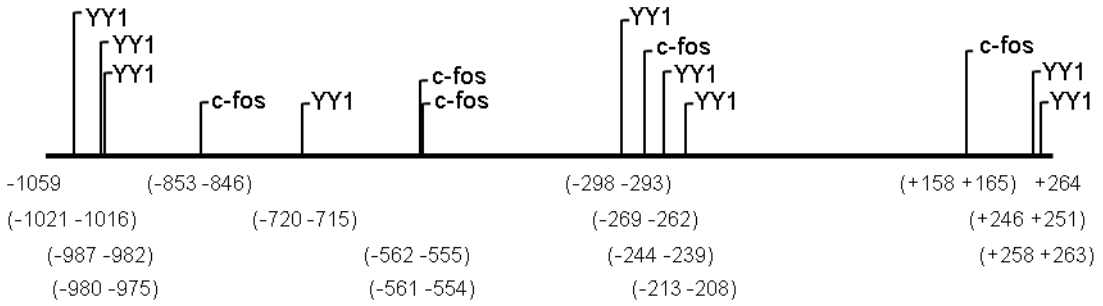



**Figure 3.3.1. YY1 and c-fos binding sites in the NPM1 promoter:** Schematic representation of the human *NPM1* gene promoter region (−1059/+264) denoting the positions of c-fos- and YY1-binding sites (at 80% TF score cut-off) identified through the Consite database. Numbers denote the positions of the nucleotides. Numbers within parentheses denote the positions and lengths of the individual sites.

In the context of YY1 and p53, the former has been reported to be a negative regulator of wild-type p53, where YY1 induces Hdm2- or Mdm2-mediated p53 polyubiquitination and degradation, a mechanism which is independent of its transactivation property (Gronroos et al. 2004; Sui et al. 2004). However, not much study has been done regarding the connection of YY1 and mutant p53, and its downstream effects on gene transactivation. One report suggests that YY1 positively regulates the transcription of mutant p53 in the breast cancer cell lines MDA-MB-231 (R280K p53) and BT-549 (R249S p53) (Wang et al. 2016). It would be interesting to study whether there is an interplay between mutant p53 and YY1 in regulating NPM1 expression, which would have important implications in cancer manifestation.

Based on the evidence we have so far through this study, we propose a model for regulation of NPM1 expression and its implications in cancer (Figure 3.3.2). According to our observations, c-fos/AP-1 and R175H p53 are co-recruited at the c-fos/AP-1 binding sites on the NPM1 promoter mediated by their mutual interaction. This results in the upregulation of NPM1 expression in cancer through a possible synergistic effect (Senapati et al. 2018). YY1 also has a positive effect on the expression levels of NPM1. However, the mechanism of YY1-mediated upregulation of NPM1 is uncertain. It could bind directly to its cognate sites in the NPM1 promoter or engage in indirect pathways involving c-fos, mutant p53, or other yet unknown factors, to regulate NPM1 expression in cancer.

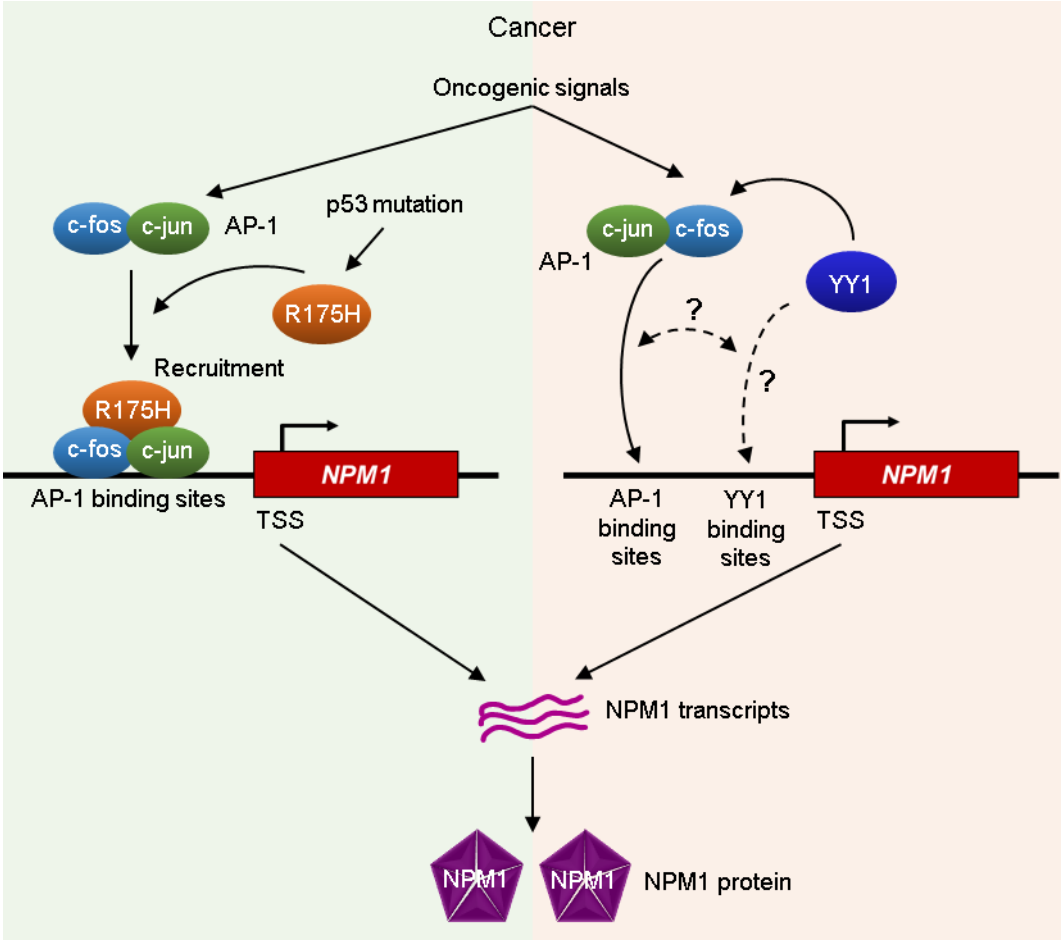



**Figure 3.3.2. Regulation of NPM1 expression in cancer: specific role of c-fos, mutant p53, and YY1:** The multi-functional histone chaperone Nucleophosmin (NPM1) is overexpressed in various types of human cancers which mostly promotes the tumorigenic properties of the cells. However, the mechanisms of NPM1 upregulation in cancers are largely unclear. (Left) The oncogenic transcription factor c-fos which forms a functional heterodimer with c-jun forming the AP-1 complex binds to NPM1 promoter leading to its upregulation in cancer. The hot-spot gain-of-function mutant of the well-known tumor suppressor p53, R175H, also activates NPM1 expression by binding to AP-1 binding sites on the NPM1 promoter via its interaction with c-fos. AP-1 and R175H thus have a synergistic effect on the enhancement of NPM1 expression (Senapati et al. 2018). (Right) YY1 can also enhance the expression NPM1 as well as c-fos in oral cancer cells

whose mechanisms are presently uncertain. YY1 could bind to its cognate sites in the NPM1 promoter which are in the vicinity of AP-1 binding sites. Therefore, there could be a possible cross-talk between the YY1- and c-fos/AP-1-mediated upregulation of NPM1 expression in cancer. Substantial evidence to support the hypothetical role of mutant p53 in this pathway is presently lacking. Hence further investigations are required in this aspect to elucidate the mechanisms precisely.

The significance of this research lies in uncovering some of the numerous molecular pathways operating in cancer which can be targeted for the diagnosis and treatment of the disease. Cancer being a very heterogeneous disease is not always diagnosed and treated using a universal set of molecular markers and therapeutics. Even within the same type of cancer, heterogeneity in molecular markers and disease outcomes is often found across different cohorts of patients. In our study using oral cancer tissue samples from the Indian cohort of patients, we have found positive correlations in the expression levels of c-fos – NPM1 and YY1 – NPM1. However, their expressions at the transcript level in global HNSC (head and neck cancer) incidence (as per the data in TCGA), was not found to correlate significantly. This suggests that the phenomenon could be specific to the patients in the Indian subcontinent, which is probably influenced by local environmental and genetic traits. Hence, understanding the molecular basis of cancer incidence and progression in a particular geographical region can potentially pave the way to develop effective region-specific diagnostics and therapeutics. This has been the motivation behind our endeavor to study the regulation of NPM1 expression with implications in cancer. The key players in this phenomenon, identified through our study, could serve as a panel of potential molecular markers and targets for the diagnosis and treatment of oral cancer in India.

# *Chapter 4: Role of NPM1 in the Regulation of RNA Polymerase II-Driven Transcription: Implications in Oral Tumorigenesis*

## *4.1. Introduction*

The role of NPM1 in the regulation of transcription has been previously described in detail in Section 1.4.1.3.6. It was earlier shown by our group that NPM1 enhances acetylation-dependent RNA Polymerase II-driven chromatin transcription *in vitro*. NPM1 itself gets acetylated by the lysine acetyltransferase (KAT) p300, which increases its affinity towards acetylated core histones and has a positive effect on its histone chaperone activity and transcription activation potential presumably through the enhanced nucleosome disassembly activity of acetylated NPM1 (AcNPM1) (Swaminathan et al. 2005). NPM1 was also found to be potent enough to even relieve the centromeric histone variant CENP-A-induced repression of chromatin transcription *in vitro* (Shandilya et al. 2014b). However, the mechanisms of NPM1-induced transcriptional activation are largely unclear. It was interesting for us to investigate the mode of regulation of RNA Polymerase II (Pol II)-mediated transcription which occurs in the nucleoplasm, by a predominantly nucleolar protein that is NPM1.

It was subsequently revealed that the acetylated pool of NPM1 (AcNPM1) localizes to the nucleoplasm along with the RNA Pol II foci, as opposed to the predominant nucleolar localization of NPM1, and regulates the transcription of genes implicated in oral cancer manifestation (Shandilya et al. 2009). Continuing on the same note, we were interested to understand the transcriptional role of NPM1/AcNPM1 at the genome-wide level, which is unknown, and deduce its gene targets in the context of oral cancer where levels of NPM1 and its acetylated form were found to be elevated (Shandilya et al. 2009). By using an in-house raised polyclonal antibody specific to NPM1 acetylated (AcNPM1) at residues K229 and K230 (Chapter 2, Section 2.5.1.1), chromatin immunoprecipitation followed by sequencing (ChIP-seq) was performed in the human cervical cancer cell line HeLa S3 (Senapati P, Ph.D. thesis, 2014). More than 82 million single-end reads from each replicate of ChIP and input samples were obtained (Table 2.4). The results of this ChIP-seq analysis would be described in the subsequent sections.

Regarding the biochemical mechanism behind NPM1-induced transcriptional activation, it was observed that the histone chaperone activity of NPM1, which depends on its oligomerization property, is important for its transcription activation potential *in vitro* (Senapati P, Ph.D. thesis, 2014). Using oligomerization-deficient mutants of NPM1, it was shown by the *in vitro* DNA supercoiling assay that they were deficient in their histone chaperone activity compared to the wild-type (WT) NPM1. These mutants of NPM1, namely L18Q and Y17T-C21F, were also deficient in their transcription activation potential compared to WT NPM1, as revealed from the *in vitro* acetylation-dependent chromatin transcription assay. NPM1 was also shown to be an inducer of autoacetylation of the transcriptional coactivator and lysine acetyltransferase protein, p300 (Arif et al. 2010), a property that enhances the enzymatic activity of p300 (Thompson et al. 2004). It was found that the oligomerization of NPM1 is important for its induction of p300 autoacetylation (Kaypee et al. 2018b) and these oligomerization-deficient mutants of NPM1, L18Q and Y17T-C21F, were unable to induce p300 autoacetylation (Senapati P, Ph.D. thesis, 2014). This could be an indirect mode of regulation by NPM1 where its inability to induce p300 autoacetylation reduces p300-mediated acetylation of nucleosomal histones at the gene promoter regions resulting in reduced activation of context-dependent gene expression.

Another possible mechanism of NPM1-mediated transcriptional regulation could be through its potential interactions with a multitude of cellular, and particularly nuclear, proteins. To test this hypothesis, a high-throughput protein-protein interaction profiling approach (Michaud and Snyder 2002) was taken (Senapati P, Ph.D. thesis, 2014). In this experiment, 3345 proteins out of 9560 proteins on the protein array, were identified to interact with NPM1 significantly. About 40% of these proteins were nuclear, out of which about 39.4% of the proteins were classified under the category of 'binding' that included the binding of 'RNA polymerase II complex', 'RNA polymerase II repressing transcription factor', 'activating transcription factor', RNA polymerase II-specific DNA-binding transcription repressors and activators, 'RNA polymerase II proximal promoter sequence-specific DNA' among others, while 12% belonged to the molecular function category of 'transcription regulator activity' according to a Gene Ontology analysis (Figure 4.1.1). Further analysis and the validation of this interactome data would be described in the subsequent sections. Collectively, these data indicated the potential role of NPM1/AcNPM1 in the recruitment, through their mutual interactions, of several transcription-related proteins at gene promoters to regulate gene transcription.

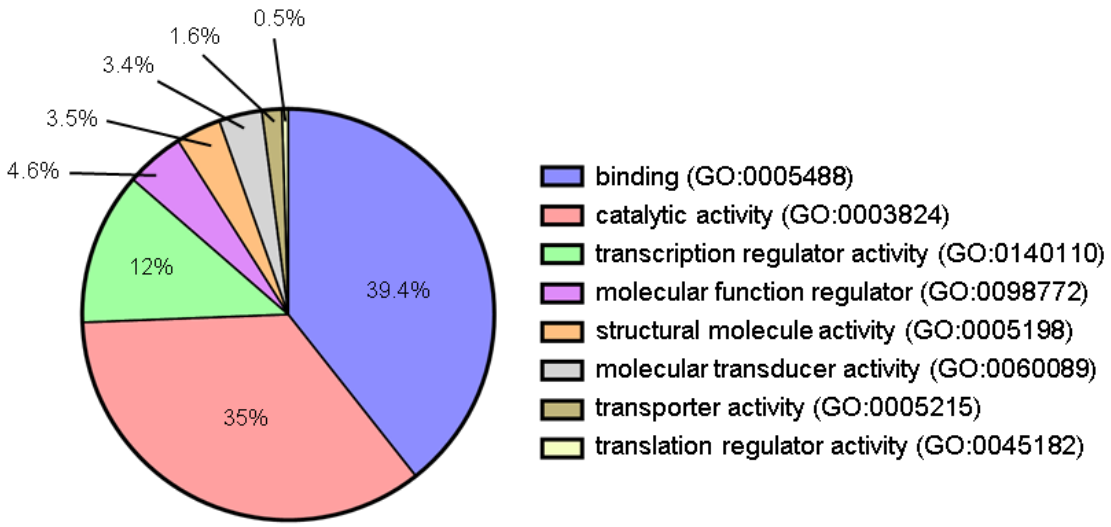



**Figure 4.1.1: NPM1 interacting proteins in the nuclear compartment of the cell:** Gene Ontology analysis of NPM1 interacting partners in the nuclear compartment categorized according to 'Molecular function'.

## *4.2. Results*

### *4.2.1. Genome-wide occupancy of acetylated NPM1 (AcNPM1)*

The AcNPM1 ChIP-seq was previously performed in HeLa S3 cells using an antibody specific towards NPM1 acetylated at sites K229 and K230 (Senapati P, Ph.D. thesis, 2014). This antibody recognized only the acetylated form of NPM1 (AcNPM1), could pull down endogenous NPM1 from the cell lysates, and exhibited nucleoplasmic staining in the cells as has been previously established for AcNPM1 (Shandilya et al. 2009). The site-specificity of this antibody was determined through peptide dot-blot and peptide competition assays (Senapati P, Ph.D. thesis, 2014).

After the ChIP-seq was performed, the sequenced reads were aligned with the hg19 assembly of the human genome using the Bowtie 2 tool (Langmead and Salzberg 2012). We called broad peaks using the Model-based Analysis of ChIP-Seq (MACS) tool, MACS2 (Zhang et al. 2008) which identified 24660 peaks conserved between the two ChIP replicates. First, we analyzed the enrichment of AcNPM1 peaks at or near the transcription start site (TSS) regions of the hg19 RefSeq genes. We found about 40.65% (10056) of the AcNPM1 peaks to be enriched within 1 kb of TSS and about 60.1% (14867) of all the AcNPM1 peaks to be enriched within 10 kb of the TSS in comparison to a random selection of peaks of the same size on the human genome where only about 16.48% of the peaks

were enriched within 10 kb of the TSS (Figure 4.2.1A). This indicated that AcNPM1 was localized at the TSS or promoter-proximal regions. A visual representation of the enrichment of AcNPM1 at the TSS of genes is shown in a genome browser snapshot of a section of chromosome 19 (Figure 4.2.1B). This fact was also clear from the heatmap of ChIP-seq tag density at TSS ± 2 kb of all the RefSeq genes (Figure 4.2.1C).

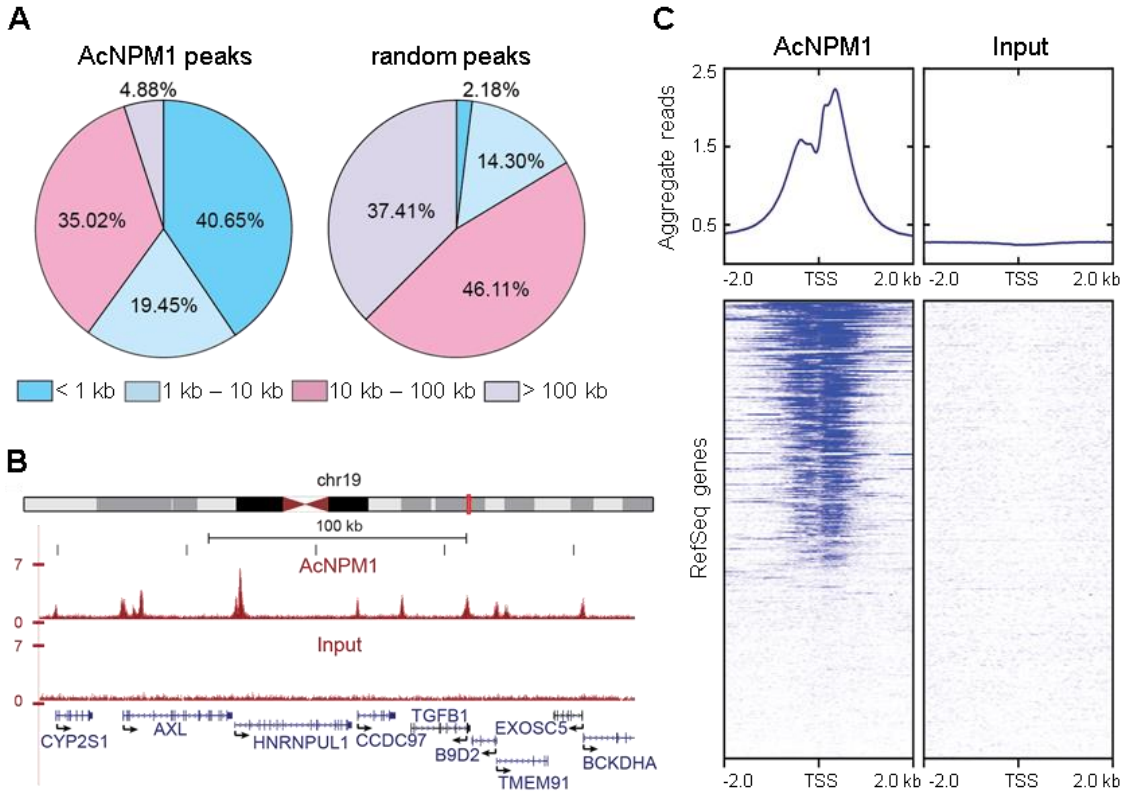



**Figure 4.2.1: Genomic localization of AcNPM1 peaks:** (A) Genomic distribution of AcNPM1 ChIP-seq peaks plotted with respect to their distance from the TSS of RefSeq genes (left pie chart) in comparison to a random distribution of peaks having approximately similar size (right pie chart). (B) UCSC genome browser snapshot showing the enrichment of normalized reads for AcNPM1 and Input signals at the TSS of genes as indicated on chromosome 19 (http://genome.ucsc.edu/). (C) Heatmaps showing the normalized reads for AcNPM1 and the Input enriched at TSS ± 2 kb regions of all RefSeq genes. The aggregate profiles of AcNPM1 and the Input read density are shown above the heatmaps.

### *4.2.2. Comparison of the genome-wide profile of AcNPM1 with those of epigenetic signatures associated with transcription*

To test if there is an association of the AcNPM1 genomic localization with epigenetic signatures of transcription activation, we compared the AcNPM1 profile with the genome-wide ChIP-seq profiles of histone modification marks in HeLa S3 cells which were

available from the ENCODE consortium (2012). The comparative analyses were done for the active histone modifications such as H3K9ac, H3K27ac, H3K4me1, H3K4me2, H3K4me3, H3K36me3, H3K79me2, and H4K20me1, and repressive histone modifications such as H3K9me3 and H3K27me3. The Jaccard index, which is a measure of similarity for two sets of data, was deduced for pairwise comparisons of AcNPM1 with these active and repressive histone modification marks, and these values were plotted in the form of a heat map with hierarchical clustering based on the similarity. It was found that there was a strong overlap of AcNPM1 with the histone modification marks H3K27ac, H3K9ac, and H3K4me3 (Figure 4.2.2) which are known to be enriched at the gene promoters. There was a moderate overlap of AcNPM1 with the histone modification marks H3K4me1 and H3K4me2 (Figure 4.2.2) which are also associated with gene promoters. However, there was no overlap of the AcNPM1 profile with active histone modification marks observed on the gene bodies namely H3K79me2, H3K36me3, and H4K20me1, as well as the repressive marks H3K9me3 and H3K27me3 (Figure 4.2.2). These results further prove that AcNPM1 is enriched at the gene promoters along with the promoter-specific histone modification marks.

To test if the occupancy of AcNPM1 at the gene promoters is associated with gene transcription, the RefSeq mRNAs were categorized into four quartiles based on their expression levels as known in the case of HeLa S3 cells. Quartile 1 (Q1) was assigned for the least expressed while quartile 4 (Q4) was assigned for the most expressed genes. It was found that the enrichment of AcNPM1 was higher at the promoters of the transcripts belonging to Q4 which contained the highly expressed transcripts. Likewise, the promoters of the transcripts showing lower levels of expression (Q3 and Q2) showed lower occupancy of AcNPM1. Going with the same logic, it was observed that there was no enrichment of AcNPM1 at the promoters of the genes with no detectable expression (Q1) (Figure 4.2.3A). When compared with the other promoter enriched marks, namely H3K27ac, H3K9ac, and H3K4me3, we found similar expected trends of enrichments of these histone modifications at the promoters of these highly expressed transcripts (Figure 4.2.3A). This correlation between the expression of the genes and the occupancy of AcNPM1, H3K27ac, H3K9ac, or H3K4me3 is also evident from the heat maps of the respective ChIP-seq tags on promoters (TSS ± 2 kb) of genes arranged by their expression levels in HeLa S3 cells (Figure 4.2.3B), as well as from the genome browser screenshot of a section of chromosome 19 depicting the enrichment of AcNPM1 with H3K27ac, H3K9ac, and H3K4me3 peaks at

promoters of expressed genes (Figure 4.2.3C). Collectively, it can be inferred from these results that the genomic localization of AcNPM1 is related to active transcription at gene promoters.

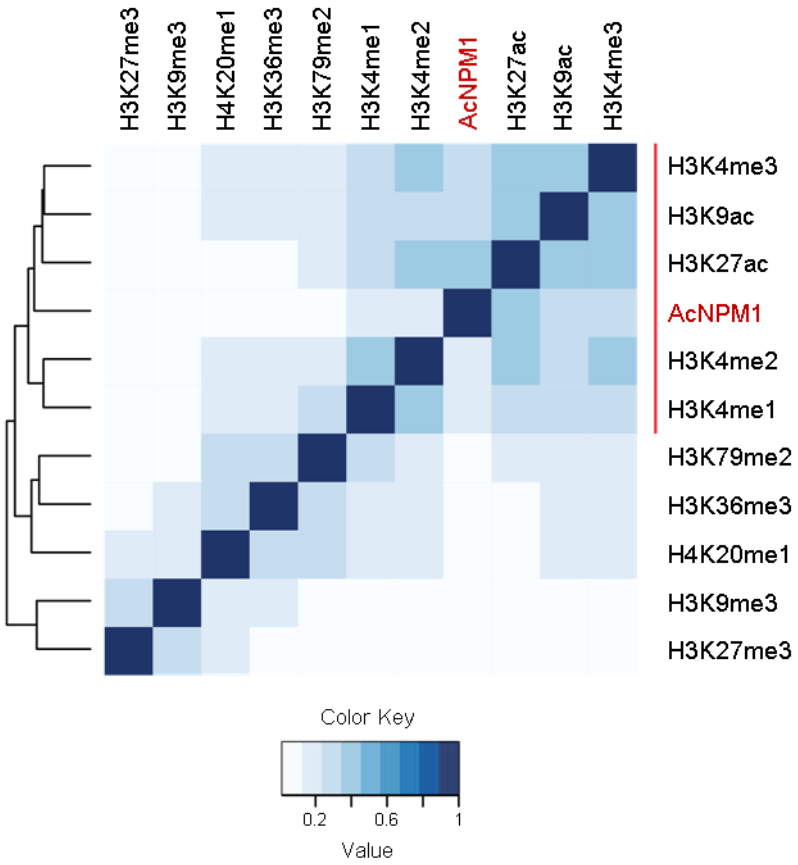



**Figure 4.2.2: Overlap of AcNPM1 ChIP-seq profile with the profiles of other transcription-related epigenetic signatures:** Clustered heatmap depicting the Jaccard similarity indexes signifying overlap between AcNPM1 ChIP-seq peaks and the transcription-associated histone modification marks as indicated. The color scale for the Jaccard indexes is shown below.

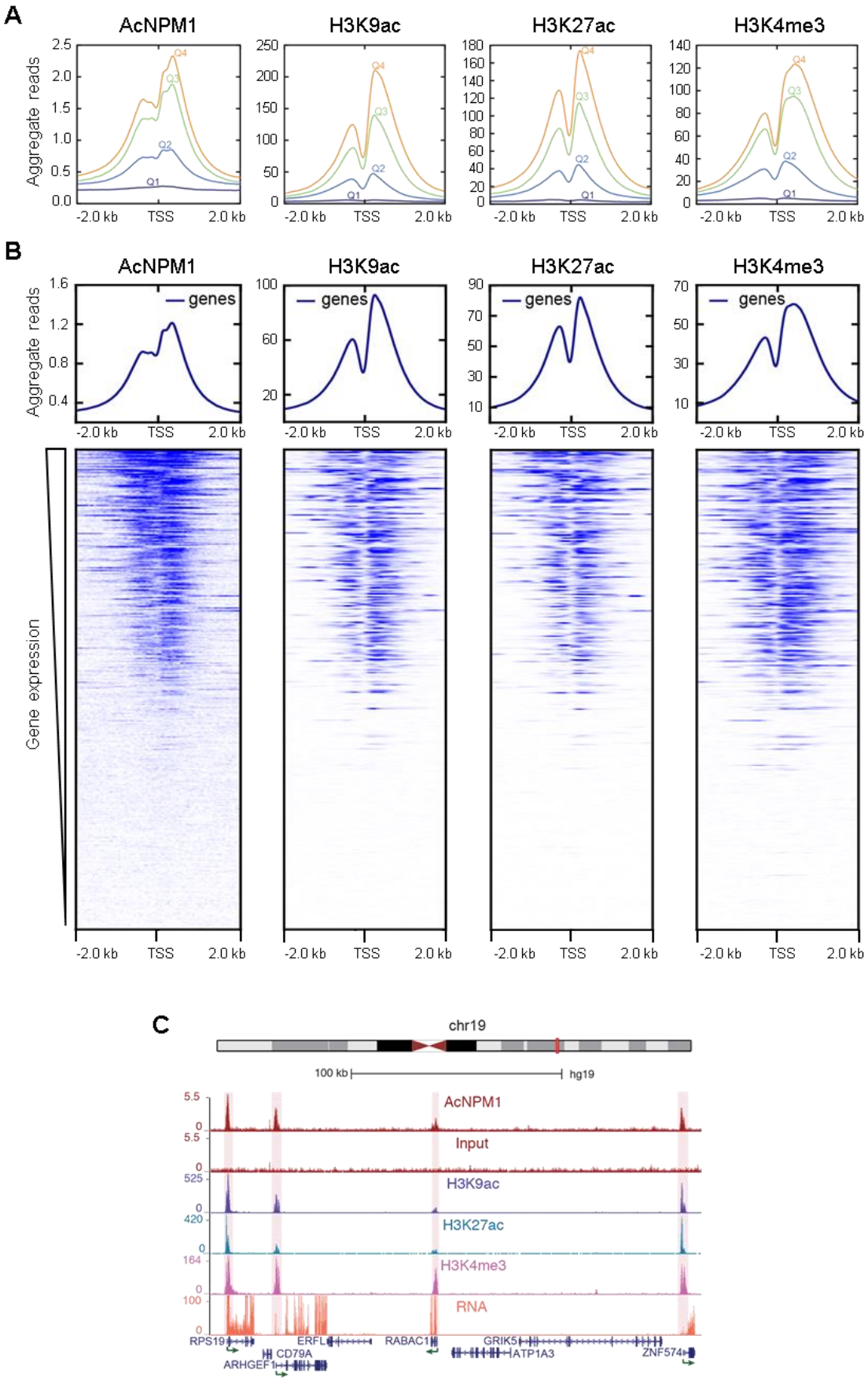

A
AcNPM1
H3K9ac
H3K27ac
H3K4me3
Aggregate reads
Q4
Q3
Q2
Q1
-2.0 kb
TSS
2.0 kb
B
AcNPM1
H3K9ac
H3K27ac
H3K4me3
Aggregate reads
genes
Gene expression
-2.0 kb
TSS
2.0 kb
C
chr19
100 kb
hg19
AcNPM1
Input
H3K9ac
H3K27ac
H3K4me3
RNA
RPS19
ERFL
RABAC1
GRIK5
ATP1A3
ZNF574
CD79A
ARHGEF1

**Figure 4.2.3: Association of AcNPM1 with active transcription:** (A) Normalized read counts for AcNPM1, H3K9ac, H3K27ac, and H3K4me3 which are plotted in 10 bp bins in a TSS ± 2 kb window for all RefSeq transcripts. Q1 (dark blue) and Q4 (orange) represent the quartiles of genes having the lowest and highest expression levels respectively. (B) Heat maps of normalized reads for AcNPM1, H3K9ac, H3K27ac, and H3K4me3, at TSS ± 2 kb regions of all RefSeq genes ordered by decreasing levels of expressions in HeLa S3 cells. The aggregate profiles of the read densities are shown above the respective heat maps. (C) UCSC genome browser snapshot of a section of chromosome 19 depicting the enrichment of normalized reads for AcNPM1, Input, histone modification marks H3K9ac, H3K27ac, and H3K4me3 at the TSS of the indicated genes as well as the mRNA track showing their respective expression levels.

### *4.2.3. Occupancy of AcNPM1 at the enhancer regions of the genome*

The AcNPM1 genome occupancy was further characterized to investigate its association with other regulatory elements of the genome. For this, the AcNPM1 peaks were compared with the combined ChromHMM and Segway segmentation (Hoffman et al. 2013) of the HeLa S3 genome retrieved from the UCSC genome browser (Kent et al. 2002). From this analysis, it was found that 41.6% of the AcNPM1 peaks were enriched at the TSS regions and 25.7% of the AcNPM1 peaks overlapped with annotated enhancer regions in the HeLa S3 genome (Figure 4.2.4A). This analysis also revealed the overlap of AcNPM1 with CTCF-bound regions in accordance with a previous report showing co-occupancy of NPM1 and CTCF at insulator regions (Yusufzai et al. 2004) (Figure 4.2.4A). Among the TSS and enhancers classified by the combined segmentation (ChromHMM + Segway) in HeLa S3 cells, about 90% of the known TSS regions were found to be occupied by AcNPM1, while the AcNPM1 enrichment observed at the annotated enhancers was about 54% (Figure 4.2.4B). The enrichment of AcNPM1 at the TSS and enhancer regions is also evident from the genome browser view of a region of chromosome 11 (Figure 4.2.4C).

Enhancer regions of the genome are characterized by the presence of DNase I hypersensitive sites (DHSs) as well as enrichment of the H3K27ac histone modification mark. The transcriptional coactivator protein p300 and RNA Pol II are also known to occupy enhancer regions in the genome. To find out the association of AcNPM1 with such enhancer markers, the AcNPM1 occupancy was determined at the DHSs which was found to be significantly higher at DHS regions in comparison to the non-DHS regions (Figure 4.2.4D). Similarly, the AcNPM1 ChIP signal showed greater overlap with the regions having higher levels of H3K27ac histone mark, p300, and RNA Pol II (Figure 4.2.4D). Collectively, these results indicated that AcNPM1 is associated with active transcription occurring at regulatory elements such as promoters and enhancers.

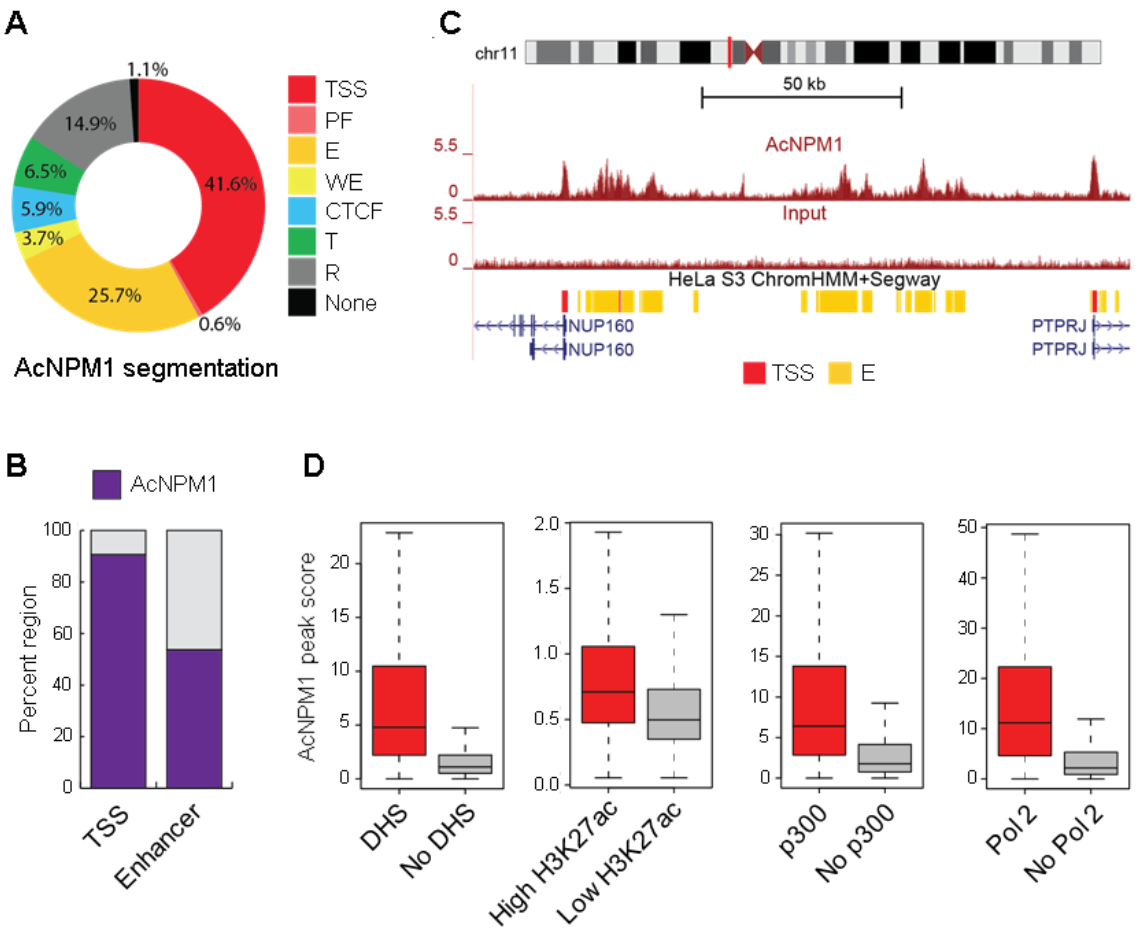


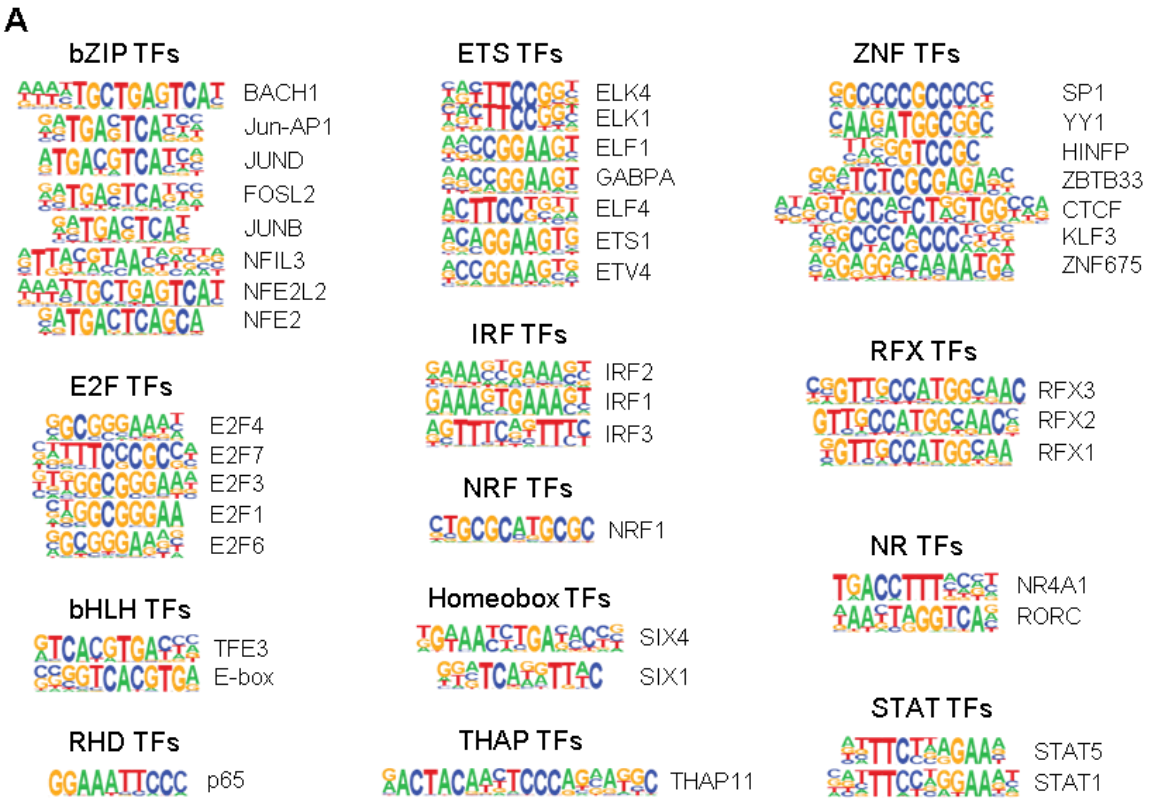

**Figure 4.2.4: Association of AcNPM1 with transcription regulatory elements such as promoters and enhancers:** (A) Pie chart showing the percent number of AcNPM1 peaks overlapping with the combined segmentation (ChromHMM + Segway) for HeLa S3 genome retrieved from the UCSC genome browser. TSS: predicted promoter region including TSS, PF: predicted promoter flanking region, E: enhancer, WE: predicted weak enhancer or open chromatin cis-regulatory element, CTCF: CTCF enriched element, T: predicted transcribed region, R: predicted repressed or low activity region, None: unclassified. (B) Percent number of TSS and enhancer regions identified by combined segmentation (ChromHMM + Segway) for the HeLa S3 genome, overlapping with AcNPM1 peaks. (C) UCSC genome browser snapshot of a section of chromosome 11 depicting the enrichment of AcNPM1 at TSS and enhancer regions as defined by the combined segmentation (ChromHMM + Segway) for HeLa S3 genome. TSS: predicted promoter region including TSS, E: enhancer. (D) Boxplots denoting the AcNPM1 read density on the AcNPM1 peaks that overlap with DNase I hypersensitive sites (DHS) or non-DHS regions, high or low enrichment of H3K27ac, p300-bound or p300-unbound regions, and RNA Pol II-occupied or Pol II-unoccupied regions, as indicated.

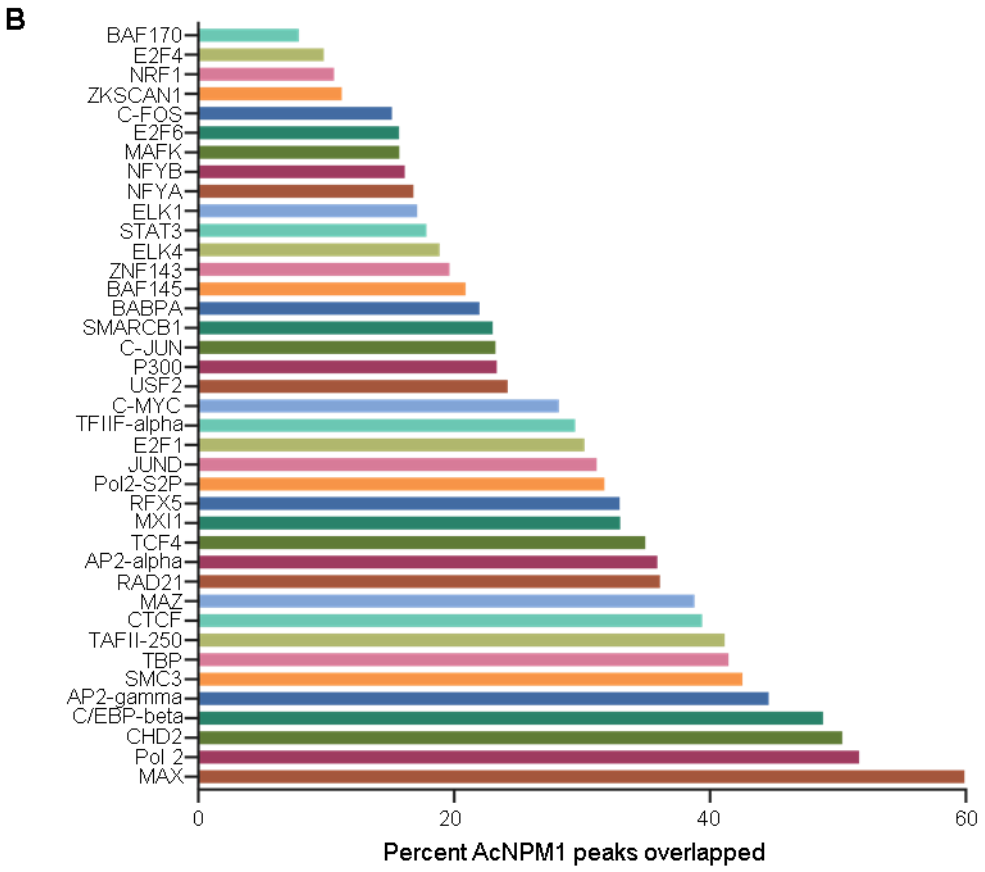
### *4.2.4. Features of the genome-wide AcNPM1peaks*

We further characterized the AcNPM1-enriched sequences to gain insights into the possible mechanisms of AcNPM1 occupancy at the active regulatory elements of the genome. We performed analyses and identified binding motifs of various transcription factors (TFs)

enriched at the AcNPM1 peaks. These TFs belonged to different families such as bZIP (basic leucine zipper), ETS (E26 transformation-specific/E-twenty-six), Znf (zinc finger), E2F (E2 transcription factor), IRF (interferon regulatory factors), RFX (regulatory factor X), Homeobox, bHLH (basic helix-loop-helix), STAT (signal transducers and activators of transcription), NR (nuclear receptor), NRF (nuclear respiratory factor), THAP (THAP domain containing), and RHD (Rel homology domain) (Figure 4.2.5A). This indicates that AcNPM1 might be co-recruited or associated with these TFs at their binding sites on their target gene promoters. We also compared the ChIP enrichment profiles of various transcription factors and chromatin-associated proteins in HeLa S3 cells available from the ENCODE database (2012) with that of AcNPM1 to validate the possibility of co-binding of such proteins with AcNPM1 in the genome. Among all the proteins whose data were available for analyses, we considered those which overlapped with at least 10% of the AcNPM1 peaks for the comparisons. We found that the transcription factor MAX (MYC-associated factor X) had the highest overlap (about 60%) at the AcNPM1-enriched regions, followed by RNA Pol II, CHD2 (chromodomain helicase DNA binding protein 2), and C/EBP-beta or CEBPB (CCAAT enhancer binding protein beta) (about 50% overlap) (Figure 4.2.5B). To represent the co-occupancy of these proteins with AcNPM1 in an unbiased manner, we plotted the correlation values for the overlap of the different factors with AcNPM1 in the form of a heat map with hierarchical clustering (Figure 4.2.5C). From this plot, we observed several clusters of co-bound factors. The largest among these clusters was the one containing the TFs MYC, MAX, MAZ (MYC associated zinc finger protein), MXI1 (MAX interactor 1, dimerization protein), along with TAF1 (TATA-box binding protein associated factor 1), CHD2, and RNA Pol II (Figure 4.2.5C). The second-largest cluster observed was the one containing the TFs c-fos, c-jun, JUND (JunD proto-oncogene, AP-1 transcription factor subunit), p300, and STAT3 (signal transducer and activator of transcription 3). The other significant clusters of TFs overlapping with the AcNPM1 peaks are as follows: the ETS Like transcription factors ELK1, ELK4, and GABP (GA-binding protein); CTCF (CCCTC-binding factor), RAD21 (RAD21 cohesin complex component), and SMC3 (structural maintenance of chromosomes 3); the E2 transcription factors E2F4, E2F6, and E2F1; the Activator Protein 2 members AP-2 alpha and AP-2 gamma; and NFYA (nuclear transcription factor Y subunit alpha) and NFYB (nuclear transcription factor Y subunit beta). The association of AcNPM1 with CTCF in the genome is in accordance with the ~6% overlap of AcNPM1 enriched regions with those of CTCF as we have observed previously (Figure 4.2.4A) and suggests that AcNPM1 might be associated

with the proteins SMC3 and RAD21 along with CTCF at the boundary elements (Parelho et al. 2008; Wendt et al. 2008). Also, the binding motifs of some of the factors which showed stronger overlap with the AcNPM1 peaks based on the clustering, such as E2F family, AP-1, JUN, STAT family, ETS family, CTCF among others (Figure 4.2.5C), were found to be enriched in the AcNPM1 peaks (Figure 4.2.5A), collectively indicating that there could be a recruitment of AcNPM1 by these TFs at their cognate sites, resulting in the respective gene expression regulation.




B
BAF170
E2F4
NRF1
ZKSCAN1
C-FOS
E2F6
MAFK
NFYB
NFYA
ELK1
STAT3
ELK4
ZNF143
BAF155
BABPA
SMARCB1
C-JUN
P300
USF2
C-MYC
TFIIF-alpha
E2F1
JUND
Pol2-S2P
RFX5
MXI1
TCF4
AP2-alpha
RAD21
MAZ
CTCF
TAFII-250
TBP
SMC3
AP2-gamma
C/EBP-beta
CHD2
Pol 2
MAX
0
20
40
60
Percent AcNPM1 peaks overlapped

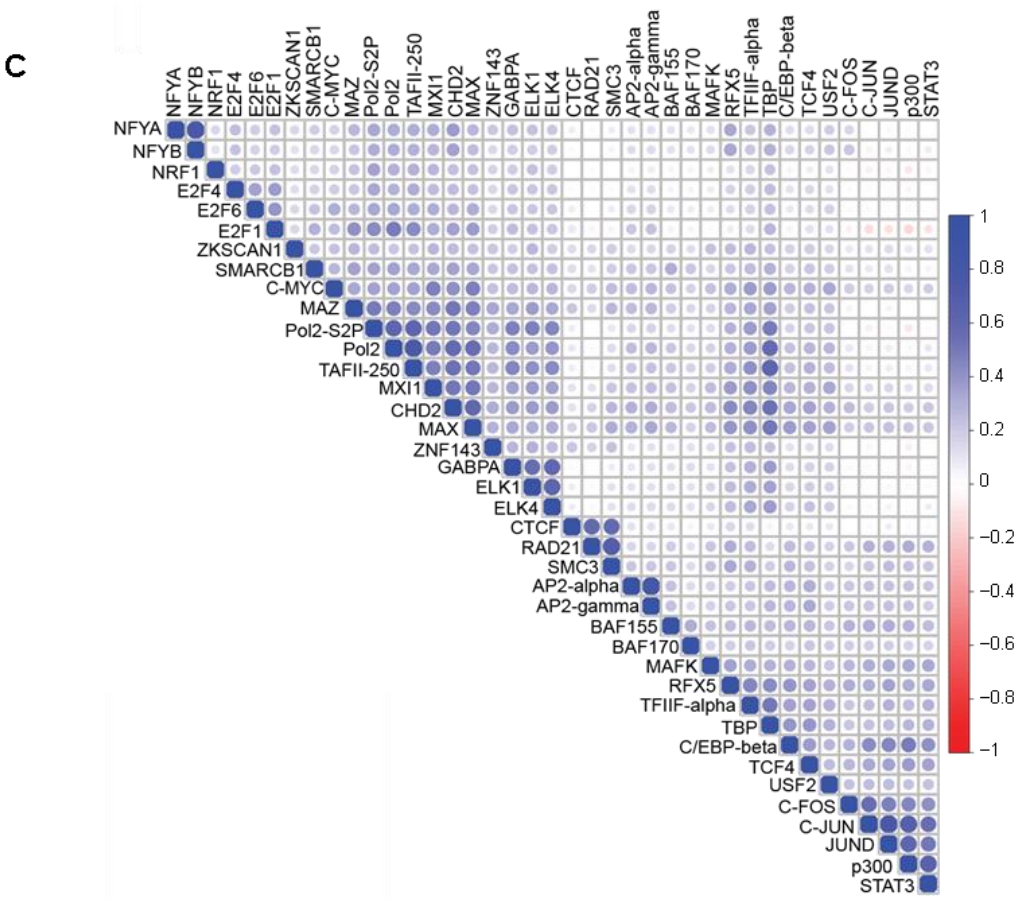

C
NFYA
NFYB
NRF1
E2F4
E2F6
E2F1
ZKSCAN1
SMARCB1
C-MYC
MAZ
Pol2-S2P
Pol2
TAFII-250
MXI1
CHD2
MAX
ZNF143
GABPA
ELK1
ELK4
CTCF
RAD21
SMC3
AP2-alpha
AP2-gamma
BAF155
BAF170
MAFK
RFX5
TFIIF-alpha
TBP
C/EBP-beta
TCF4
USF2
C-FOS
C-JUN
JUND
p300
STAT3
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1

**Figure 4.2.5. Features of the genome-wide AcNPM1 peaks and overlap analyses of binding profiles of AcNPM1 and other transcription-related proteins:** (A) Transcription factors (TFs) classified by their TF families whose binding motifs were found to be enriched in the AcNPM1 peaks. $P < 0.05$. (B) Percent number of AcNPM1 peaks that overlapped with the peaks of the indicated transcription factors and chromatin-associated proteins, as obtained from the ENCODE data. (C) Heat map representing the Pearson correlation plot for the co-occupancy of various transcription factors and chromatin-associated proteins with AcNPM1. Hierarchical clustering was performed to identify the factors that co-occupied AcNPM1 peaks. The color scale depicts the Pearson correlation coefficient values.

### *4.2.5. Analysis and validation of NPM1 interactome*

The comparative analyses of the genome-wide occupancy profiles of AcNPM1 and other transcription-related proteins suggested that AcNPM1 co-occupies the transcriptional regulatory elements for RNA Pol II, TFs and other chromatin-associated proteins. In support of this possibility, we tested if NPM1 could directly interact with these factors. An immunoprecipitation (IP)-mass spectrometry assay from whole-cell lysates was previously performed to identify the different NPM1-interacting partners (Senapati P, Ph.D. thesis, 2014). However, since the majority of the NPM1 protein is present in the nucleoli of the cells, through this approach, it was difficult to detect the proteins interacting with the relatively smaller pool of AcNPM1 in the nucleoplasm. Also, the mass spectrometry approach potentially identified numerous proteins that might be interacting with NPM1 indirectly within complexes. For this reason, a subsequent high-throughput protein-protein interaction profiling approach was undertaken (Michaud and Snyder 2002) using an array consisting of 9560 human proteins (Senapati P, Ph.D. thesis, 2014). In this assay, the native proteins spotted on the array was incubated with or without (negative control) recombinant NPM1, followed by detection and quantification of the interaction by using a specific primary antibody (anti-NPM1) and a fluorophore-conjugated secondary antibody. As mentioned in Section 4.1, by this approach, 3345 proteins out of 9560 proteins spotted on the array, were identified to interact with NPM1 significantly. A Gene Ontology analysis of the interacting proteins identified about 40% being nuclear proteins, out of which about 39.4% were classified under the category of 'binding' that included the binding of 'RNA polymerase II complex', 'RNA polymerase II repressing transcription factor', 'activating transcription factor', RNA polymerase II-specific DNA-binding transcription repressors and activators, 'RNA polymerase II proximal promoter sequence-specific DNA' among others, while 12% belonged to the 'molecular function' category of 'transcription regulator activity' (Figure 4.1.1). Several interacting proteins were identified to be novel (Figure

4.2.6A), belonging to different classes such as core histones and histone variants, ribosome complex, cyclin-dependent protein kinase holoenzyme complex, DNA polymerases, histone acetyltransferase complexes, RNA Pol II holoenzyme, transcription elongation machinery, transcription-coupled NER (nucleotide excision repair) machinery, proteins associated with condensed chromosome centromeric region and so on (data not shown). The validity of the analysis was assessed by the presence of known interacting partners of NPM1 which were also identified in this approach. Among the 623 known interacting proteins of NPM1 as retrieved from the BioGRID database (Stark et al. 2006), 279 were present on the array, out of which only 143 proteins passed the stringent threshold parameters for detecting a positive signal and indicating high confidence in the identified NPM1 interacting partners (Senapati P, Ph.D. thesis, 2014). We selected four novel candidate proteins from the analyzed dataset, namely transcription factor AP-4, RNA polymerase II subunit K (POLR2K), positive coactivator 4 (PC4), and snail family transcriptional repressor 1 (SNAI1), and experimentally validated their interactions with NPM1 and AcNPM1 *in vitro*. The *in vitro* interaction assays were performed using $His_6$-tagged interacting partner proteins and FLAG-tagged NPM1 protein purified after expression in *E. coli* cells. The FLAG-NPM1 recombinant protein was pre-acetylated using p300 or mock acetylated (without p300) so as to test if these NPM1-interacting proteins could interact with AcNPM1 as well (Figure 4.2.6B). The mass acetylation of NPM1 was confirmed by western blotting with anti-AcNPM1 antibody (Chapter 2, Section 2.11.2, Figure 2.30). The $His_6$-tagged proteins were pulled down using Ni-NTA resin and interaction of NPM1 with the $His_6$-tagged proteins was detected by western blotting using anti-NPM1 antibody. The results of the individual experiments showed that AP-4, POLR2K, and PC4 interacted with both mock acetylated and acetylated forms of NPM1 (Figure 4.2.6C – E) whereas SNAI1 was found to interact with only mock acetylated NPM1 and not AcNPM1 (Figure 4.2.6F) under the experimental conditions tested. This showed that the NPM1 interacting proteins identified through the profiling array could be validated in our *in vitro* interaction assays.

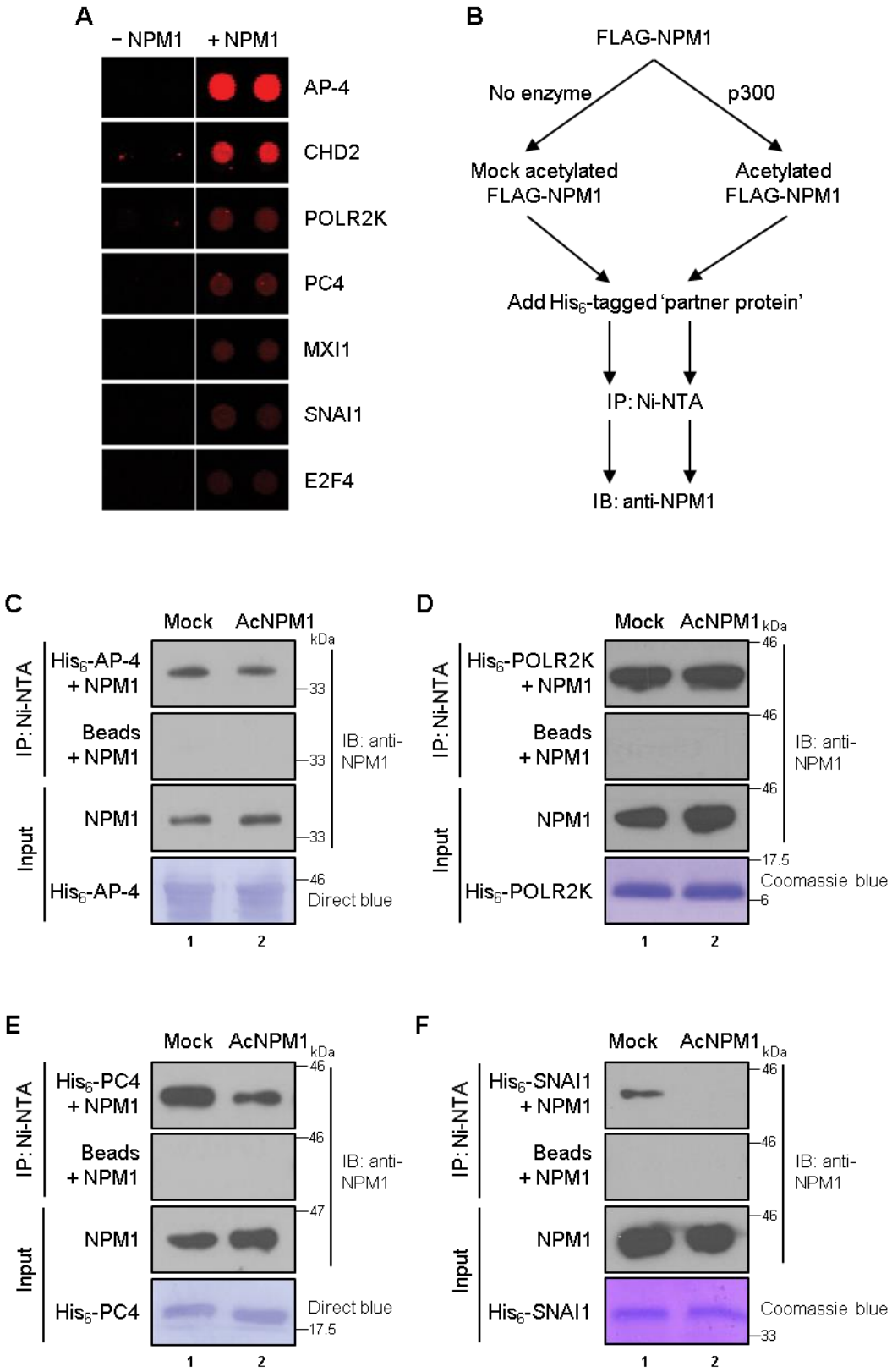



**Figure 4.2.6: Interactions of NPM1 with transcription-related proteins *in vitro*:** (A) Representative images of the protoarray probed with (right) or without NPM1 protein (left) showing spots (red) of selected interacting proteins as indicated. (B) Schematic representation of the experimental steps followed in the *in vitro* interaction assays for the validation of interactions identified through the protoarray. (C – F) *In vitro* Ni-NTA pull-down (IP) of FLAG-tagged recombinant

mock acetylated NPM1 (Mock) or acetylated NPM1 (AcNPM1) using (C) $His_6$-AP-4 (D) $His_6$-POLR2K (E) $His_6$-PC4, and (F) $His_6$-SNAI1 proteins. Western blots (IB) were performed with the NPM1 antibody after the pull-downs. Equal loading of the $His_6$-tagged partner proteins is shown using Direct blue staining of the membrane or Coomassie blue staining of the gel as indicated. Input is 20% of NPM1.

As mentioned previously, through Gene Ontology analysis of the NPM1-interacting proteins, we identified important biological processes associated with these proteins. Several such novel interacting partners were identified as subunits of different histone acetyltransferase (HAT) complexes (Figure 4.2.7A), RNA Pol II (POLR2C, POLR2E, POLR2I, POLR2K, POLR2M), GTFs (GTF2A2, GTF2E1, GTF2E2, GTF2F1), TFIID associated factors (TAF5L, TAF6, TAF7L, TAF8, TAF9), and Mediator complex (MED4, MED8, MED11, MED12, MED17, MED19, MED22, MED31) (Figure 4.2.7B) among others. Collectively, these results indicate that NPM1 possibly interacts with the transcription machinery, HAT complexes, remodelers, and other transcription-related proteins and could potentially function as a coactivator during transcription.

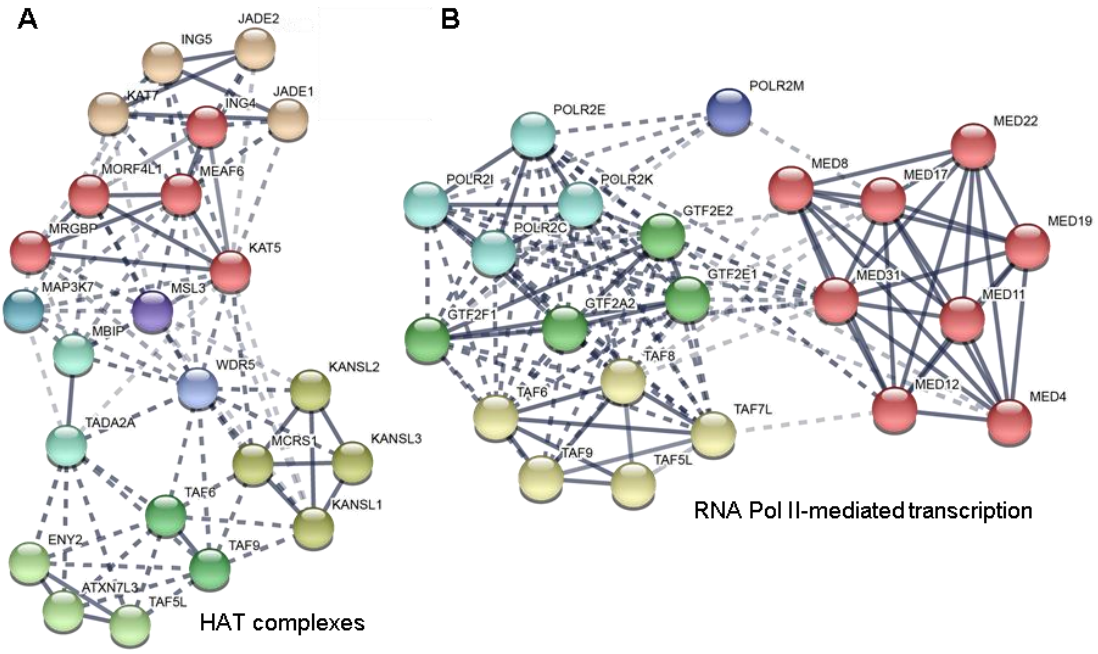



**Figure 4.2.7. Classification of the NPM1 interacting partners:** (A – B) String database (version 11.0) network showing NPM1 interacting proteins identified by the protein-protein interaction profiling. Proteins present in histone acetyltransferase (HAT) complexes are shown in (A) and those involved in RNA Pol II-mediated transcription belonging to the classes of RNA Pol II subunits, general transcription factors (GTFs), TFIID associated factors (TAFs), and Mediator complex subunits, are shown in (B).

### *4.2.6. Role of NPM1 downregulation on the tumorigenic properties of oral cancer cells*

Our investigation so far has dealt with gaining insights into the molecular mechanisms of NPM1-mediated regulation of gene expression where we have presented evidence of multiple possible modes of NPM1/AcNPM1 functioning as a transcriptional coactivator. Previous studies have reported the role of NPM1 regulating RNA Pol I- and Pol II-mediated gene transcription *in vivo* (elaborated in Section 1.4.1.3.6). Earlier reports from our group implicated NPM1 in the regulation of gene expression in oral cancer where the levels of NPM1 and its acetylated pool were found to be elevated (Shandilya et al. 2009). To test the involvement of AcNPM1 in the transcriptional regulation of genes in oral cancer, we studied some of the phenotypic and molecular changes following shRNA-mediated downregulation of NPM1 expression in oral cancer. After screening various available oral cancer cell lines for the expression status of NPM1, we selected the line AW13516 having high endogenous NPM1 expression and generated an inducible Tet-On NPM1 shRNA stable cell line in that background (Section 2.3.4.3.1). We found that doxycycline treatment to these cells resulted in significant downregulation of NPM1 mRNA (Figure 2.10B) and protein levels (Figure 2.10C). To study the effects of NPM1 knockdown on the tumorigenic properties of the cells, we performed a growth curve assay using these cells with (Dox) and without (UT: untreated) doxycycline treatment. We found that doxycycline-treated cells exhibited an appreciable decrease in their growth rate compared to the untreated cells (Figure 4.2.8A). The reduced proliferation of the doxycycline-treated cells compared to the untreated cells was also evident from the smaller size of the colonies formed by the 'Dox' cells compared to the 'UT' cells in the colony formation assay (Figure 4.2.8B) although the there was no significant difference in the total number of colonies in the two groups. The doxycycline-treated cells also had slower wound closure rates compared to the untreated cells as assessed by the wound closure assay (Figure 4.2.8C). The healing of the wound can be attributed to both proliferation as well as migration-invasion properties of the cells. Hence, to test if NPM1 knockdown as any effect on the migratory properties of these oral cancer cells, we performed the wound closure assay after pre-treating the cells with mitomycin C to block their cell division. In this experiment, we observed that doxycycline-treated cells could not heal the would whereas the untreated cells partially closed the wound (Figure 4.2.8D) post-12 h of wound creation owing to their higher migration-invasion properties. The positive effect of NPM1 specifically on migration and invasion processes

was also confirmed using the transwell assays where the doxycycline-treated sample showed significantly reduced number of migrated and invaded cells in the assay compared to the untreated cells (data not shown). These results indicate that overexpressed NPM1 in oral cancer enhances the proliferation, migration and invasion features of the tumor cells.

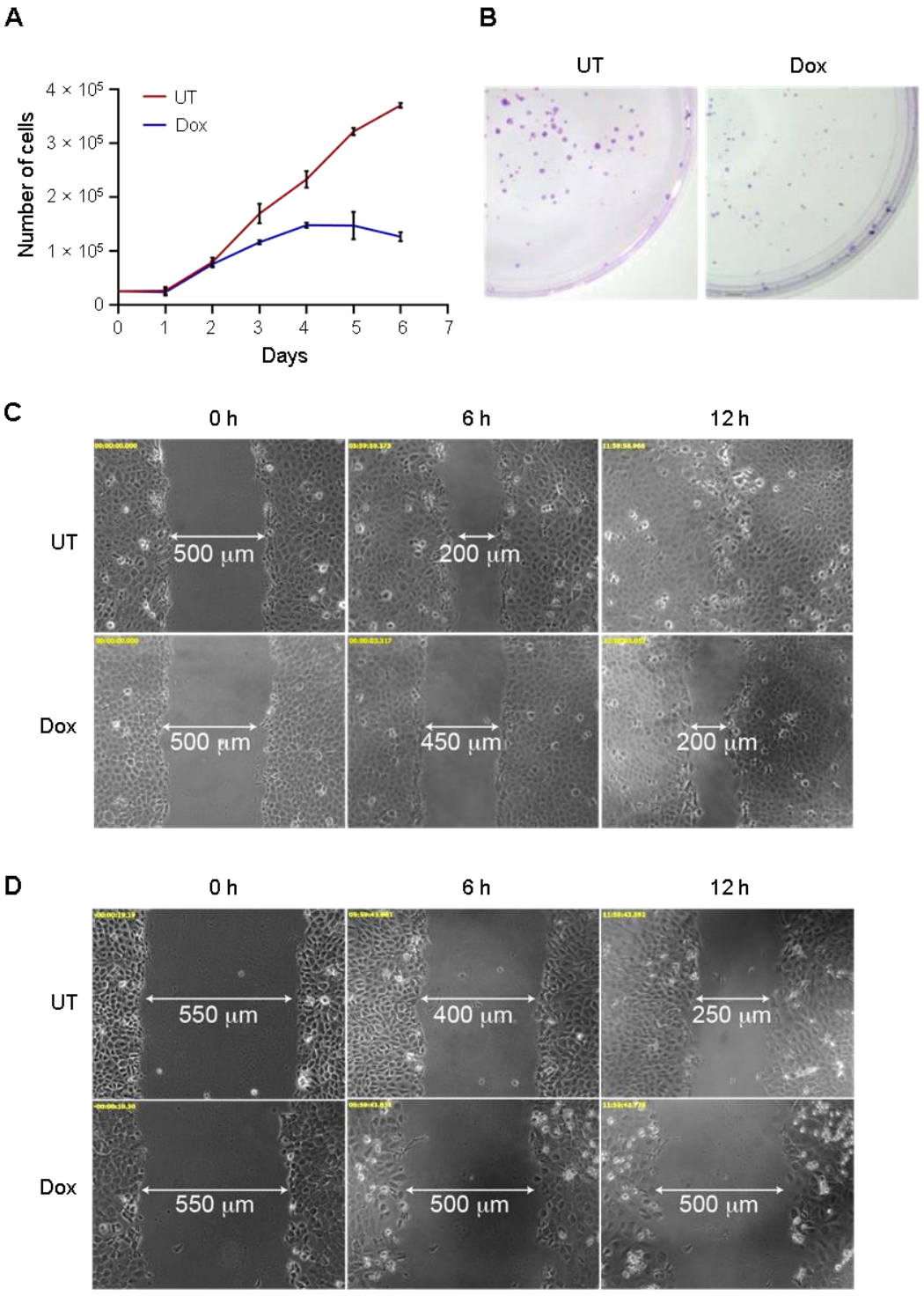



**Figure 4.2.8. Effect of NPM1 knockdown on the proliferation, migration, and invasion of oral cancer cells:** (A) Line graphs represent the number of cells with (Dox) or without (UT) doxycycline

treatment to the AW13516-shNPM1 cells. Doxycycline treatment at 2 µg/ml every 24 h induces the expression of NPM1 shRNA resulting in the knockdown of NPM1. Untreated cells were seeded at an initial seeding density of $2.5\times10^4$. Values are mean ± SEM from two independent experiments. (B) Representative photographic images showing size and number of colonies formed by the AW13516-shNPM1 cells with (Dox) or without (UT) doxycycline treatment. (C) Representative photomicrographs of AW13516-shNPM1 cells either treated (Dox) or untreated (UT) with doxycycline and captured in real-time for a period of 12 h post-wound creation showing their relative propensity to close the wound. The wound length was measured in microns post-6 h and 12 h in the two conditions (UT and Dox). (D) Representative photomicrographs of AW13516-shNPM1 cells, pre-treated with mitomycin C (5 µg/ml ) for 2 h, followed by treatment with (Dox) or without (UT) doxycycline, and captured in real-time for a period of 12 h post-wound creation showing their relative propensity to close the wound. The wound length was measured in microns post-6 h and 12 h in the two conditions (UT and Dox).

### *4.2.7. Regulation of the gene network involved in oral tumorigenesis by NPM1/AcNPM1*

Having established that NPM1 plays a role in the tumorigenic properties of oral cancer cells, we wanted to understand the molecular transcriptional changes brought about by NPM1 knockdown in the AW13516 oral cancer cells. For this, we performed RNA-sequencing with (Dox) and without (UT) doxycycline treatment to these stable cells. The RNA-seq analysis revealed that NPM1 knockdown resulted in significant alterations in gene expression (Figure 4.2.9A) and a large number of differentially expressed genes (DEGs) were obtained (Figure 4.2.9B). Out of 925 DEGs, 663 genes were found to be downregulated and 262 genes were found to be upregulated due to NPM1 knockdown (Figure 4.2.9C). Gene Set Enrichment Analyses (GSEA) (Subramanian et al. 2005) of these downregulated genes revealed that they were enriched in the gene sets related to several cancer pathways such as positive regulation of cell proliferation, positive regulation of cell migration, negative regulation of apoptosis, angiogenesis, hypoxia, epithelial-to-mesenchymal transition (EMT), among others (Figure 4.2.9D). We also found that genes in the AP-1 and HIF-1α transcription factor networks, and the TNF-α signaling through the NF-κB signaling pathway were enriched (Figure 4.2.9D). Closer analyses of the promoters of the genes downregulated by NPM1 knockdown showed that the binding motifs of these TFs namely AP-1, HIF-1α, and NF-κB, were enriched in those promoters (data not shown). This was in accordance with a previous report that showed that NPM1 downregulation resulted in a decrease in NF-κB-mediated transcription of selected target genes where NPM1 mediated the recruitment of NF-κB p65 to the gene promoters. In the same study, it was demonstrated that NPM1 was required for the expression of inflammatory genes

induced by TNF-α and lipopolysaccharide in fibroblasts and macrophages (Lin et al. 2017). This indicated that NPM1 possibly co-regulates the gene targets of these TFs.

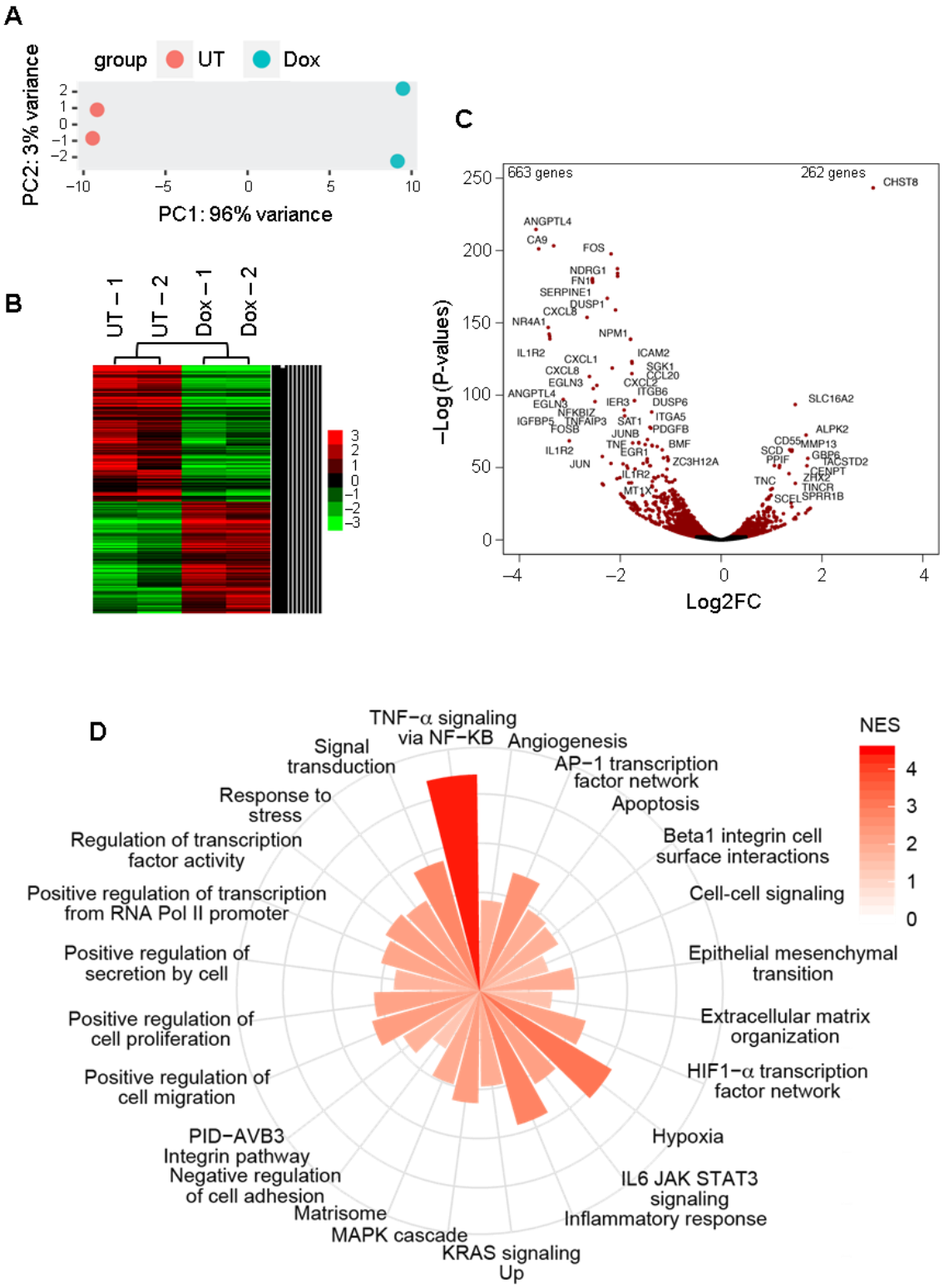



**Figure 4.2.9. Analysis of the gene network regulated by NPM1 in oral cancer:** RNA-seq was performed for AW13516 cells with (Dox) or without (UT) doxycycline treatment for the inducible expression of NPM1 shRNA resulting in NPM1 knockdown by doxycycline treatment. (A) Principal

component analysis (PCA) plot showing a good replicate agreement for the RNA-seq data. (B) Heat map depicting the differentially expressed genes between the replicates of UT and Dox samples. (C) Volcano plot showing the genes with significantly altered expression (red) after NPM1 knockdown in AW13516 cells. (D) Gene sets enriched in genes downregulated after NPM1 knockdown. The absolute value of normalized enrichment score (NES) from the Gene Set Enrichment Analysis (GSEA) is shown in the right. $P < 0.05$.

We validated the analyzed RNA-seq data by checking the expression levels of a few candidate genes from the pathways such as cell proliferation, cell migration, EMT, angiogenesis, and negative regulation of apoptosis (Table 4.2.1). We observed significant downregulation of all the tested genes which were found to be differentially expressed with highly negative rank metric score values in the analyzed RNA-seq data (Table 4.2.1 and Figure 4.2.10A). In addition, we also selected some of the genes from the same pathways whose expressions were not significantly altered in the RNA-seq analysis (having rank metric score values near zero), such as *CD44* and *MMP1* (Table 4.2.1), and found that their mRNA levels were unaffected by NPM1 knockdown (Figure 4.2.10A). This validated the RNA-seq analysis and showed that NPM1 regulates the expressions of several genes in different cancer pathways, in oral cancer cells. To test if there is a direct role of AcNPM1 in the regulation of expressions of these genes as has been indicated from our ChIP-seq data, we checked the occupancy of AcNPM1 at the promoters of few candidate genes in the AW13516 cells with and without NPM1 knockdown by doxycycline-mediated induction of NPM1 shRNA expression (Figure 4.2.10B). We observed that AcNPM1 was enriched at the promoters of all these genes tested (Figure 4.2.10C – M). For several of the gene promoters such as those of *FOS*, *EGR1*, *ANGPTL4*, and a low enrichment site (NPM1_S1) at the *NPM1* gene promoter itself, there was a significant decrease in the occupancy of AcNPM1 due to NPM1 knockdown (Figure 4.2.10C – F) while for some gene promoters such as those of *NR4A1*, *SOD2*, *PDGFB*, *SERPINE1*, *JUN*, *VEGFA*, and a low enrichment site at the *NPM1* gene promoter (NPM1_S2), the downregulation of NPM1 did not result in a concomitant decrease in AcNPM1 occupancy as expected (Figure 4.2.10G – M), indicating that the regulation of expressions of these specific genes could be indirect, involving multiple pathways and factors in play. In other words, the occupancy of AcNPM1 at the promoters of some genes (*NR4A1*, *PDGFB*, *SERPINE1*, *JUN*, and *VEGFA*), is not necessarily translated to the extent of downregulation of those genes due to NPM1 knockdown (Figure 4.2.10A, H, J – M). It is to be noted that the regions at the NPM1 promoter tested here, had comparatively lower enrichment of AcNPM1 as per the

analyzed AcNPM1 ChIP-seq data, which is also observed in our experimental data (Figure 4.2.10F – G). Interestingly, these regions were also found to be targeted by the AP-1 transcription factor c-fos as was previously demonstrated (Section 3.2.6, Figure 3.2.8). This is in accordance with another previous observation where we found enrichment of AP-1 binding motifs in the promoters of genes regulated by NPM1 and bound by AcNPM1 as analyzed from the RNA-seq (data not shown) and ChIP-seq (Figure 4.2.5) data respectively. This indicates that several genes could be co-regulated by NPM1/AcNPM1 and other protein factors by binding to the same regions in the target promoters through simple or complex mechanisms.

| **Angiogenesis** | **Cell Migration** | **EMT** | **Negative Regulation of Apoptosis** | **Cell Proliferation** |
|---|---|---|---|---|
| *NR4A1* (-3.430) | *NR4A1* (-3.430) | *CXCL8* (-2.659) | *ANGPTL4* (-3.672) | *NR4A1* (-3.430) |
| *CXCL8* (-2.659) | *CXCL8* (-2.659) | *FN1* (-2.559) | *SERPINE1* (-2.257) | *FN1* (-2.559) |
| *SERPINE1* (-2.257) | *FN1* (-2.559) | *COL16A1* (-2.357) | *NPM1* (-1.798) | *NPM1* (-1.798) |
| *SAT1* (-1.411) | *TNF* (-1.754) | *SERPINE1* (-2.257) | *TNF* (-1.754) | *TNF* (-1.754) |
| *HMOX1* (-1.404) | *HBEGF* (-1.708) | *JUN* (-1.163) | *ZC3H12A* (-1.719) | *HBEGF* (-1.708) |
| *VEGFA* (-0.730) | *PDGFB* (-1.208) | *VEGFA* (-0.730) | *HMOX1* (-1.404) | *HMOX1* (-1.404) |
| *FOS* (-2.182) | *VEGFA* (-0.730) | *CD44* (-0.440) | *VEGFA* (-0.730) | *PDGFB* (-1.208) |
| *PDGFB* (-1.208) | *CD44* (-0.440) | *MMP1* (0.657) | *SOD2* (-0.574) | *EGR1* (-1.168) |
| *JUN* (-1.163) | *MMP1* (0.657) | | *CD44* (-0.440) | *F3* (-1.097) |
| *F3* (-1.097) | | | | *VEGFA* (-0.730) |

**Table 4.2.1. Genes from different cancer pathways selected for experimental validation of RNA-seq data by RT-qPCR analysis.** The rank metric score of each gene, indicative of its fold change in expression as analyzed from the RNA-seq data, is mentioned within parenthesis adjacent to it. Higher negative values indicate higher downregulation, lesser negative or positive values indicate less significant changes in expression, and higher positive values indicate greater upregulation in expression.

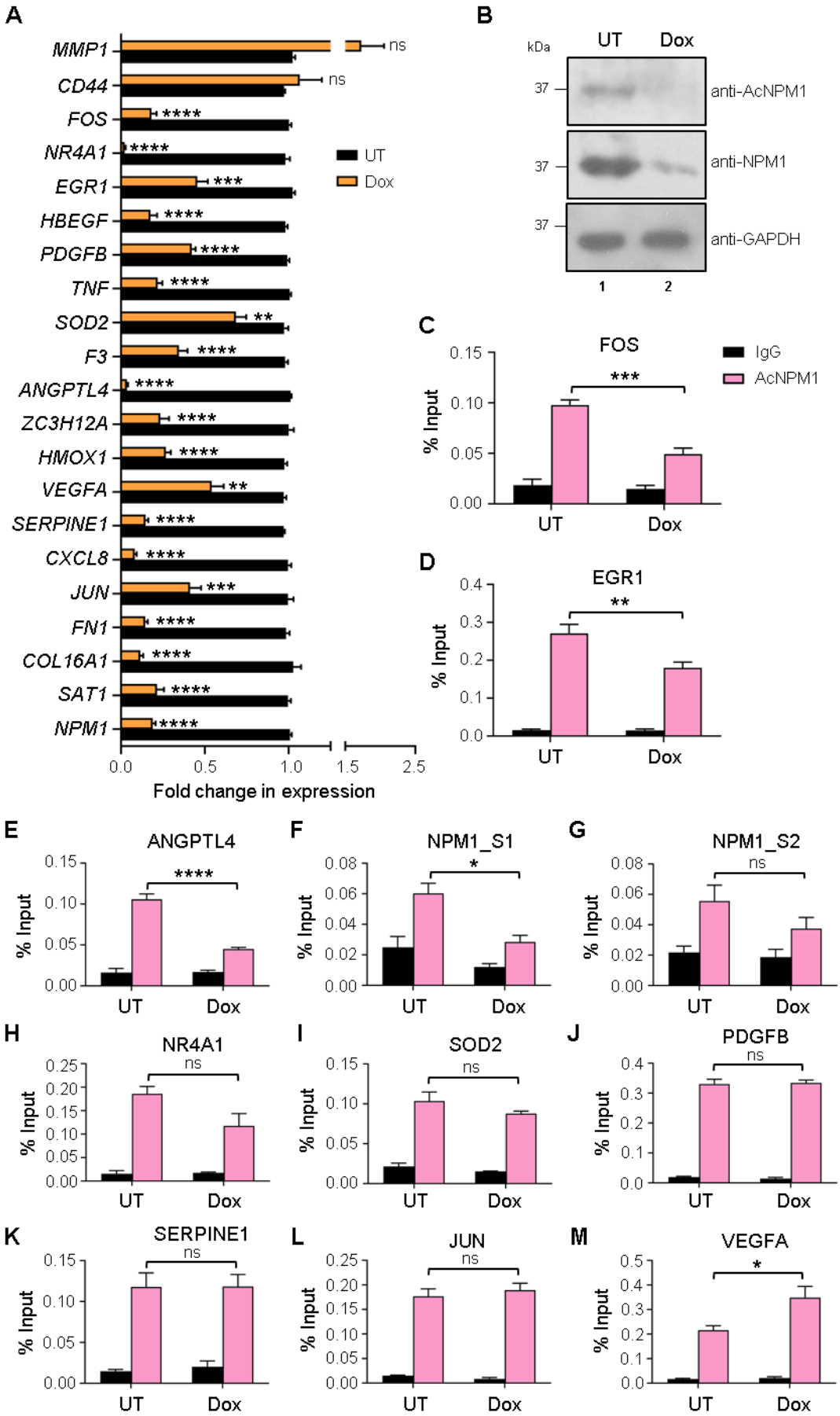

A
MMP1
CD44
FOS
NR4A1
EGR1
HBEGF
PDGFB
TNF
SOD2
F3
ANGPTL4
ZC3H12A
HMOX1
VEGFA
SERPINE1
CXCL8
JUN
FN1
COL16A1
SAT1
NPM1
ns
UT
Dox
0.0 0.5 1.0 1.5 2.5
Fold change in expression
B
kDa
UT Dox
37
anti-AcNPM1
anti-NPM1
anti-GAPDH
1 2
C
FOS
IgG
AcNPM1
% Input
D
EGR1
E
ANGPTL4
F
NPM1_S1
G
NPM1_S2
H
NR4A1
I
SOD2
J
PDGFB
K
SERPINE1
L
JUN
M
VEGFA

**Figure 4.2.10. Validation of gene targets of NPM1/AcNPM1 in oral cancer cells:** (A) Bars represent fold change in expression levels of indicated genes belonging to cancer pathways such as cell proliferation, cell migration, EMT, angiogenesis, and negative regulation of apoptosis (Table 4.2.1), as measured by RT-qPCR analysis after NPM1 knockdown by doxycycline treatment (Dox) in AW13516 cells compared to untreated (UT) control. Internal normalization was done with housekeeping gene β-actin levels. Values are mean + SEM from four independent experiments. Statistical analysis was performed using unpaired two-tailed Student's *t*-test. ***P* < 0.01, ****P* < 0.001, *****P* < 0.0001, ns: non-significant. (B) Western blot analysis showing the levels of AcNPM1 and NPM1 in AW13516 cells with (Dox) and without (UT) doxycycline-induced NPM1 knockdown. The upper panel shows western blot with anti-AcNPM1, middle panel, anti-NPM1, and bottom panel with anti-GAPDH antibodies respectively. (C – M) Bars represent the enrichment of AcNPM1 at the promoter regions of indicated genes represented as percent input (% Input) and assessed by ChIP-qPCR analysis after NPM1 knockdown by doxycycline treatment (Dox) in AW13516 cells compared to untreated (UT) control. (F – G) correspond to two regions namely site 1 (S1) and site 2 (S2), in the NPM1 promoter with lesser enrichment of AcNPM1 as determined from the analyzed AcNPM1 ChIP-seq data. Values are mean + SEM from three independent experiments. Statistical analysis was performed using one-way ANOVA, Tukey's multiple comparisons test. **P* < 0.05, ***P* < 0.01, ****P* < 0.001, *****P* < 0.0001, ns: non-significant.

## *4.3. Discussion*

The results presented in this study have shown that the acetylation of NPM1 plays an important role in the regulation of gene expression which is implicated in the tumorigenesis of oral cancer. NPM1 being a histone chaperone regulates several nuclear processes related to the histone dynamics of the chromatin such as DNA replication, repair, and transcription. Its role as a transcriptional coactivator was first demonstrated using an *in vitro* assembled chromatin template and its acetylation was proposed to play an important function presumably by disassembling the nucleosomes required for the activation of transcription *in vivo* (Swaminathan et al. 2005). However, in cells, the majority of NPM1 exists as a phosphoprotein and localized to the nucleoli. Therefore, the mechanism of regulation of RNA Pol II-driven transcription in the nucleoplasm by a predominantly nucleolar protein was unclear. It was later discovered that the acetylated pool of NPM1 (AcNPM1) is localized in the nucleoplasm as opposed to the nucleoli, including the transcription foci with RNA Pol II (Shandilya et al. 2009) which was the first indication of the role of acetylation of NPM1 in RNA Pol II-mediated transcription in cells. In this study, through genome-wide analysis of AcNPM1 occupancy in HeLa S3 cells, we find that AcNPM1 is indeed enriched at the gene regulatory elements such as promoters and enhancers, overlapping with the occupancy of RNA Pol II, p300 and epigenetic signatures of active chromatin such as the histone modification marks H3K9ac, H3K27ac, and H3K4 methylations (Figure 4.2.2 and Figure 4.2.4). Interestingly, we found enrichment of

AcNPM1 primarily at the promoters of the genes and not on the gene bodies (Figure 4.2.1), indicating its predominant role in transcription initiation. However, we cannot rule out the possibility of the presence of unmodified NPM1 on gene bodies since multiple attempts to perform NPM1 ChIP-seq have been unsuccessful due to the unavailability of ChIP-grade polyclonal antibodies against unmodified NPM1. Also, we did not find enrichment of AcNPM1 on the rDNA regions (data not shown), as opposed to NPM1 in general, which was shown to occupy rDNA regions (Murano et al. 2008). This is in accordance with our previous observation where AcNPM1 was found to be localized in the nucleoplasm including the transcription foci while NPM1 is majorly present in the nucleoli (Shandilya et al. 2009).

We have done investigations towards understanding the biochemical mechanisms of NPM1/AcNPM1-mediated transcriptional coactivation. Previous studies from our group have shown that NPM1 is acetylated by p300, which increases its histone chaperone activity and transcription activation potential (Swaminathan et al. 2005). AcNPM1 was found to have increased binding affinity towards acetylated histones, which potentially could lead to enhanced nucleosome disassembly at the promoters (Swaminathan et al. 2005). The crystal structures of NPM1 and its homologs in the Nucleoplasmin family have helped us understand the mode of interactions between NPM1 and the core histones and their implications in the histone chaperone activity of NPM1 (Dutta et al. 2001; Namboodiri et al. 2004a; Lee et al. 2007a; Platonova et al. 2011). From such studies it was clear that the oligomerization of NPM1 increases the efficiency of its interactions with core histones, thereby enhancing its histone chaperone activity. To understand if this histone chaperone activity of NPM1 is important for its transcription coactivation function, oligomerization deficient mutants of NPM1 namely NPM1 L18Q and NPM1 Y17T-C21F, were previously generated, which were found to be deficient in their histone chaperone activity as well as transcription activation potential (Senapati P, Ph.D. thesis, 2014). Hence, histone chaperone function mediated through its oligomerization was found to be important for NPM1 to activate transcription from the chromatin template.

Another hypothesized mechanism of NPM1-mediated transcription regulation was through its interaction and recruitment of the transcription machinery and other transcription-related proteins at the gene promoters to regulate gene expression. To test this possibility, we have compared the genome-wide occupancy profiles of various proteins with roles in transcription, with that of AcNPM1. The overlap analyses of these proteins with AcNPM1

revealed that AcNPM1 co-occupies the genomic regions with several such proteins to different extents, the highest being for the transcription factor MAX (~60%), followed by RNA Pol II (~52%) and the remodeler CHD2 (~50%) (Figure 4.2.6). There was also an appreciable overlap of the AcNPM1 peaks with those of TFIIF-alpha (GTF2F1) (~30%) and the subunits of the TFIID complex namely TBP and TAFII-250 (TAF1) (> 40%) (Figure 4.2.6), which are components of the RNA Pol II machinery, further indicating that AcNPM1 might act as a recruiter of the RNA Pol II transcription complex to the gene regulatory elements. Besides these, the co-occupancy of AcNPM1 with various other TFs and enrichment of those TFs' binding motifs in the AcNPM1 peaks (Figure 4.2.6) suggests the co-regulation of gene expression from their promoters by AcNPM1 along with the other specific protein factors. To gain insights into this angle, high-throughput approaches were undertaken to identify potential interacting partners of NPM1 (Senapati P, Ph.D. thesis, 2014). Through this approach, a multitude of proteins was identified to interact with NPM1. Analysis of the biological functions of these interacting partners of NPM1 revealed that several factors belonged to the classes such as histone acetyltransferase complexes, kinases, transcription factors, remodelers, coactivators, RNA Pol II transcription machinery, core and variant histones, among others (Figure 4.2.7 and Figure 4.2.8). In accordance with the overlap ChIP-seq analysis for AcNPM1 and CHD2 discussed previously, we found strong interaction of NPM1 and CHD2 in the protein interaction profiling, suggesting that NPM1/AcNPM1 might associate with remodelers such as CHD2 to facilitate nucleosome disassembly. We have validated the high-throughput protein-protein interaction analysis with selected candidate interacting partners namely the transcriptional activator AP-4, the RNA Pol II subunit POLR2K, the transcriptional coactivator PC4, and the transcriptional repressor SNAI1 (Figure 4.2.7), which indicates that NPM1/AcNPM1 could potentially interact with such transcription-related proteins and regulate gene transcription.

As a functional validation of the transcriptional coactivation role of NPM1/AcNPM1, we studied its involvement and its gene targets in oral cancer where it was found to be overexpressed with a concomitant increase in its acetylated pool (Shandilya et al. 2009). We found that knockdown of NPM1 in the oral cancer cell line AW13516 reduces the proliferation, migration, and invasion properties of the cells (Figure 4.2.9), at least by regulating the expression of several genes involved in these cancer pathways (Figure 4.2.10 and Figure 4.2.11). Several of these gene promoters are occupied by AcNPM1 which shows

that AcNPM1 in also involved in the regulation of expression of these genes through direct or indirect mechanisms (Figure 4.2.11).

The positive role of NPM1 in oral tumor manifestation has been validated in an orthotopic mouse model. We have investigated if knockdown of NPM1 would affect oral tumor growth in mice. We xenografted cells from a stable cell line expressing the Luciferase gene in the AW13516-shNPM1 background, into the floor of the mouth region of 2-month old male NOD.CB17-*Prkdc*$^{scid}$/J mice (n = 14). These injected AW13516-shNPM1-luc+ cells developed into oral tumors in 5 days. One group of mice were then fed with doxycycline (1 mg/ml with 5% w/v sucrose in water) (Dox) to induce the expression of NPM1 shRNA in the tumor cells, while the control group was fed with the vehicle (Veh). The growth of the tumors was assessed by measuring the tumor flux (Luciferase activity) using a live animal imaging system. It was found that at Day-16 or 11 days post-doxycycline treatment, there was a significant decrease in the tumor flux of the Dox group compared to the Veh group (Figure 4.3.1A – B), indicating that downregulation of NPM1 in tumor cells could result in slower tumor growth. To rule out the possibility of doxycycline itself on the tumor growth, a parallel experiment was set up using another control group of mice (n = 5) which were injected with AW13516-luc+ cells (lacking the NPM1 shRNA) and fed similarly with doxycycline or the vehicle. There was no significant difference in the tumor flux of the Dox and Veh group from this set of animals (data not shown), indicating that there is no effect of doxycycline per se on the oral tumor growth. To confirm the downregulation of NPM1 in these tumor cells, we checked the levels of NPM1 as well as the proliferation marker Ki67, in the tumor tissue samples from the Dox and Veh groups of the set of mice injected with the AW13516-shNPM1-luc+ cells, by immunohistochemistry analysis. As expected, we observed a significant decrease in the levels of NPM1 protein (Figure 4.3.1C – D) in the Dox group compared to the Veh group. We also observed a decreased staining for Ki67 in the Dox samples compared to Veh, indicating that NPM1 knockdown indeed results in the reduction in proliferation of the cells (Figure 4.3.1C). Since we observed that NPM1 knockdown results in a decrease in EMT gene expression (Figure 4.2.10 and Figure 4.2.11), we checked the protein levels of E-cadherin, an epithelial marker, and Fibronectin, a mesenchymal marker, in the tumor samples and found that there was an increase in E-cadherin and decrease in Fibronectin levels in the Dox samples compared to the Veh samples, showing the reduction in EMT at the molecular level due to NPM1 knockdown (data not shown) consistent with our *ex vivo* cellular assays (Figure 4.2.9). In accordance

with this observation, we indeed noticed metastasis in adjoining lymph nodes in some of the mice from the Veh group, which was very less in the Dox group of animals (data not shown).

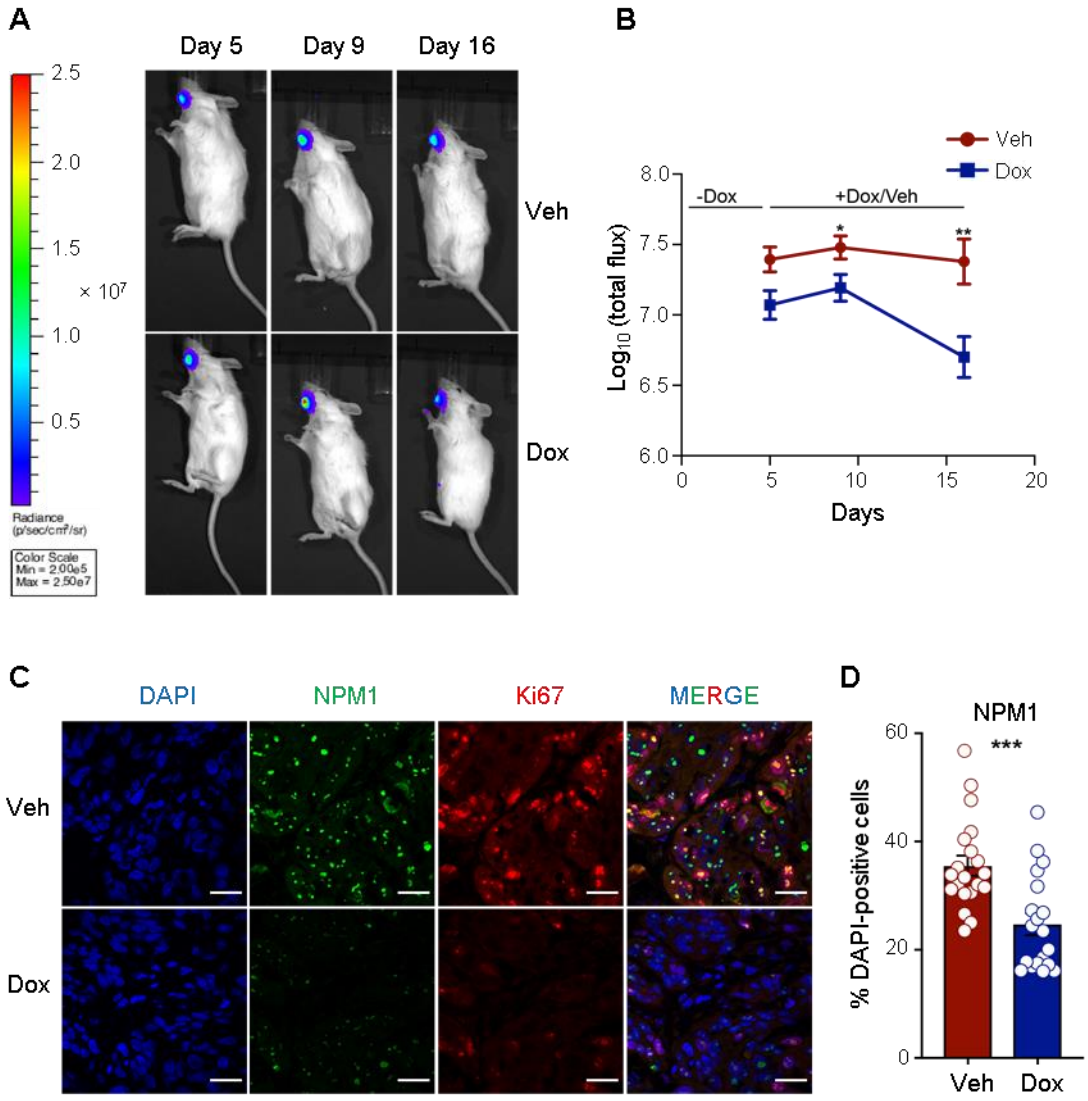



**Figure 4.3.1. Effect of NPM1 knockdown on the oral tumor growth in mice:** (A) Representative bioluminescent images of Vehicle (Veh) or doxycycline (Dox)-fed mice at 5, 9 and 16 days post-xenograft of the floor region of the mouth of each mouse with $1\times10^6$ AW13516-shNPM1-luc+ cells. Units shown at the left, represent relative light units. (B) Line curves represent bioluminescence intensity measured at 5, 9 and 16 days post-injection of the cells, for the Veh and Dox groups. Values are mean ± SEM of $\log_{10}$(total flux), n = 7 animals in each group. Statistical analysis was performed using unpaired two-tailed Student's *t*-test. **P* < 0.05, ***P* < 0.01. (C) Representative immunohistochemical images showing staining of the tumor tissues from Veh and Dox groups with NPM1 (green) and Ki67 (red) antibodies. Nuclei were stained with DAPI. Scale bar is 25 µm. (D) Bars represent the number of NPM1-positive cells expressed as a percentage of DAPI-positive cells. n = 5 animals from each group with 3 – 4 fields from each animal. Statistical analysis was performed using the two-tailed Mann-Whitney test. ****P* < 0.001.

In summary, these results suggest that NPM1 overexpression in oral cancer might enhance the expression of tumor-promoting genes presumably through AcNPM1-mediated transcriptional regulation (Figure 4.3.2). However, it should be kept in mind that the transcriptional regulation of cancer-associated genes might not be the only mode of oncogenic functioning of NPM1 in these cells. The overexpression of NPM1 in different cancers including oral cancer may enhance the rDNA transcription and ribosome biogenesis. In addition, the molecular chaperone function of NPM1 might be important in this context in catering to the increased protein synthesis requirement in the rapidly proliferating tumor cells, thereby avoiding a misfolded protein stress response. The other functions of NPM1 such as enhancing DNA synthesis, inhibition of pro-apoptotic proteins, and facilitating the progression through mitosis, could also contribute to the survival of the tumor cells (Grisendi et al. 2006). Presumably, one or more of such mechanisms could be employed by NPM1 to contribute to tumorigenesis.

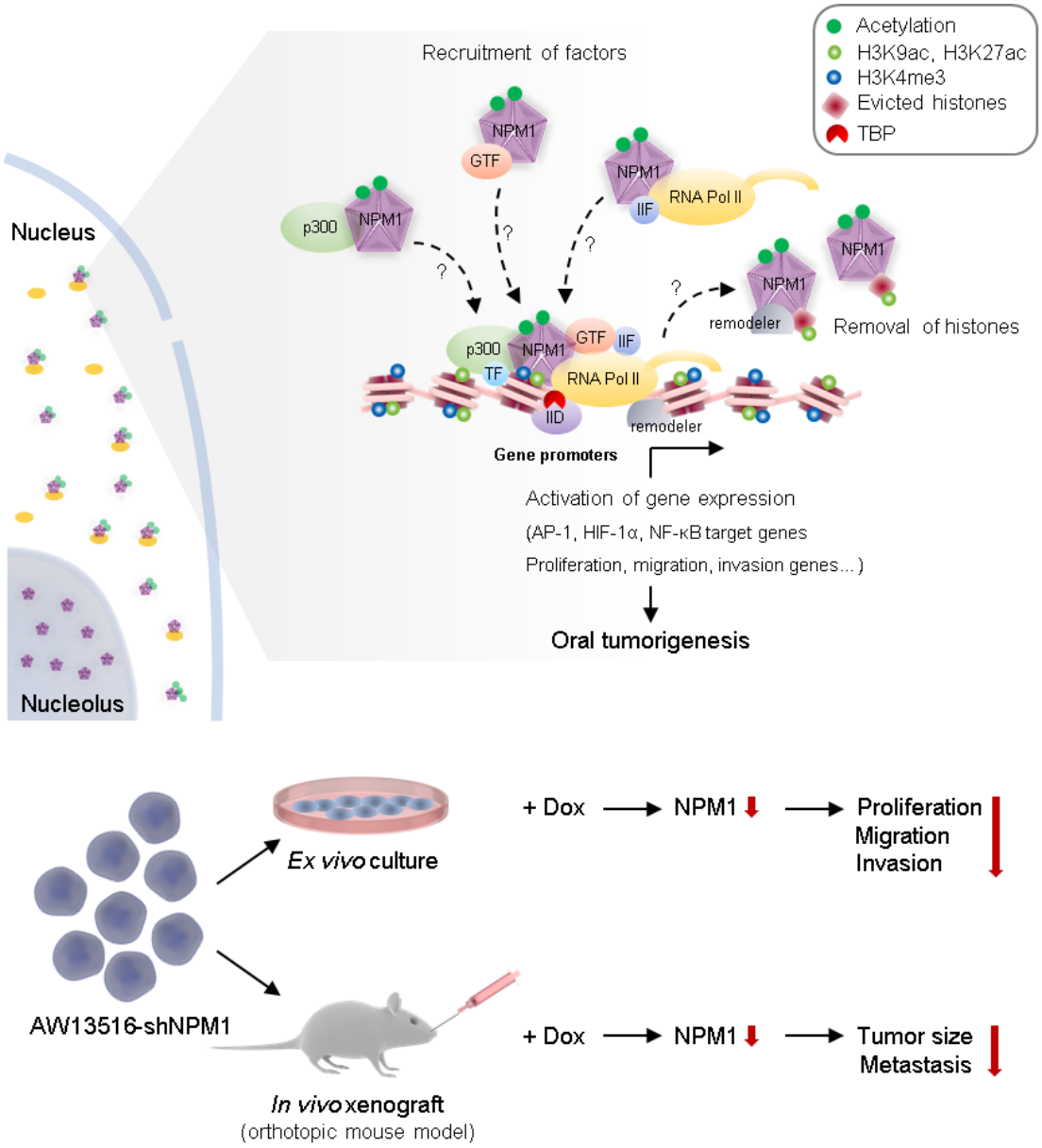



**Figure 4.3.2. Proposed model depicting the role of NPM1/AcNPM1 in the regulation of expression of genes and their implications in oral tumorigenesis:** NPM1 is predominantly present in the nucleolus while the acetylated NPM1 (AcNPM1) pool localizes in the nucleoplasm including the transcription foci along with RNA Polymerase II. Specifically, AcNPM1 is enriched at the gene promoters, overlapping with the occupancy of RNA Pol II, as well as activating histone modification marks H3K9ac, H3K27ac, and H3K4me3. AcNPM1 might function as a coactivator by recruiting the transcription machinery through direct interactions with RNA Pol II subunits, p300, and general transcription factors (GTFs) such as TFIID and TFIIF. It might co-regulate the expressions of genes that are targets of other transcription factors (TFs) such as AP-1, HIF-1α, NF-κB, among others, through protein-protein interactions. It might also act through its histone chaperone activity and/or associate with remodelers such as CHD2 to facilitate nucleosome disassembly during transcription. Several genes implicated in processes such as cell proliferation, migration, and invasion, are regulated by NPM1/AcNPM1, which are manifested in oral tumorigenesis. The process is evident from the decrease in cell proliferation, migration, and invasion of cultured AW13516 oral cancer cells, and reduced oral tumor size and lymph node metastasis in an orthotopic mouse model, after doxycycline-induced shRNA-mediated knockdown of NPM1.

# *Chapter 5: Functional Characterization of Mammalian Nucleoplasmin (NPM2) as a Potential Regulator of Transcription*

## *5.1. Introduction*

NPM2 is the mammalian ortholog of *Xenopus* Nucleoplasmin (NP), which was the first histone chaperone to be described (Laskey et al. 1978). The properties of Nucleoplasmin with respect to its interaction with histones have been extensively characterized biochemically and biophysically, and its role in egg competence and sperm chromatin remodeling in the *Xenopus* and mouse models has been well-established (described in detail in Section 1.4.2.2.1). The paralog of NPM2, that is Nucleophosmin (NPM1/B23/Numatrin) also exhibits properties of a histone chaperone among other functions (as described in Section 1.4.1.3). Among the many functions, NPM1 is a potent regulator of RNA Polymerase I- and RNA Polymerase II-driven transcription which could be brought about through various mechanisms such as its histone chaperone activity, interaction with transcription-related proteins under different contexts, induction of autoacetylation of the lysine acetyltransferase and transcriptional coactivator protein p300, and so on (elaborated in Section 1.4.1.3.6). Human Nucleoplasmin (NPM2) shares about 40% identity in protein sequence with NPM1, but is structurally quite similar to NPM1 in its N-terminal oligomerization domain, the acidic tracts and the bipartite nuclear localization signal (Lee et al. 2007a; Platonova et al. 2011). Owing to such structural similarity and molecular properties, it can be expected that these proteins belonging to the Nucleoplasmin family of histone chaperones, could have some common functions and mechanisms of action. However, at the same time, it is wise to expect variations and uniqueness in the exhibition of the functions due to the presence of unique regions in the NPM1 protein that is not shared with its paralogs such as NPM2, and expression pattern of these proteins. NPM2 lacks an acidic tract in the N-terminal oligomerization domain, and a substantial portion in the C-terminal disordered region including the nuclear export signal, nucleolar localization signal, and nucleic acid-binding domain, which are present in NPM1 that contribute to the functional diversity of NPM1. NPM2, unlike NPM1, is also expressed in the cells in a tissue-restrictive manner. While NPM1 is a ubiquitously expressed protein, NPM2 is

reported to be expressed only in the oocytes in different models tested such as the mouse, *Xenopus*, *Bos taurus*, and zebrafish. The other member of the Nucleoplasmin family, NPM3, is also known to be ubiquitously expressed. However, not many functions have been discovered for this protein. With respect to its basic molecular properties, it was found that NPM3 lacks an inherent histone chaperone activity, a property that is exhibited significantly by NPM1, but can act as an enhancer of activator-dependent transcription in a cellular system (Gadad et al. 2010). Such differences in the expression patterns of these NPM proteins should have some evolutionary significance which is yet to be uncovered. Given the fact that both NPM1 and NPM3 had positive effects on the activation of transcription, *in vitro* and/or in a cellular system, we hypothesized that NPM2 could be a regulator of transcription through common or unique mechanisms. Previous reports have hinted towards a transcription regulation function of Nucleoplasmin and its orthologs. Nucleoplasmin was found to stimulate transcription factor binding to nucleosomes and factor-induced nucleosome disassembly, implicating its potential ability to activate transcription (Chen et al. 1994; Walter et al. 1995). It was found that hyperphosphorylated Nucleoplasmin isolated from *Xenopus* eggs could efficiently unfold sperm and somatic chromatin by removing the chromosomal proteins from linker DNA regions (Ramos et al. 2005), a mechanism which can be potentially utilized for activating transcription from the chromatin. In another amphibian model, *Pleurodeles waltl*, Nucleoplasmin was found to be localized on the lampbrush loops, that is on the sites of transcription and apparently associated with the ribonucleoprotein (RNP) particles which suggested that Nucleoplasmin may play a role in transcriptional activity in cells (Moreau et al. 1986). Nucleoplasmin orthologs in mouse and zebrafish have been shown to have critical roles in zygotic gene expression (Burns et al. 2003; Bouleau et al. 2014) and human phospho-mimic NPM2 was found to regulate the expression of genes involved in naïve stem cell stage during induced pluripotent stem cell generation (Huynh et al. 2016). However, the precise role of NPM2 as a regulator of chromatin transcription and the underlying mechanisms are unknown. It is unclear if NPM2 has intrinsic transcription regulation properties which could be demonstrated in *in vitro* transcription assays, or the actual mechanism is more complicated than this where this apparent transcription regulation function is manifested under specific contexts. We intended on finding the answers to such questions described above, and conceived and designed our study towards this angle.

## 5.2. Results

### 5.2.1. Histone chaperone activity of NPM2

At the inception of this study, we first characterized the biochemical properties of human NPM2 as a histone chaperone. We purified the $His_6$-tagged NPM2 protein after expressing it in *E. coli* cells and checked its intrinsic histone chaperone activity by the *in vitro* histone transfer or the plasmid supercoiling assay (Section 2.11.6). The assay was performed taking 5, 10 and 15 pmol dose of $His_6$-NPM2, keeping $His_6$-tagged NPM1 as a positive control for the assay (Okuwaki et al. 2001b; Swaminathan et al. 2005). The histone chaperone activity was assessed by the ability of the protein to recover the supercoiled form of a pre-relaxed plasmid by assembling the nucleosomes on the template. The activity was quantified by densitometric scanning of the band corresponding to the most supercoiled form recovered in the assay and expressed as a percentage of or fold change over the input.

The results revealed that NPM2 had minimal intrinsic histone chaperone activity (Figure 5.2.1A lanes 6 – 8) when compared to the basal level ('no chaperone', Figure 5.2.1A lane 3). The positive control NPM1 showed higher histone chaperone activity (Figure 5.2.1 lanes 9 – 11). We repeated the assay with higher doses of NPM2 (in the range of 5 – 45 pmol) to check if there is any specific stoichiometry of core histones to NPM2 that results in the highest exhibition of the histone chaperone activity of NPM2. In our standardized protocol (Section 2.11.6), we have observed the highest histone chaperone activity of NPM1 in the range of 5 – 10 pmol dose that was consistent across experiments. Hence, we kept 10 pmol dose of NPM1 as the positive control for our assay. We found that there was no noticeable increase in the activity of NPM2 in any of the doses tested in this assay (Figure 5.2.1B, lanes 4 – 12) compared to the basal level (Figure 5.2.1B, lane 3) or the positive control (Figure 5.2.1B, lane 13). We have tested the activity of NPM2 at a dose lower than 5 pmol (that is 2.5 pmol) but did not observe any significant increase in its histone chaperone activity (data not shown). The activity quantified from the results of multiple independent experiments showed that NPM2 had minimal histone chaperone activity (Figure 5.2.1C) which was about 2-fold at the doses tested, but statistically significant when compared to the basal level ('no chaperone', Figure 5.2.1D). These observed results were in accordance with the observations reported previously (Okuwaki et al. 2012).

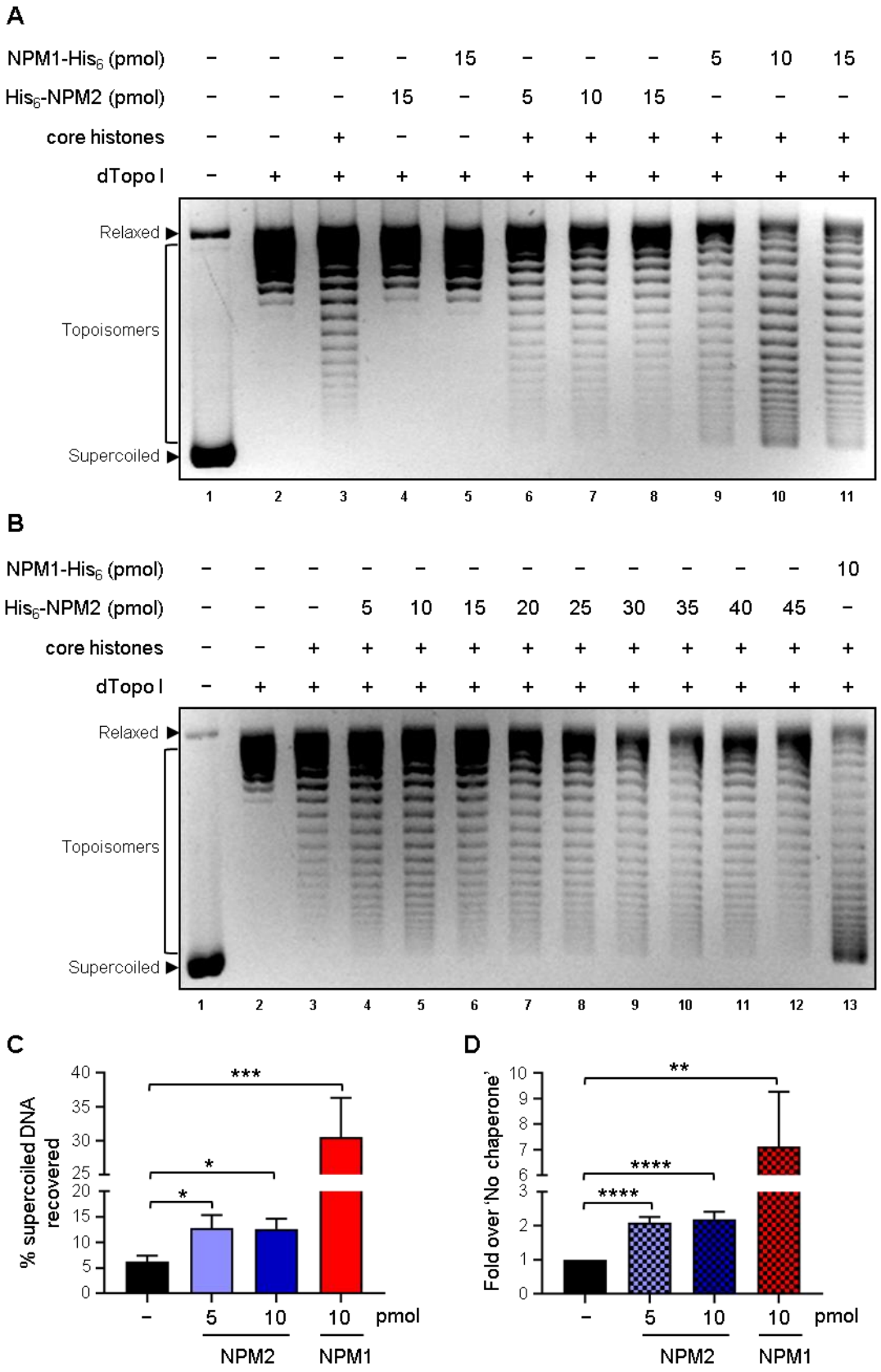



**Figure 5.2.1. Histone chaperone activity of NPM2:** (A) Histone transfer or plasmid supercoiling assay with increasing doses (5, 10 and 15 pmol) of $His_6$-NPM2 (lanes 6 – 8) and NPM1-$His_6$ (lanes 9 – 11). Lane 1 (Input) indicates the reference position of the relaxed and supercoiled form of the

plasmid. Lane 2 (negative control) indicates the sample without core histones. Lane 3 indicates the sample with core histones only (without any chaperone) that marks the basal level of supercoiled DNA generated in the absence of any chaperone in the assay. Lanes 4 and 5 are other negative controls (with the highest dose of histone chaperone but no core histones). (B) Histone transfer or plasmid supercoiling assay with increasing doses (5 – 45 pmol as indicated) of $His_6$-NPM2 (lanes 4 – 12) and 10 pmol of NPM1-$His_6$ (lanes 13, serving as the positive control). Lane 1: Input, lane 2: negative control (without core histones), lane 3: basal activity (with core histones only, without any chaperone). (C – D) Bars represent histone chaperone activity of indicated doses of $His_6$-NPM2 and NPM1-$His_6$ quantified in terms of (C) percentage supercoiled DNA recovered with reference to Input (lane 1) and (D) fold change in activity measured in (C) over the basal level ('No chaperone', lane 3). Values are mean + SEM from 12 independent experiments. Statistical analysis was performed using Student's *t*-test. **P* < 0.05, ***P* < 0.01, ****P* < 0.001, *****P* < 0.0001.

We hypothesized that the low histone chaperone activity of NPM2 in comparison to NPM1 was probably due to the absence of one of the acidic tracts, A1, in NPM2 which is present in NPM1 (Section 1.4, Figure 1.9). These acidic stretches are important for the interaction of the histone chaperone with the core histones thereby affecting their histone chaperone activity. To test this hypothesis, we wished to perform the assay using recombinant human, mouse (which also lacks the A1 acidic tract; Section 1.4.2, Figure 1.14) and *Xenopus* (which contains the A1 acidic tract; Section 1.4.2, Figure 1.14) Nucleoplasmin proteins to compare their activities. We purified the $His_6$-tagged mouse (Npm2) and *Xenopus* (NP) Nucleoplasmin proteins and performed the assay along with human $His_6$-tagged NPM2 and NPM1 (positive control). Results showed that all three orthologs of NPM2 exhibited a lower histone chaperone activity compared to that of NPM1 at the doses tested (2.5, 5 and 10 pmol of the histone chaperone proteins) (Figure 5.2.2). Among the NPM2 orthologs, mouse Nucleoplasmin (Npm2) showed the lowest activity whereas the activities of human (NPM2) and *Xenopus* (NP) Nucleoplasmin were similar. The assay was repeated three times and showed a similar trend in the results (quantification not shown). This showed that in this *in vitro* system using the recombinant $His_6$-tagged proteins, there was no marked difference in the histone chaperone activity of the NPM2 orthologs that could have resulted due to the presence or absence of the first acidic (A1) tract.

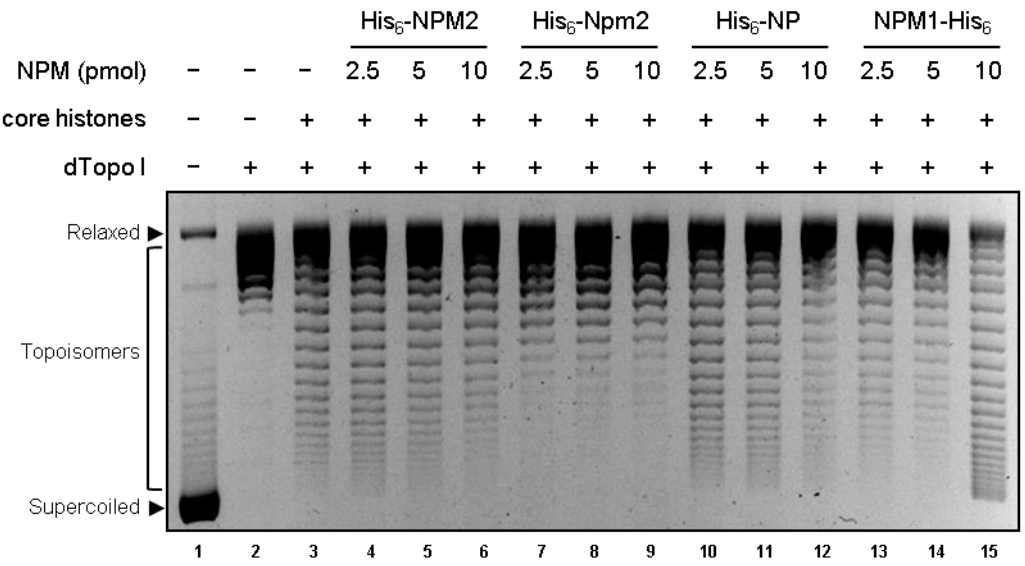



**Figure 5.2.2. Histone chaperone activity of NPM2 orthologs:** Histone transfer or plasmid supercoiling assay with increasing doses (2.5, 5 and 10 pmol) of $His_6$-NPM2 (human ortholog of Nucleoplasmin, lanes 4 – 6), $His_6$-Npm2 (mouse ortholog of Nucleoplasmin, lanes 7 – 9), $His_6$-NP (*Xenopus* Nucleoplasmin, lanes 10 – 12) and NPM1-$His_6$ (lanes 13 – 15). Lane 1 (Input) indicates the reference position of the relaxed and supercoiled form of the plasmid. Lane 2 (negative control) indicates the sample without core histones. Lane 3 indicates the sample with core histones only (without any chaperone) that marks the basal level of supercoiled DNA generated in the absence of any chaperone in the assay.

From the results presented so far, we found that recombinant unmodified N-terminal $His_6$-tagged Nucleoplasmin did not show significantly enhanced histone chaperone activity in the supercoiling assay. We noted that the NPM2 construct used for the protein purification had an N-terminal $His_6$ tag, while the NPM1 protein was C-terminal $His_6$-tagged. The N-terminal domain of the NPM proteins consists of the oligomerization domain which primarily interacts with the core histones, while the C-terminal disordered tail helps in stabilizing the interaction of NPM with the core histones, which was first elucidated in the case of *Xenopus* Nucleoplasmin (Dutta et al. 2001). We hypothesized that the N-terminal $His_6$ tag for recombinant NPM2 could contribute to minor variations in the three-dimensional structure of NPM2 and interfere in the proper conformation of the protein that is required for the N-terminal oligomerization domain-mediated interaction with the core histones. To test this possibility, we generated a bacterial expression construct for C-terminal $His_6$-tagged NPM2 (NPM2-$His_6$) (Section 2.2.1.1) similar to the NPM1-$His_6$ construct, and purified the protein after expressing in *E. coli* cells. On a parallel note, we speculated that the native form of Nucleoplasmin (with its native post-translational modifications and associated factors) could exhibit a better histone chaperone activity

compared to the recombinant unmodified protein purified from *E. coli* cells. To test this, we generated a stable cell line in HEK-293 background for the constitutive overexpression of 3xFLAG-NPM2 (Section 2.3.4.2.1) and purified the FLAG-tagged NPM2 from these cells using anti-FLAG M2 agarose resin. A similar purification from an empty vector (EV)-stable cell line in HEK-293 was performed to serve as a negative control for the FLAG-tagged NPM2 complex.

We performed the supercoiling assay using normalized amounts of the NPM2-$His_6$ and 3xFLAG-NPM2 (along with the EV negative control) and compared it with the histone chaperone activity of $His_6$-NPM2, keeping NPM1-$His_6$ as the positive control for the assay. Preliminary results showed that NPM2-$His_6$ did not exhibit any increased histone chaperone activity compared to $His_6$-NPM2 (Figure 5.2.3, compare lanes 6 – 7 with 4 – 5). This rules out the possibility that the orientation of the $His_6$ tag has any major influence on the histone chaperone activity of recombinant NPM2 purified from the *E. coli* cells. We also did not observe any increased activity of the FLAG-tagged NPM2 *in vivo* pulled down from cells (which was presumed to be a native form of the protein) compared to $His_6$-NPM2 (Figure 5.2.3, compare lanes 10 – 11 with 4 – 5) as we had hypothesized, although it was marginally higher than the complex pulled down the empty vector negative control cells and used in equivalent amounts (Figure 5.2.3, compare lanes 10 – 11 with 8 – 9). Other positive (Figure 5.2.3, lanes 1, 12 – 13), and negative controls (Figure 5.2.3, lanes 2, 14 – 15) and basal activity ('no chaperone', Figure 5.2.3, lane 3) for the assay were taken, as described before.

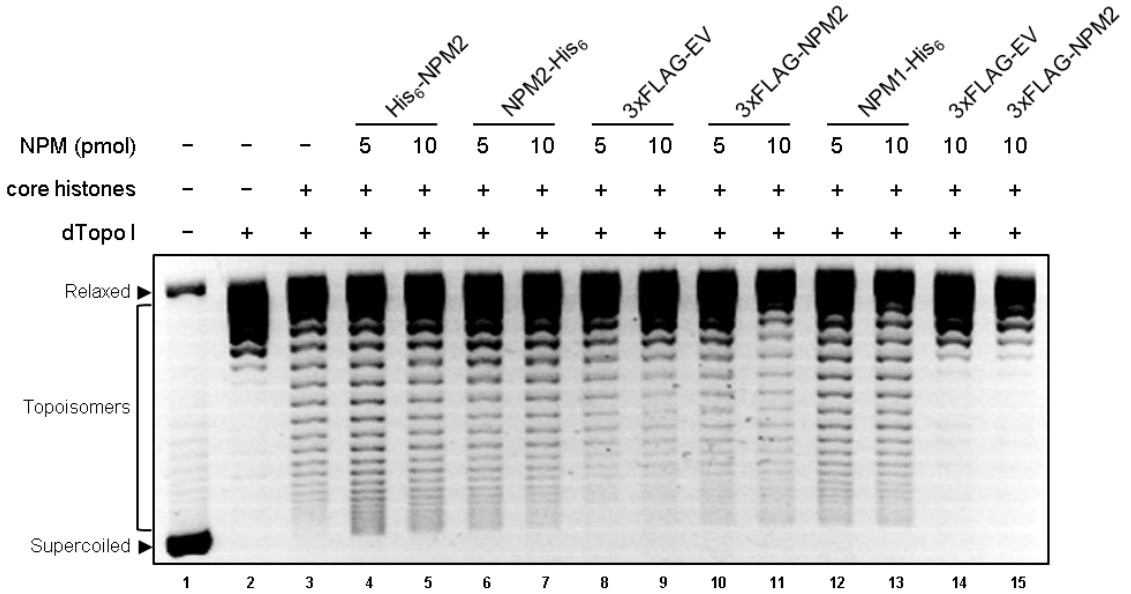


**Figure 5.2.3. Histone chaperone activity of $His_6$- and FLAG-tagged NPM2:** Histone transfer or plasmid supercoiling assay with increasing doses (5 and 10 pmol or their equivalent amounts) of $His_6$-NPM2 (lanes 4 – 5), NPM2-$His_6$ (lanes 6 – 7), FLAG immuno-pull-down (IP) complex from empty vector (EV) stable HEK-293 cells (lanes 8 – 9), 3xFLAG-NPM2 IP complex from HEK-293 stable cells (lanes 10 – 11), and NPM1-$His_6$ (lanes 12 – 13). Lane 1 (Input) indicates the reference position of the relaxed and supercoiled form of the plasmid. Lane 2 (negative control) indicates the sample without core histones. Lane 3 indicates the sample with core histones only (without any chaperone) that marks the basal level of supercoiled DNA generated in the absence of any chaperone in the assay. Lane 14 – 15 (negative controls) indicate the samples without core histones but with the FLAG-IP complexes pulled down from HEK-293 stable cells as indicated.

From the results presented so far, we speculate that the histone chaperone activity of NPM2 could be manifested in a more *in vivo* context where it is naturally expressed. It is possible that the protein requires its natural niche, with its post-translational modifications and tissue-specific associated factors for manifestation of its function. On the same note, it is also possible that NPM2 could be a special chaperone for the germ cell-specific histone variants besides the core histones. These hypotheses need further investigations for validation.

### *5.2.2. Transcription regulation potential of NPM2*

As discussed in the preceding sections, there have been few reports about the role of Nucleoplasmin and its orthologs in the regulation of transcription, which could be direct or indirect. Biochemically, Nucleoplasmin purified in its native form from *Xenopus* eggs, was found to stimulate transcription factor binding to nucleosomes and factor-induced nucleosome disassembly, a mechanism that can potentially activate transcription *in vivo* (Chen et al. 1994; Walter et al. 1995). Hyperphosphorylated Nucleplasmin isolated from *Xenopus* eggs was found to efficiently unfold sperm and somatic chromatin by removing the chromosomal proteins from linker DNA regions (Ramos et al. 2005), another mechanism which can be implicated in transcriptional activation from the chromatin. In another amphibian model, *Pleurodeles waltl*, Nucleoplasmin was found to be localized on the lampbrush loops, at the sites of transcription and associated with the ribonucleoprotein particles, suggesting a possible role of Nucleoplasmin in transcriptional activity in cells (Moreau et al. 1986). Further, Nucleoplasmin orthologs in mouse and zebrafish were found to be important for zygotic gene expression (Burns et al. 2003; Bouleau et al. 2014) and human phospho-mimic NPM2 was found to regulate the expression of genes involved in

naïve stem cell stage during induced pluripotent stem cell generation (Huynh et al. 2016). This suggests that NPM2 could act as a potential regulator of transcription, in which case, the mechanisms are yet to be uncovered. We were interested in addressing this question and wanted to test if NPM2 by itself could activate transcription from the chromatin template *in vitro*.

Previous studies reported by our group and others have shown that the paralog of NPM2, that is NPM1, is a regulator of RNA Polymerase I- (Murano et al. 2008) and RNA Polymerase II- (Swaminathan et al. 2005; Shandilya et al. 2009) driven transcription. The mechanisms of NPM1-mediated transcription regulation could be varied and context-dependent as discussed in the previous chapters 1 (Section 1.4.1.3.6) and 4 (Sections 4.1, 4.3), such as through its histone chaperone activity (Murano et al. 2008), interactions with transcription-related proteins (Section 4.2.5), modulation of activity of epigenetic enzymes such as p300 (Arif et al. 2010; Kaypee et al. 2018b) which play roles in transcription activation, and so on. While histone chaperone activity of NPM1 was found to be important for its transcription activation potential, we do not observe the same scenario in the case of the other paralog of NPM2, which is NPM3. While NPM3 was not found to possess intrinsic histone chaperone activity, it could enhance activator-dependent transcription of a reporter gene in a cellular system. We wished to test if NPM2 could activate chromatin transcription despite not being an efficient histone chaperone as NPM1. We performed the *in vitro* chromatin transcription assay (Section 2.11.7) with $His_6$-tagged NPM2 keeping NPM1 as the positive control for the experiments (Swaminathan et al. 2005). The activity was measured by densitometric analysis of the intensities of the bands corresponding to $G_5$ML transcript normalized to that of the ML200 control transcript. Results from multiple experiments showed that NPM2 did not have a significantly positive effect on the transcription of the template at the doses tested, over the basal level (Figure 5.2.4A, compare lanes 5 – 7 to 4) while there was at least 2-fold activation of transcription in the presence of the positive control NPM1 (Figure 5.2.4, compare lanes 8 and 4). The transcription activation potential of NPM2 and NPM1, quantified using data from at least three independent experiments (Figure 5.2.4B), confirmed this observation.

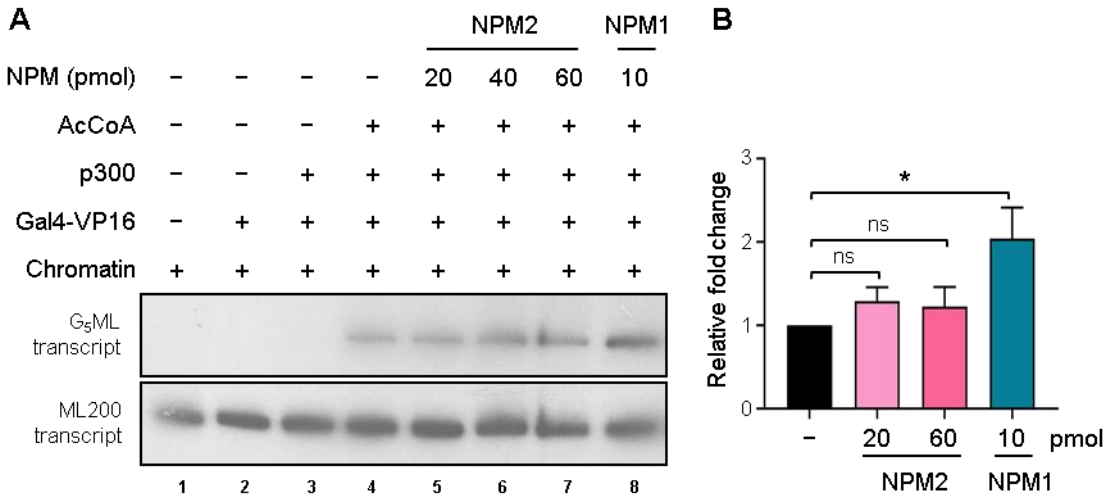
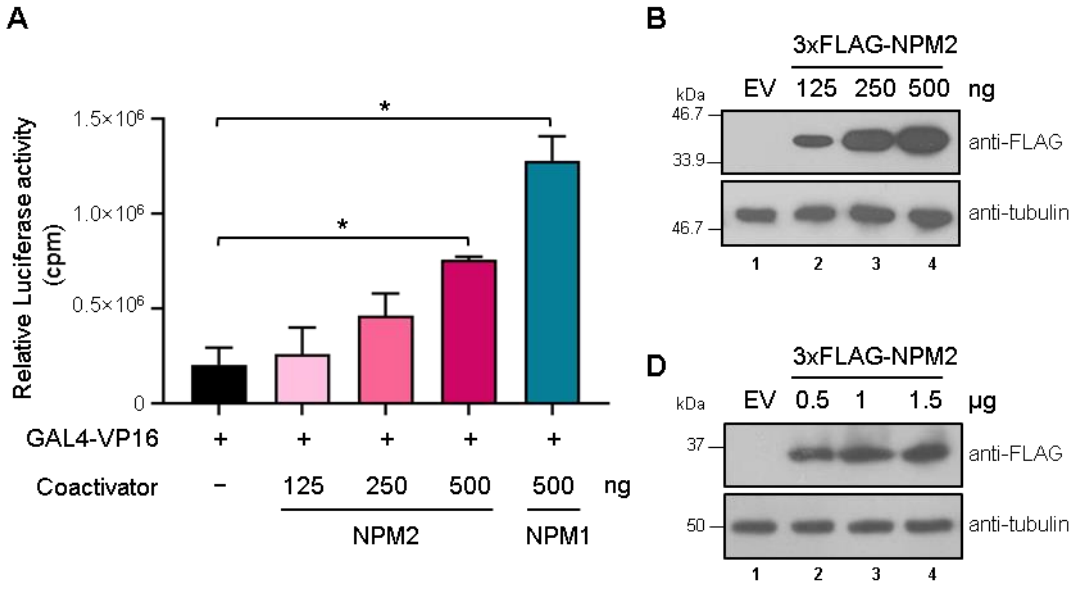



**Figure 5.2.4. Transcription regulation (*in vitro*) potential of NPM2:** (A) A representative result showing the effect of recombinant $His_6$-tagged NPM2 on the transcription of a chromatinized template in the *in vitro* transcription assay. Chromatin (28 ng) was freshly assembled using the NAP1 assembly system. Lanes 1 – 3 correspond to the samples lacking the specific components used in the assay as indicated. Lane 4 (with all the basic components needed for transcription and lacking the remodeler NPM) serves as the basal level. NPM2 was added in increasing doses (20, 40 and 60 pmol, lanes 5 – 7) to test its effect on chromatin transcription, while 10 pmol of NPM1 was used as a positive control for the assay (lane 8). The upper panel shows transcription of the $G_5ML$ template and the bottom panel shows the transcription of ML200 control (loading) template. (B) Bars represent the relative transcription per lane (in fold activation over the basal acetylation-dependent transcription lane (lane 4) as determined by densitometric analysis. Values are mean + SEM from at least three independent experiments. Statistical analysis was performed using Student's *t*-test. $*P < 0.05$, ns: non-significant.

These results show that recombinant $His_6$-tagged NPM2 does not exhibit an intrinsic ability to activate transcription under the present experimental conditions. Possibly the low inherent histone chaperone activity of $His_6$-tagged NPM2 could be a determining factor in its ability to activate transcription. On a parallel note, we can also assume that NPM2's function as a histone chaperone and potential transcription regulator is better manifested when it is post-translationally modified and/or associated with other factors in its native form. To test this hypothesis, we performed an activator (Gal4-VP16)-dependent Luciferase reporter assay (Section 2.4.2) after overexpressing FLAG-tagged NPM2 in HEK-293 cells. We observed that in the presence of increasing doses of NPM2 (Figure 5.2.5B), there was a moderate but dose-dependent enhancement in the expression of the Luciferase reporter gene (Figure 5.2.5A). This enhancement of the reporter expression by 500 ng of NPM2 was not obtained to similar levels as that of the positive control NPM1 (Swaminathan et al. 2005) of the same dose, which suggests that the transcription

regulatory function of NPM2 is either not as efficient as NPM1 or it is not manifested in a true sense in a system where it is not naturally present. To test the latter point, we overexpressed FLAG-tagged NPM2 in HEK-293 cells and checked the mRNA expressions of a few candidate genes by RT-qPCR analysis. We selected few genes related to pluripotency pathway (c-myc and *NANOG*) where NPM2 was found to play a role (Huynh et al. 2016), few ubiquitously expressed (*NPM1*, *GAPDH* and *18S* rRNA) and signal-dependent (c-fos and *NR4A1*) genes, and checked their transcript levels after transfection of HEK-293 cells with increasing doses of NPM2. We found that NPM2 overexpression in HEK-293 cells did not lead to a significant alteration in the expressions of these genes (Figure 5.2.5C). This suggests that the regulation of expressions of these genes in the cell line used here, might not be mediated by NPM2. In other words, possibly NPM2 requires its natural environment to manifest its function as a potential transcriptional regulator, presumably in its native post-translationally modified form or in a complex with other proteins.



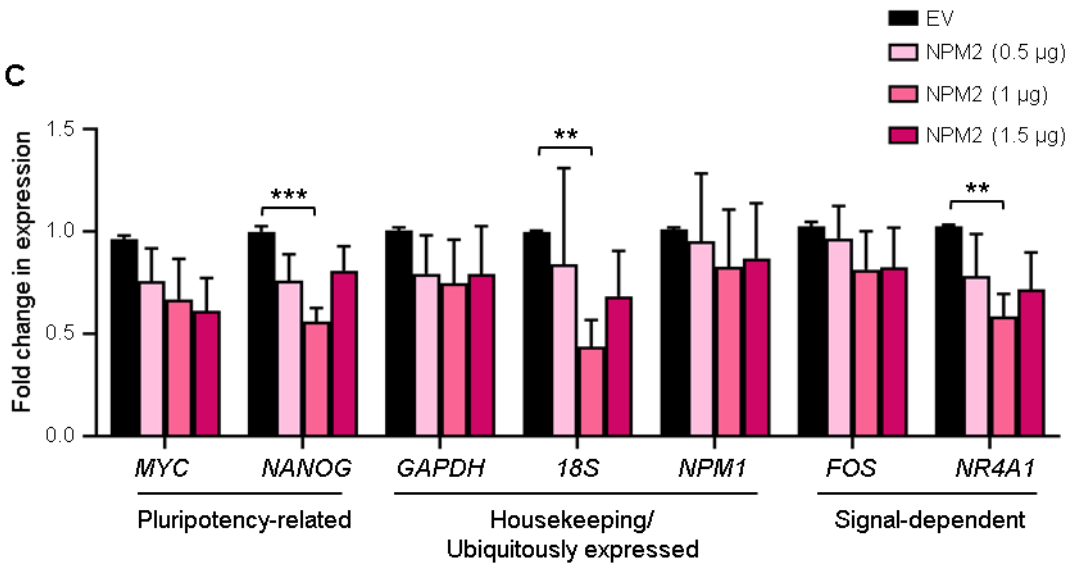



**Figure 5.2.5. Transcription regulation (*ex vivo*) potential of NPM2:** (A) Bars represent relative Luciferase activity after transfection of HEK-293 cells with different doses of expression constructs of NPM2 and NPM1 as indicated along with 100 ng of reporter plasmid $pG_{10}luc$, 100 ng of internal control plasmid pCMV-LacZ, and 10 ng of activator Gal4-VP10 plasmid for 24 h. The total amount of the plasmid transfected was the same in all the samples which were made up with the empty vector plasmid. Data were normalized to internal transfection control β-galactosidase. Values are mean + SEM from two independent experiments with two technical replicates per experiment. Statistical significance was calculated using Student's *t*-test. $*P < 0.05$. (B) Western blot analysis showing dose-dependent expression of FLAG-tagged NPM2 after transfection of different amounts of the plasmid as indicated for the experiment described in (A). The upper panel shows western blot with anti-FLAG and the bottom panel shows western blot with anti-tubulin antibody. (C) Bars represent fold change in expression of indicated genes measured by RT-qPCR analysis after transfection of HEK-293 cells with increasing doses of FLAG-tagged NPM2 or the empty vector (EV) as indicated, for 24 h. Values are mean + SEM from four independent experiments. Statistical significance was calculated using Student's *t*-test. $**P < 0.01$, $***P < 0.001$. The non-significant outcomes have not been specifically indicated to avoid cluttering. (D) Western blot analysis showing expression of FLAG-tagged NPM2 after transfection of different amounts of the plasmid as indicated for the experiment described in (C). The upper panel shows western blot with anti-FLAG and bottom panel shows western blot with anti-tubulin antibody.

### *5.2.3. Post-translational modifications (acetylation, methylation and phosphorylation) of NPM2*

Based on our results presented so far, we hypothesized that the activity of NPM2 as a histone chaperone and potential transcription regulator could be impacted by its post-translational modifications. To date, there have been few studies related to the post-translational modifications and their effect on Nucleoplasmin function. *Xenopus* Nucleoplasmin (NP) has been found to be majorly phosphorylated. Phosphorylation is higher in the laid eggs compared to the oocytes and it occurs in multiple sites throughout

the protein which has been found to increase the activity of Nucleoplasmin (Cotten et al. 1986; Leno et al. 1996; Banuelos et al. 2007). One of the kinases phosphorylating NP is Casein Kinase II (Taylor et al. 1987; Vancurova et al. 1995). Human NPM2 also gets phosphorylated, more efficiently by mitotic kinases (Okuwaki et al. 2012). However, the specific kinase enzymes are not reported. Nonetheless, the protein harbors consensus sites for several cell cycle-dependent kinases.

Among other PTMs that have been reported, NP has been shown to get symmetrically di-methylated at Arg187 by PRMT5-Mep50 complex, which modulates the enzyme's activity towards histones, indicative of a regulatory role of NPM2 *in vivo*. This methylated NPM2 is physiologically detected late in oogenesis and is believed to contribute towards the development of pluripotency (Wilczek et al. 2011). In a relatively recent report, NP has been shown to get glutamylated as well, other than the previously reported PTMs of phosphorylation and arginine methylation. These PTMs are developmentally regulated and they greatly impact the conformation of NP causing it to either deposit the histones onto DNA (mediated by C-terminal tail phosphorylation, PRMT5-mediated arginine methylation and TTLL4-mediated glutamylation of Nucleoplasmin in the oocytes) or sequester them (mediated by hyperphosphorylated Nucleoplasmin in the eggs) (Onikubo et al. 2015).

To test the role of post-translational modifications (PTMs) of NPM2 on its activity, we screened for a few PTMs such as acetylation, methylation, and phosphorylation with specific enzymes by *in vitro* assays. We normalized the activity of common lysine acetyltransferases (KATs) such as p300, PCAF, Tip60 and HAT domain (HD) of MOZ purified from Sf21 cells infected with the respective baculoviruses, by the *in vitro* filter-binding assay using rat liver core histones as the substrate (Section 2.11.3) (Figure 5.2.6A). Using these enzymes, we performed a gel assay (Section 2.11.2) to check if NPM2 could be acetylated *in vitro* by any of these enzymes. We performed parallel acetylation reactions using core histones purified from rat liver tissue as the positive control substrates for acetylation. Results showed that NPM2 was not acetylated by any of the KATs used in the assay (Figure 5.2.6B, upper panel) while the positive control core histones could be acetylated (Figure 5.2.6B, third panel from the top). We repeated the assay using a non-histone positive control substrate, NPM1, of the most common KAT that is p300 (Swaminathan et al. 2005). We found that the acetylation of NPM2 (Figure 5.2.6C, upper panel, lane 3) was very less compared to the signal corresponding to NPM1 acetylation

(Figure 5.2.6C, upper panel, lane 4). This indicated that NPM2 was not a substrate for acetylation by p300.

Since Nucleoplasmin is known to be a substrate for the protein arginine methyltransferase complex PRMT5-Mep50, we were interested to test if human NPM2 could be methylated by another arginine methyltransferase that is coactivator associated arginine methyltransferase 1 (CARM1) or PRMT4. We performed an *in vitro* methylation assay (Section 2.11.4) using CARM1 and NPM2 or histone H3 (as the positive control). The results showed that NPM2 was not arginine-methylated by CARM1 (Figure 5.2.6D, upper panel) while H3 was readily methylated in the assay (Figure 5.2.6D, third panel from the top).

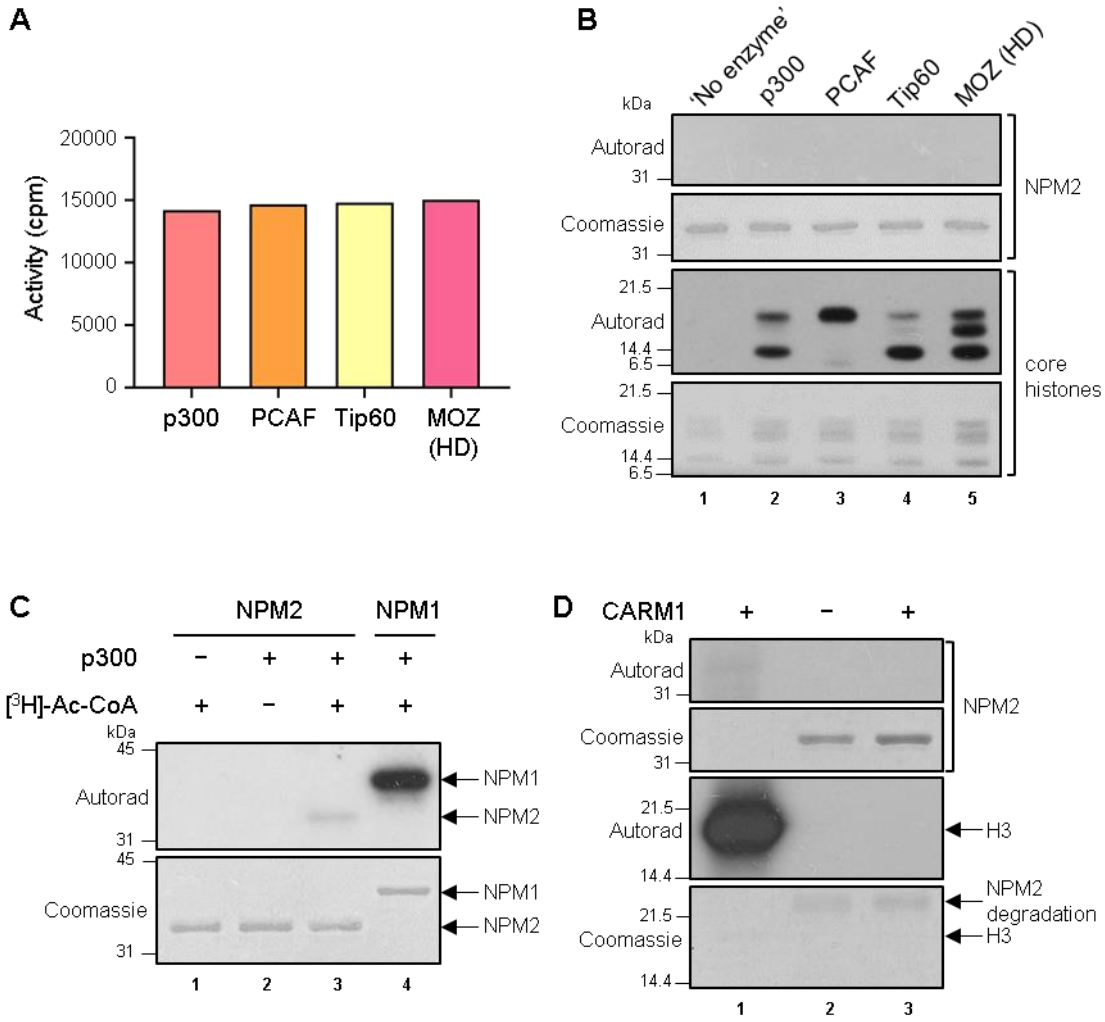



**Figure 5.2.6. Post-translational modifications (acetylation and methylation) of NPM2:** (A) Bars represent the normalized activity of the specific lysine acetyltransferases (KATs) as indicated. cpm: counts per min, HD: HAT domain. (B) *In vitro* acetylation (gel) assay using normalized amounts (as shown in (A)) of the specific KATs as indicated, and 1 µg each of NPM2 (upper panel) and core

histones (positive control, third panel from the top) as the substrates. The second panel from the top and the bottom panel show the Coomassie-stained images of NPM2 and core histones respectively denoting the loading controls. (C) *In vitro* acetylation (gel) assay using the specific KAT p300, and 1 µg each of NPM2 and NPM1 (positive control) as the substrates. The bottom panel shows the Coomassie-stained image of the substrated denoting the loading controls. (D) *In vitro* methylation (gel) assay using the specific arginine methyltransferase enzyme CARM1, and 1 µg of NPM2 and 100 ng of histone H3 (positive control) as the substrates. The second panel from the top and the bottom panel show the Coomassie-stained images of the substrates as indicated, denoting the loading controls. The amount of H3 (100 ng) is not distinctly visible by Coomassie staining.

Since Nucleoplasmin is known to be phosphorylated both *in vitro* and *in vivo* (Cotten et al. 1986; Leno et al. 1996; Banuelos et al. 2007); (Taylor et al. 1987; Vancurova et al. 1995), and human NPM2 could be phosphorylated *in vitro* by mitotic kinases (Okuwaki et al. 2012), we wished to investigate if NPM2 was a substrate of the mitotic kinases Aurora Kinase A and B. We performed an *in vitro* phosphorylation (gel) assay (Section 2.11.1) using these kinases purified from baculovirus-infected Sf21 cells and NPM2. We used commercially procured rat liver Casein Kinase II as a positive control kinase for NPM2 in the assay. The results showed that NPM2 could be phosphorylated by all these three kinases *in vitro* (Figure 5.2.7A). We repeated the assay using increasing amounts of the enzymes Aurora Kinase A and B and found that there was a dose-dependent increase in the signal corresponding to phosphorylated NPM2 in the presence of Aurora Kinase A and B (Figure 5.2.7B). This showed that NPM2 was a substrate of Aurora Kinase A and B *in vitro*.

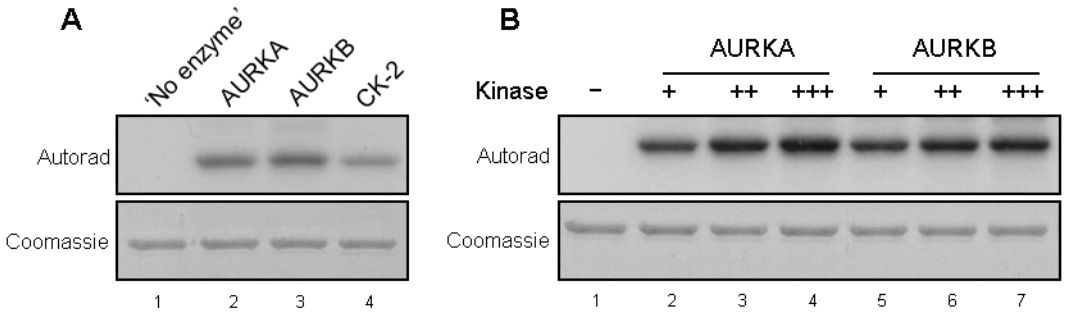



**Figure 5.2.7. *In vitro* phosphorylation of NPM2:** (A – B) *In vitro* phosphorylation (gel) assay using specific kinases as indicated and 1 µg of NPM2 as the substrate. Activites of the kinases have not been normalized and hence the indicated results are qualitative. Lower panels show the Coomassie-stained images of the gels indicating the loading control. AURKA: human Aurora Kinase A, AURKB: human Aurora Kinase B, CK-2: rat liver Casein Kinase II.

### *5.2.4. Effect of Aurora Kinase-mediated phosphorylation of NPM2 on its histone chaperone activity*

Phosphorylation of Nucleoplasmin (NP) has been shown to enhance its activity in terms of sperm chromatin decondensation (Leno et al. 1996; Okuwaki et al. 2012). Hyperphosphorylated NP isolated from *Xenopus* eggs could efficiently unfold sperm and somatic chromatin by removing the chromosomal proteins from linker DNA regions (Ramos et al. 2005). Endogenous hyperphosphorylated, as well as artificial phospho-mimicking mutant forms of Nucleoplasmin, are more active towards sperm chromatin decondensation compared to the unphosphorylated or hypophosphorylated forms (Hierro et al. 2002; Banuelos et al. 2007; Okuwaki et al. 2012). Phosphorylation of Nucleoplasmin at Tyr 124 was also found to mediate apoptotic chromatin condensation in *Xenopus* egg extracts (Lu et al. 2005). NPM2 phosphorylated in the presence of mitotic cell extracts, and phospho-mimic NPM2 were also found to have reduced non-specific binding to DNA and enhance nucleosome core particle (NCP) assembly compared to unmodified recombinant NPM2 (Okuwaki et al. 2012). However, the role of specific kinase-mediated phosphorylation of human NPM2 on its histone chaperone activity in terms of nucleosome assembly has not been studied so far.

We were interested in investigating this angle. For this, we mass phosphorylated recombinant NPM2 with Aurora Kinase A or B (Section 2.11.1) and checked its histone chaperone activity by the supercoiling assay. As controls for the assay, we kept unphosphorylated NPM2 (which was not subjected to the process of mass phosphorylation), mock phosphorylated NPM2 (which was subjected to the process of mass phosphorylation in the absence of an enzyme), and recombinant NPM1 as the positive control for the assay. We found that there was a modest but statistically significant increase in the histone chaperone activity of NPM2 upon phosphorylation by Aurora Kinase A and compared to the unphosphorylated or mock phosphorylated NPM2 (Figure 5.2.8A, compare lanes 8 – 11 with 4 – 7, and C) at the normalized doses tested (5 and 10 pmol, Figure 5.2.8B). This shows that Aurora Kinase-mediated phosphorylation of NPM2 could have a positive effect on its histone binding and nucleosome assembly activity, which needs to be further validated.

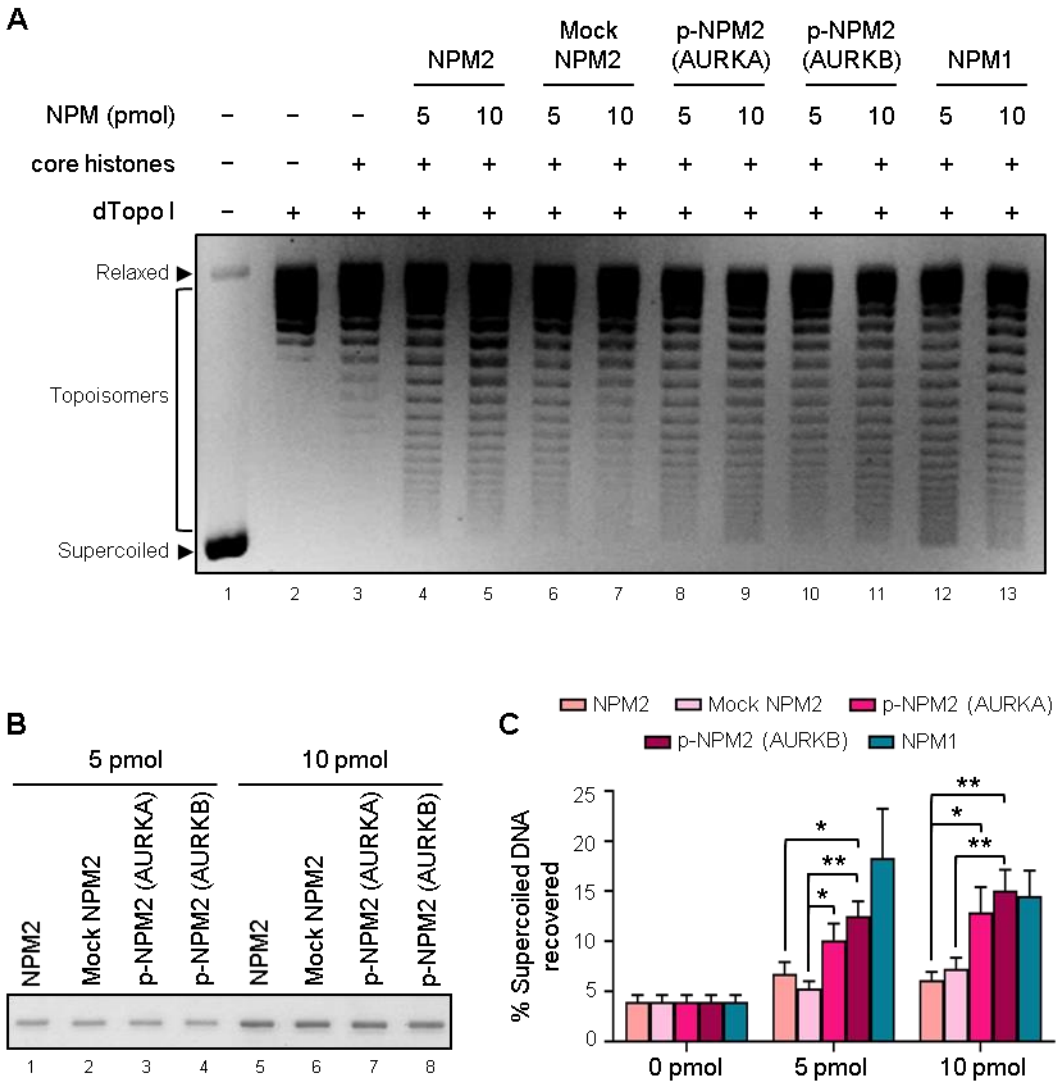



**Figure 5.2.7. *In vitro* phosphorylation of NPM2:** (A) Histone transfer or plasmid supercoiling assay with increasing doses (5 and 10 pmol) of unphosphorylated $His_6$-NPM2 (NPM2, lanes 4 – 5), mock phosphorylated (without enzyme) $His_6$-NPM2 (Mock NPM2, lanes 6 – 7), $His_6$-NPM2 mass phosphorylated by human Aurora Kinase A (p-NPM2 (AURKA), lanes 8 – 9), $His_6$-NPM2 mass phosphorylated by human Aurora Kinase B (p-NPM2 (AURKB), lanes 10 – 11), and positive control NPM1-$His_6$ (lanes 12 – 13). Lane 1 (Input) indicates the reference position of the relaxed and supercoiled form of the plasmid. Lane 2 (negative control) indicates the sample without core histones. Lane 3 indicates the sample with core histones only (without any chaperone) that marks the basal level of supercoiled DNA generated in the absence of any chaperone in the assay. (B) Gel profile showing normalized amounts of the different NPM2 samples (5 and 10 pmol) used in the assay depicted in (A) as indicated. (C) Bars represent histone chaperone activity of indicated doses of NPM2 samples and NPM1 quantified in terms of percentage supercoiled DNA recovered with reference to Input (lane 1). Values are mean + SEM from 6 independent experiments. Pairwise statistical analysis was performed using Student's *t*-test. **P* < 0.05, ***P* < 0.01.

### *5.2.5. Identification of sites of NPM2 phosphorylated by Aurora Kinase A and B in vitro*

To further probe the role of Aurora Kinase-mediated phosphorylation of NPM2, we wanted to identify the sites of NPM2 phosphorylated by the Aurora Kinase A and B. We performed an *in silico* generic prediction analysis for putative sites of phosphorylation in the NPM2 protein based on consensus target sequences of well-characterized kinases using the tool NetPhos 3.1 (Blom et al. 1999). This program predicted 31 putative phosphorylated sites of NPM2 including Ser, Thr and Tyr residues (Figure 5.2.8A), based on the consensus target sequences of the kinases namely ATM, CKI, CKII, CaMKII, DNAPK, EGFR, GSK3, INSR, PKA, PKB, PKC, PKG, RSK, SRC, cdc2, cdk5 and p38MAPK. Parallelly, we identified two residues in NPM2 by mass spectrometry analyses namely S174 and S196 as putative sites phosphorylated by the Aurora Kinases (Section 2.12.2). S174 was found to be phosphorylated by both Aurora Kinase A and B, while S196 was found to be phosphorylated by Aurora Kinase A *in vitro*. S196 was found to be conserved across most of the model organisms (Figure 5.2.8B) indicating that its phosphorylation could have evolutionary significance. Its role as a putative phosphorylated site of NPM2 was previously studied where it was found that its phosphomimetic mutation to Asp in combination with six other putative residues could reduce NPM2's non-specific DNA binding activity and increase its nucleosome core particle (NCP) assembly activity (Okuwaki et al. 2012). Similar phospho-mimic mutant of NPM2 where Ser 196 along with nine other residues were mutated to Asp, was found to enhance induced human pluripotent stem cell generation in the presence of oocyte enriched histone variants TH2A and TH2B, and the pluripotency factors OCT4, SOX2, KLF4 and c-MYC (OSKM) over and above the OSKM factors alone (Huynh et al. 2016). These studies indicated that phosphorylation at S196 could have functional significance, yet the kinases responsible for phosphorylating this site are unknown. To confirm if these sites were phosphorylated by the Aurora Kinase enzymes, we generated phospho-deficient mutants of NPM2 where these Ser residues were mutated to Ala and performed the *in vitro* kinase assays with Aurora Kinase A or B. While we did not find a significant alteration in the phosphorylations of the S174A mutant compared WT NPM2, both by Aurora Kinase A and B (data not shown), we observed that S196A mutant was significantly less phosphorylated by Aurora Kinase A (Figure 5.2.8C) and Aurora Kinase B (Figure 5.2.8D) compared to WT NPM2. The results from two independent experiments using different batches of the purified substrate and enzyme

proteins have been quantified and found to be statistically significant (Figure 5.2.8E). This confirms that S196 was a major site of phosphorylation by Aurora Kinase A and B *in vitro*.

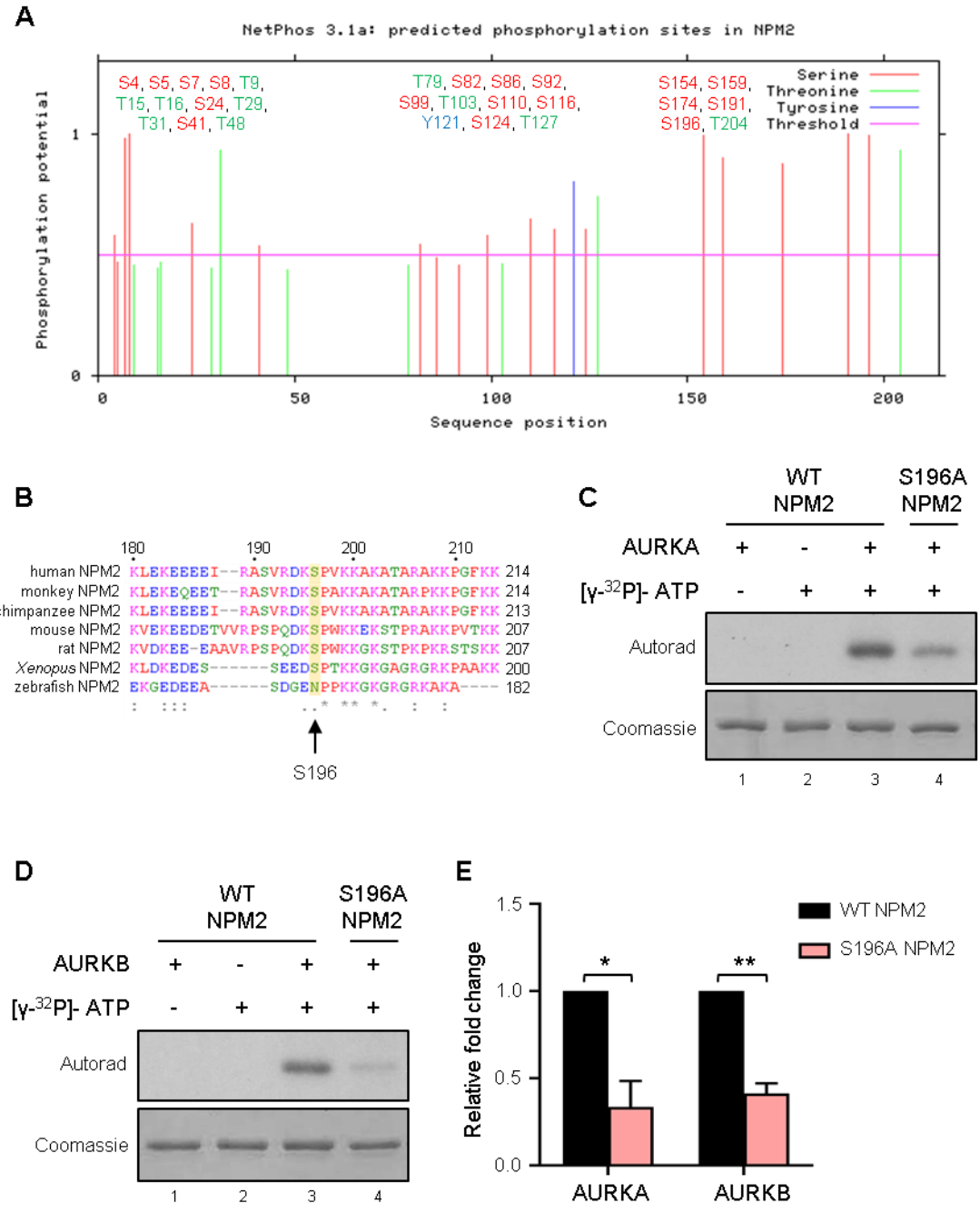



**Figure 5.2.8. Identification of sites of NPM2 phosphorylated by Aurora Kinase A and B *in vitro*:** (A) Graphical representation of the generic prediction of the putative phosphorylated sites of NPM2 generated using the NetPhos 3.1 tool. The Ser, Thr, and Tyr residues are denoted in red, green, and blue respectively. The length of the lines denotes the output score (value in the range 0.000 – 1.000) for each of the predicted sites. The threshold score is 0.500, denoted in pink. Note that S196 was among the top-scoring residues for phosphorylation. (B) Multiple sequence alignment of the NPM2 protein sequence generated using the Clustal Omega (1.2.4) tool (Sievers et al. 2011) showing the high conservation of S196 (highlighted in yellow) across the different species. (C – D) *In vitro* phosphorylation (gel) assay using (C) Aurora Kinase A (AURKA) and (D)

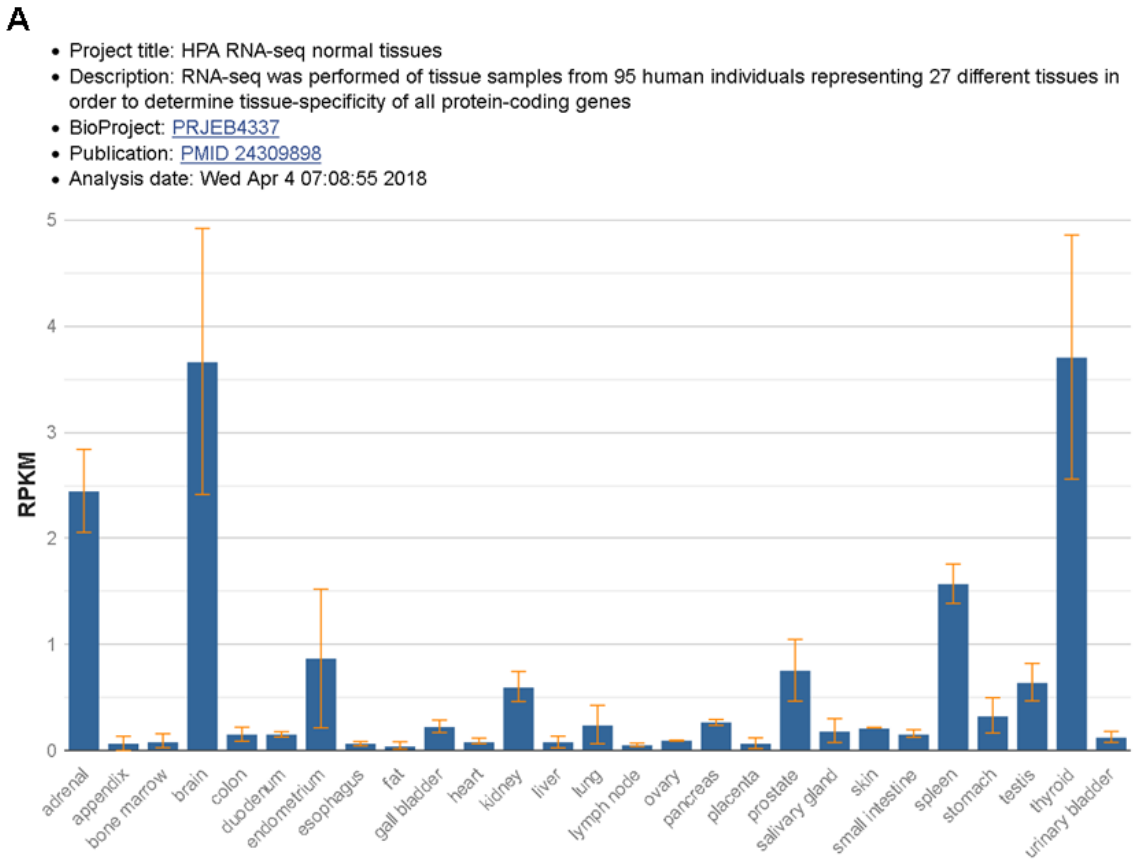
Aurora Kinase B (AURKB) and 1 µg of NPM2 WT or S196A as the substrate. Lower panels show the Coomassie-stained images of the gels indicating the loading control. (E) Bars represent relative fold changes in the phosphorylation signal of NPM2 (WT or S196A) in the presence of AURKA or AURKB. Values are mean + SEM from two biological replicates. Statistical analysis was performed using Student's *t*-test. **P* < 0.05, ***P* < 0.01.

The functional role of Aurora Kinase-mediated phosphorylation of NPM2 needs to be further investigated. Mass phosphorylation of NPM2 by Aurora Kinase resulted in a significant increase in its histone chaperone activity determined by the *in vitro* histone transfer or plasmid supercoiling assay. It needs to be validated if the phospho-mimic mutant of NPM2 (S196D or S196E) can replicate this result and its effect on the regulation of chromatin transcription *in vitro*. As reported previously, Nucleoplasmin is an oocyte-specific protein, hence, it could also be a substrate of Aurora Kinase C, which is expressed in the germ cells in mammals. Multiple studies have shown that the Aurora Kinase enzymes play important roles in oocyte development (Nguyen and Schindler 2017) and induced pluripotency (Lee et al. 2012a), where the functions of NPM2 and its orthologs have also been implicated. It would be interesting to study if Aurora Kinase-mediated phosphorylation of NPM2 has functional relevance in the context of oocyte development and induced pluripotency.

### *5.2.6. Expression analysis of mammalian Nucleoplasmin in normal (non-cancerous) cells*

To study the phosphorylation of NPM2 by Aurora Kinase and its function in a cellular system, we took cues from our observations and the reported literature and searched for a suitable cellular model to test our hypotheses made so far. Nucleoplasmin and its orthologs have been reported to be expressed in the eggs and oocytes in non-human model species such as the mouse (Burns et al. 2003), zebrafish (Bouleau et al. 2014), *Xenopus* (Litvin and King 1988), and *Bos taurus* (Lingenfelter et al. 2011). In humans, such strict tissue-specificity for NPM2 transcript expression appears to be absent. Transcript expression data compiled from RNA-seq performed of normal tissue samples from 95 human individuals representing 27 different tissues as depicted in the NCBI Gene database, shows that NPM2 mRNA is expressed in various tissues with high levels in the adrenal gland, brain, and thyroid tissue (Figure 5.2.9A). Since our study has been focussed on the functional

characterization of mammalian Nucleoplasmin, we wished to study the *in vivo* role of NPM2 in the animal model which was derived from human cells or is close to humans, that is the mouse. We first did an expression analysis of Npm2 and its paralog Npm1 in normal tissues from different organs of male and female adult mice. Semi-quantitative PCR (data not shown) and RT-qPCR data compiled from 4 male and 3 female mice belonging to BALB/c, CD-1 and outbred wild-type strains showed that Npm2 mRNA expression was almost negligible in most of the organ tissues tested except for ovary and testis where it was expressed significantly (Figure 5.2.9B). Npm1, as expected showed ubiquitous expression in these tissues (Figure 5.2.9C). The expression of Npm2 in mouse testis was not observed in a previous study (Burns et al. 2003), possibly because its lesser expression compared to that in the ovary, could not be detected by Northern blotting analysis performed in that study. When we checked the expression data from the mouse ENCODE project (Yue et al. 2014) as depicted in the NCBI Gene database (data not shown) or that retrieved from the WashU EpiGenome Browser (Figure 5.2.9D), we indeed found that Npm2 transcript expression was high in normal adult mouse ovary as well as testis tissues.



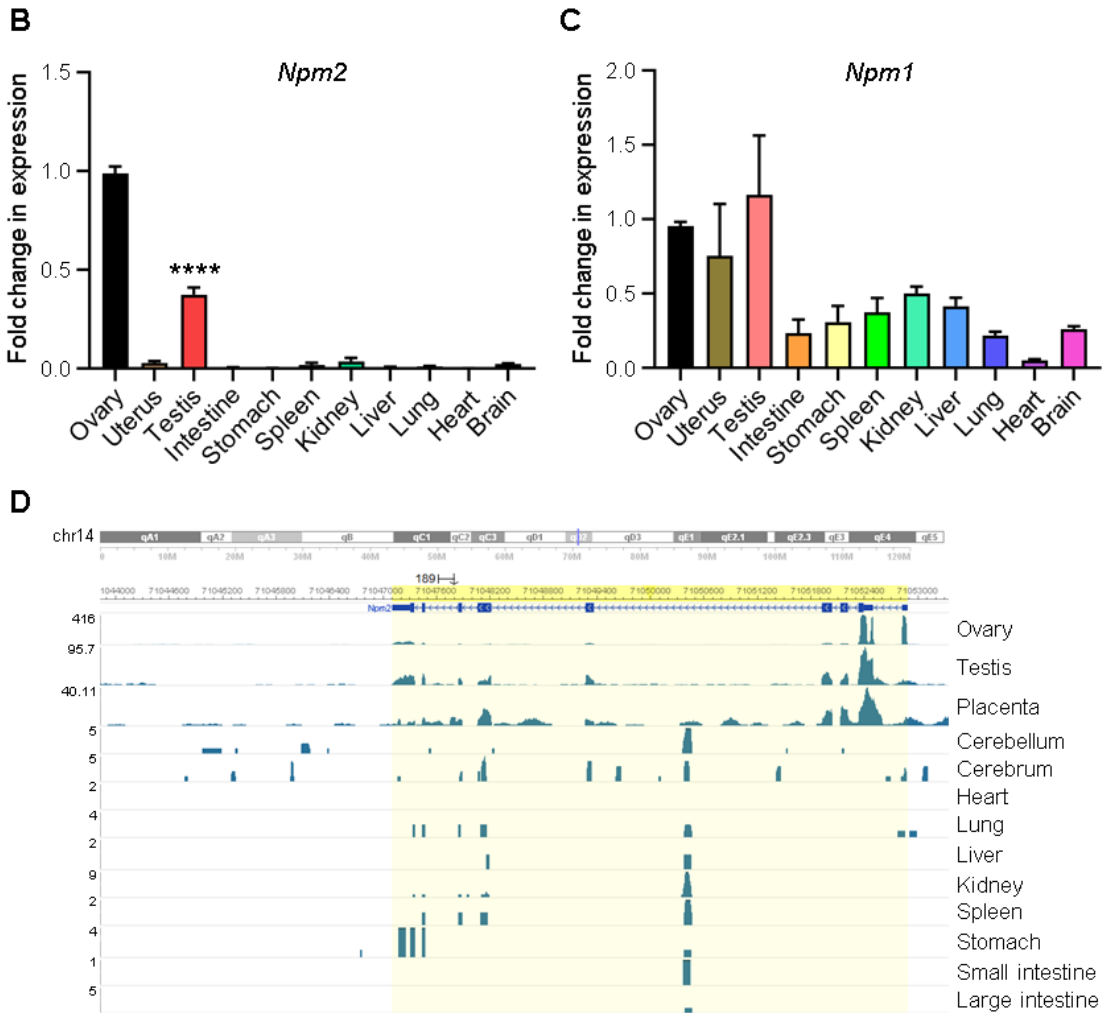



**Figure 5.2.9. Identification of sites of NPM2 phosphorylated by Aurora Kinase A and B *in vitro*:** (A) Screenshot from NCBI Gene database (https://www.ncbi.nlm.nih.gov/gene) showing the mRNA expression of human NPM2 in various normal tissues as indicated and described. (B – C) Bars represent fold change in expression of mRNA levels of (B) Npm2 and (C) Npm1 in different organ tissues of normal adult mice as indicated. Values are mean + SEM from 4 male and 3 female mice. (B) Statistical analysis between any two samples was done using Student's *t*-test. ****$P < 0.0001$. (D) WashU Epigenome Browser (v46.2) screenshot showing the mRNA expression levels of Npm2 (gene body highlighted in yellow) in various tissues from normal adult mice as indicated and retrieved from the mouse ENCODE datasets. Note the scale of the individual tracks showing that the Npm2 transcript is expressed highest in the ovary, followed by testis and placenta, while almost negligible in the other organ tissues tested here. The lower value of the scale of every track is zero.

The presence of Npm2 protein in the mouse testis is yet to be confirmed. Initial attempts to detect the protein by western blotting and immunohistochemistry (IHC) analysis have not yielded conclusive results probably due to lower sensitivity of the anti-NPM2 antibody towards mouse Npm2, hence confirmatory analysis using mass spectrometry needs to be done. The expression of Npm2 in mouse testis would provide a convenient model to further

explore the biological functions of this histone chaperone, especially in the process of male germ cell development, which could be conducted *in vitro*. Although it is reported that in zebrafish, Nucleoplasmin is expressed only in the ovary, the scenario is slightly different in other cyprinid fish models such as *Carassius auratus gibelio* and *Carassius auratus*. The characterization of the Nucleoplasmin/Nucleophosmin homolog in these fish revealed that the homolog was expressed in both the gonads and specifically in the oogenic and spermatogenic cells (Wu et al. 2009). On a similar note, a very early study reported more than three decades ago, showed that a polyclonal anti-serum against Nucleoplasmin (Krohne and Franke 1980) (that cross-reacted with its other paralogs, presumably the approximately 40 kDa protein NO38 or Nucleophosmin), detected faint bands corresponding to the 30 kDa protein that is Nucleoplasmin, in whole-tissue lysates from adult *Xenopus* heart and testis, by western blotting (Litvin and King 1988), which was undetected using anti-Nucleoplasmin monoclonal antibody. Using this monoclonal antibody, it was found that Nucleoplasmin was localized in the nuclei of post-mitotic cells in the brain, retina, spinal cord and striated muscle (but not in the smooth muscle of the gut) in the tadpole stage (stage 49) of *Xenopus* but not in the adult tissues namely brain, spinal cord, eye, skin, lung, liver and gut (Litvin and King 1988). These facts indicate that the expression of Nucleoplasmin might not be strictly oocyte-specific as commonly understood, in all the species at all developmental stages and there could be a possibility of undetectably lower expression of this protein in non-germ cell types as well.

An important tissue type that has not been precisely studied in any organism for the expression analysis of Nucleoplasmin, is the muscle tissue. Skeletal myogenesis, that is the process of development of the skeletal muscle tissue, has been an interesting model to study the underlying molecular regulatory mechanisms which dictate the development process. Transcription factors known as myogenic regulatory factors (MRFs) are the key players in setting up the different transcriptional programs during the differentiation process. These are MyoD, Myf5, Myf6, and Myogenin (Myog). Besides these, other proteins belonging to the classes of myocyte enhancer factors (MEFs) and serum response factors (SRFs) also play roles in the myogenesis process (Comai and Tajbakhsh 2014). Besides transcription factors, chromatin remodelers, chromatin modifiers, histone chaperones, non-coding RNAs and other factors are also known to play important roles in this highly-coordinated process (Braun and Gautel 2011; Robinson and Dilworth 2018; Sartorelli and Puri 2018). The role of Nucleoplasmin and its homologs in this process have not been studied much.

Nucleophosmin was found to be a negative regulator of skeletal myogenesis process. The destabilization of Npm1 mRNA by the HuR/KSRP complex, resulting in reduced levels of Npm1 protein, was shown to be required for muscle fiber formation in the mouse C2C12 cell line system (Cammas et al. 2014). Further, in the human vascular smooth muscle system, it was found that full-length NPM1 levels were reduced by Granzyme B-mediated proteolytic cleavage in the differentiated aortic smooth muscle cells (Ulanet et al. 2004). The expression status of Nucleoplasmin in muscle tissue is not clearly known except for a few indirect studies which showed that Nucleoplasmin expression was detected in the striated muscle tissue of the tadpole and heart muscle tissue of the adult *Xenopus laevis* (Litvin and King 1988), as well as in zebrafish and *Xenopus tropicalis* muscle tissues (Cheung et al. 2018), albeit at a very low level. However, the expression of Nucleoplasmin in the course of muscle cell differentiation and its signification is not known.

We were interested to study the expression pattern of Npm2 in the course of differentiation of the mouse myoblast cell line C2C12. These cells proliferate in the form of spindle-shaped myoblasts in the high serum-containing medium which can be induced to differentiate into elongated myotubes under low serum conditions (Figure 5.2.10A). This provides a convenient *in vitro* model to study the molecular pathways operating in the process of myogenesis. We cultured the C2C12 myoblasts in the high serum-containing medium until about 70% confluency, after which we changed the medium to reduced serum-containing medium and allowed the cells to differentiate for 4 days. Samples corresponding to 'Day 0' (prior to changing the medium from high serum- to low serum-containing) and 'Day 1' through 'Day 4' (post-medium change from high serum- to low serum-containing) were collected and mRNA expression analysis was carried out. Beside Npm2, we checked the expression of Npm1 and c-fos whose expressions are reported to be downregulated (Thinakaran et al. 1993; Trouche et al. 1993; Trouche et al. 1995; Cammas et al. 2014), MyoD whose expression is known to be mildly upregulated or unchanged, and Myogenin (Myog) whose expression is known to be highly upregulated (Montarras et al. 1991) during myoblast to myotube differentiation in the C2C12 cells, as controls. The changes in the morphology of the C2C12 cells in the course of differentiation were distinctly visible under a bright field microscope (Figure 5.2.10B). Gene expression analysis showed that Npm2 mRNA levels were significantly upregulated as the differentiation progressed (Figure 5.2.10C). The mRNA levels of Myogenin were highly upregulated (Figure 5.2.10D) during this differentiation process and that of MyoD

remained more or less stable and not significantly altered (Figure 5.2.10E), in accordance with the previous studies reported. Contrary to the published reports, we did not observe downregulation in the mRNA expressions of c-fos (Figure 5.2.10F) and Npm1 (Figure 5.2.10G) in the course of myoblast to myotube differentiation, however, the changes in their expressions were not consistent and non-significant across the different biological replicates of the experiment. We checked the status of Npm2 mRNA expression in mouse skeletal muscle tissue, myocytes and C2C12 cells from the publicly available RNA-seq data sets in the mouse ENCODE database and found that low levels of expression of Npm2 have been detected in these three systems (Figure 5.2.11A, top three RNA-seq tracks) which partially validate our experimental observation and indicate that Npm2 might be indeed expressed at low levels in mouse muscle cells. We were interested to find out which factors could be resulting in the expression of Npm2 mRNA in the C2C12 cells. Publicly available ChIP-seq data for various epigenetic marks such as H3K4me3, H3K36me3, H3K27me3, H3ac, H3K79me2, and H3K79me3 showed almost no enrichment in the Npm2 promoter or gene body region (Figure 5.2.11A, scales of the ChIP-seq tracks show that enrichments of these marks are less than 1). DNase-seq performed in the mouse skeletal muscle tissue as reported in mouse ENCODE, showed that most of the Npm2 proximal promoter and gene body region is not sensitive to DNase (data not shown). The lack of activating epigenetic modifications and DNase hypersensitivity sites in the Npm2 promoter region explains the low expression of Npm2. The low expression, however, could be brought about at a homeostatic level by the action of specific transcription factors, presumably the myogenesis-specific factors. ChIP-seq data related to MyoD (Figure 5.2.11A, MYOD1 ChIP-seq track) and Myogenin (Figure 5.2.11A, MYOG ChIP-seq track) from the same database showed negligible enrichments of these factors at the Npm2 promoter which indicate that these MyoD and Myog might not be involved in the regulation of Npm2 gene transcription. We retrieved about 1 kb sequence of the Npm2 promoter DNA (-1000/+250) from the mm10 assembly of the mouse genome and analyzed for binding motifs of transcription factors involved in myogenesis using the Consite online prediction tool. At 80% cut-off, we found several high-scoring binding motifs for the myogenic factors (Myf) which positively regulate myogenesis, as well as c-fos protein which can negatively regulate myogenesis (Figure 5.2.11B). It would be interesting to investigate if these factors have any role in regulating Npm2 expression during myogenesis and its potential implications.

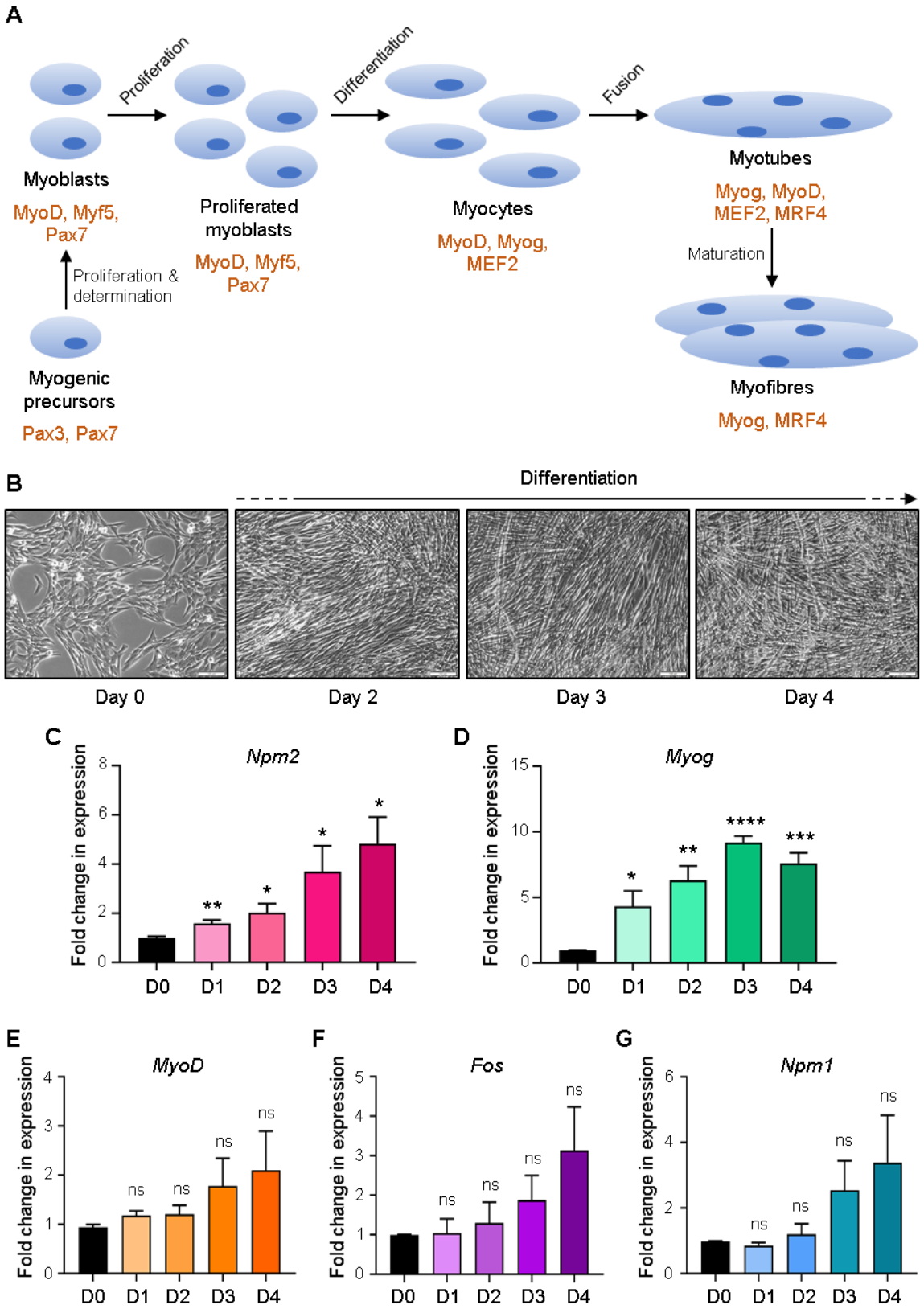



**Figure 5.2.10. Transcriptional changes during skeletal muscle cell differentiation:** (A) Schematic representation of the different stages of skeletal muscle cell differentiation and the expression of marker genes as indicated. Figure adapted from (Beaudry et al. 2016). (B) Photomicrographs showing the morphology of the C2C12 cells growing in high serum-containing medium (Day 0) and those growing in the presence of low serum-containing medium for different

time points (Day 2, Day 3, and Day 4) as indicated. Scale bar is 100 µm. (C – G) Bars represent fold change in expression of the mRNA levels of (C) Npm2, (D) Myogenin (Myog), (E) MyoD, (F) c-fos (Fos), and (G) Npm1 as determined by RT-qPCR analysis after differentiation of the C2C12 cells by growing them in low serum-containing medium for different time points (Day 2 through 4) in comparison to undifferentiated cells grown in high serum-containing medium (Day 0). Values are mean + SEM from four independent experiments. Statistical analysis was performed using Student's *t*-test. **P* < 0.05, ***P* < 0.01, ****P* < 0.001, *****P* < 0.0001, ns: non-significant, D: Day.

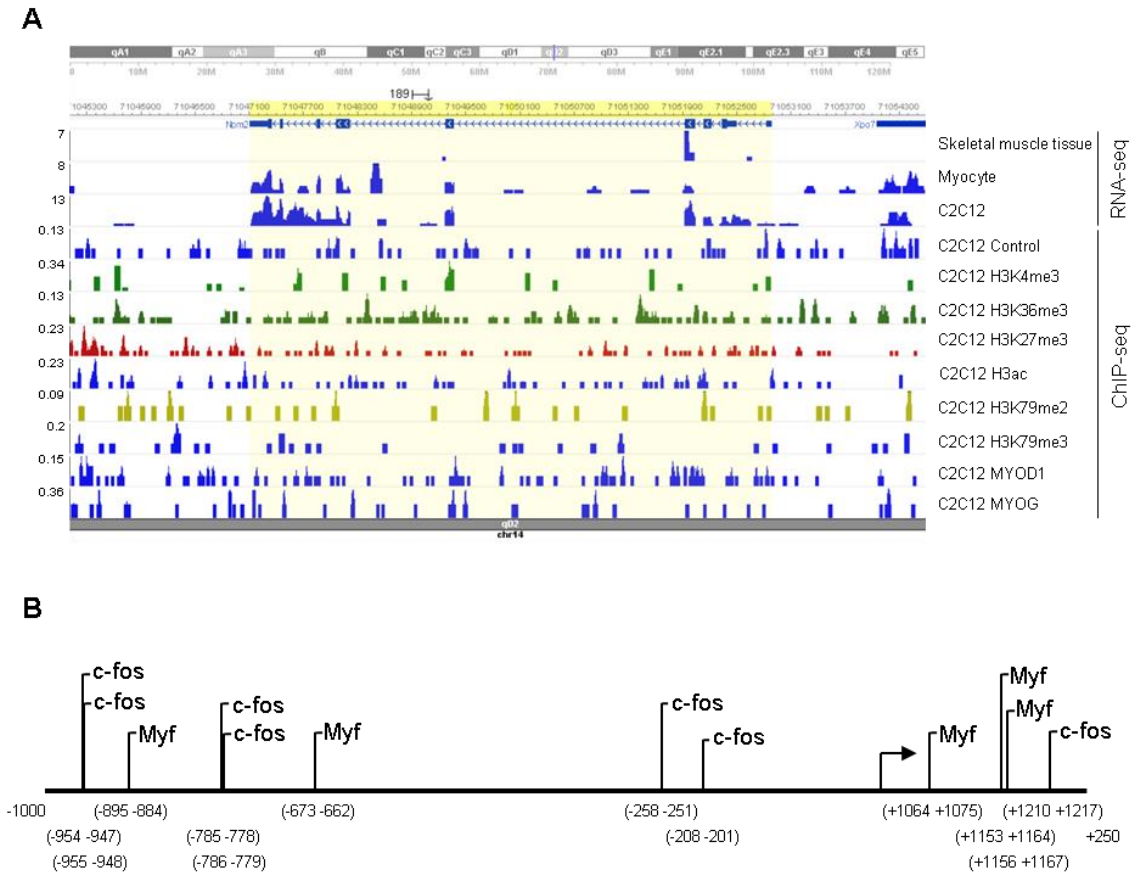



**Figure 5.2.11. Potential regulation of Npm2 transcript expression in mouse muscle cells:** (A) WashU Epigenome Browser (v46.2) screenshot showing the mRNA expression levels of Npm2 (gene highlighted in yellow) in mouse skeletal muscle tissue, myocytes and C2C12 cells determined through RNA-seq analysis, and the enrichment of various epigenetic marks as indicated and transcription factors (MyoD and Myogenin) determined by ChIP-seq analysis in C2C12 cells, retrieved from the mouse ENCODE database. Note the scale of the individual tracks showing that the Npm2 transcript is expressed in low levels in the three systems shown here, enrichment of various epigenetic marks and transcription factors are almost negligible. The lower value of the scale of every track is zero. (B) Schematic representation of the various c-fos- and Myf-binding sites at the mouse Npm2 promoter DNA sequence (-1000/+250), obtained at 80% cut-off using the Consite prediction tool. Numbers indicate the start and endpoints of the DNA sequence used for the analysis and those within parentheses, the start, and endpoints of the transcription factor-binding motifs. The arrow indicates the transcription start site.

The results obtained so far need to be further confirmed, especially at the protein levels. Preliminary western blotting analysis indicates the presence of Npm2 protein in the

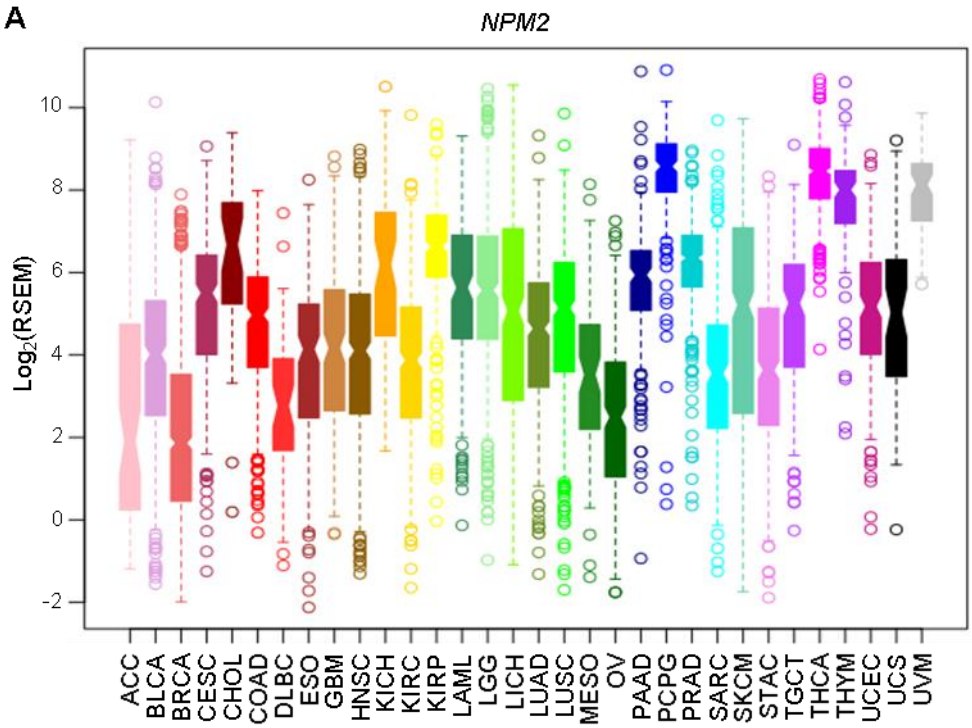
differentiated C2C12 cells (data not shown), however, due to the presence of non-specific reactivity of the anti-NPM2 antibody towards other proteins in the whole-cell lysates of C2C12, the results require confirmation by mass spectrometry approaches. Conclusive analysis of the expression of Npm2 in the model of muscle differentiation would provide significant leads in the study of the role of Nucleoplasmin in cellular differentiation.

### *5.2.7. Expression analysis of NPM2 in cancer cells*

Studies conducted by us and other groups so far have established that Nucleoplasmin is expressed at significantly higher levels in the germ cells. It is often been observed that the expressions of various germ cell-specific proteins such as DDX4 or VASA (Hashimoto et al. 2008; Schudrowitz et al. 2017), MAGE (Fon Tacer et al. 2019), GAGE (Simpson et al. 2005; Gjerstorff and Ditzel 2008), TEX19 (Planells-Palop et al. 2017), Aurora Kinase C (Quartuccio and Schindler 2015), to name a few, are misregulated in cancer (Hofmann et al. 2008; Lifantseva et al. 2011; Bruggeman et al. 2018). An important reason for such deregulated expressions of germ cell-specific proteins in cancer is due to a global alteration of the epigenetic landscape and rewiring of the transcriptional networks in the tumor cells. To date, there has not been much study related to the expression status of Nucleoplasmin in cancer. In a genome-wide study of DNA methylation in melanoma, the NPM2 promoter DNA was found to be hypermethylated compared to normal melanocytes and nevi, which was moderate in early but significantly high in advanced-stage melanomas. The expression of NPM2 mRNA also showed an inverse correlation with its promoter DNA methylation status, that is, NPM2 expression was higher in the normal melanocytes but downregulated in the melanomas, which could be reversed by treatment of the cancer cells with 5-Aza-2'-deoxycytidine (Koga et al. 2009). The downregulated expression of NPM2 in melanoma compared to normal melanocytes and benign nevi was also evident at the protein level (Fujiwara et al. 2018). NPM2 promoter methylation was also found in acute myeloid leukemia, both in cell lines, as well as patient samples, where its methylation was associated with decreased platelet counts at diagnosis and hence, could be used as a diagnostic marker for this cancer (Kroeger et al. 2008). The loci of NPM2 was found to be deleted in colorectal cancer resulting in reduced expression of NPM2 in this cancer (Xi et al. 2016). Further, methylation of the NPM2 promoter region was found to be associated with worse overall survival in neuroblastoma (Decock et al. 2016).

Expression data obtained from databases such as GeneCards (Stelzer et al. 2016) and the Human Protein Atlas (Uhlen et al. 2015) (http://www.proteinatlas.org) indicated that NPM2 was expressed at negligible to low levels in most human cancer cell lines (a moderate level of mRNA expression reported in the liver cancer cell line Hep G2 and the transformed keratinocyte cell line HaCaT), while its mRNA is expressed at least at a basal amount in most normal tissues and the protein is expressed at medium to high levels in some tissues such as respiratory epithelia, endocrine glands, brain, melanocytes, Leydig cells, testis, and oocytes (data not shown). The reports suggest that NPM2 expression might not be strictly germ cell-specific in humans and it might have a negative correlation with cancer.

We carried out bioinformatics analyses using the NPM2 expression and promoter DNA methylation datasets of various human cancers from the TCGA database (TCGA Research Network: https://www.cancer.gov/tcga). The analyses were also done for the NPM1 gene as well for comparison with those of NPM2. Expression analyses showed that the transcript expression of NPM2 (Figure 5.2.12A) was generally much lower than that of NPM1 (Figure 5.2.12B) in all the cancers tested.



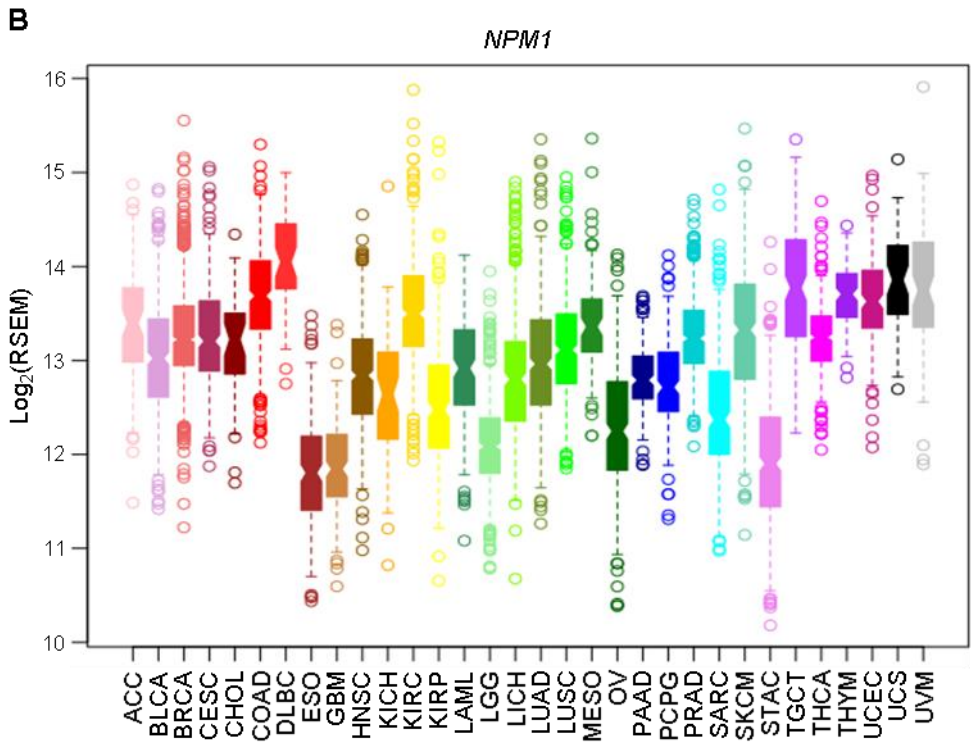



**Figure 5.2.12. Expression analyses of NPM2 and NPM1 transcripts in human cancers:** (A – B) Expression analysis of absolute levels of (A) NPM2 and (B) NPM1 transcripts determined by RNA-seq analyses in patient tissue samples of different cancers as indicated obtained from TCGA database. Details of the abbreviations of different cancer types have been mentioned elsewhere (Appendix, Table A.1). Note the scales on the Y-axis indicating the low expression of NPM2 and higher expression of NPM1 in the various cancers.

To corroborate these reports from the published studies and databases, we carried out screening for NPM2 expression in various human cancer cell lines available in our laboratory. Preliminary results from different cell lines showed that NPM2 mRNA was indeed expressed at low or negligible levels in most of the cancer cell lines tested in comparison to the liver cancer cell line Hep G2 where NPM2 mRNA is reported to be expressed (Figure 5.2.13A – B). We did not detect any NPM2 protein by western blotting analysis in the Hep G2 cells (Figure 5.2.13C) or several other cell lines such as Hep3B, SK-HEP-1, PLC/PRF/5, Huh-7, A2780, HEK-293, and AW8507 (data not shown). The mRNA expression of NPM2 in Hep G2 could be resulting from a transcriptionally permissive landscape at the NPM2 gene promoter evident by the higher enrichment of activating histone marks at the NPM2 promoter (Figure 5.2.14A) compared to that in HeLa S3 cells (Figure 5.2.14B), and unmethylated promoter CpG island in Hep G2 cells compared to hypermethylated CpG islands in HeLa S3 cells (Figure 5.2.14C).

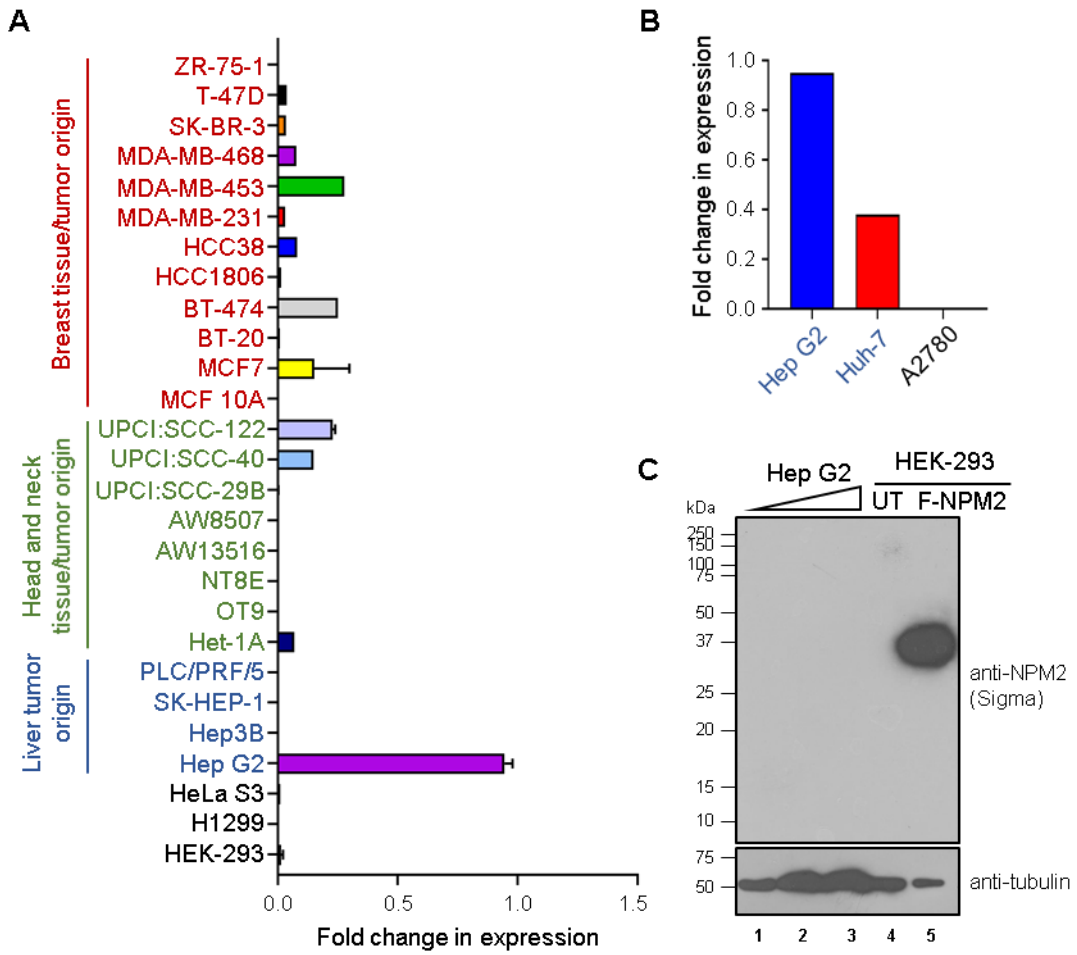



**Figure 5.2.13. Expression analyses of NPM2 in human cancer cell lines:** (A – B) Bars represent fold change in expression of NPM2 transcript in different human cancer cell lines as indicated, determined by RT-qPCR analysis. Internal normalization was done by (A) β- actin and (B) 18S rRNA levels. (C) Western blot analysis showing the expression status of NPM2 protein (upper panel) in increasing amounts of whole-cell lysates from Hep G2 cells (lanes 1 – 3), untransfected (UT) HEK-293 cells (negative control, lane 4) and HEK-293 stably expressing FLAG-tagged NPM2 (positive control, lane 5). Lower panel shows western blot with anti-tubulin antibody.

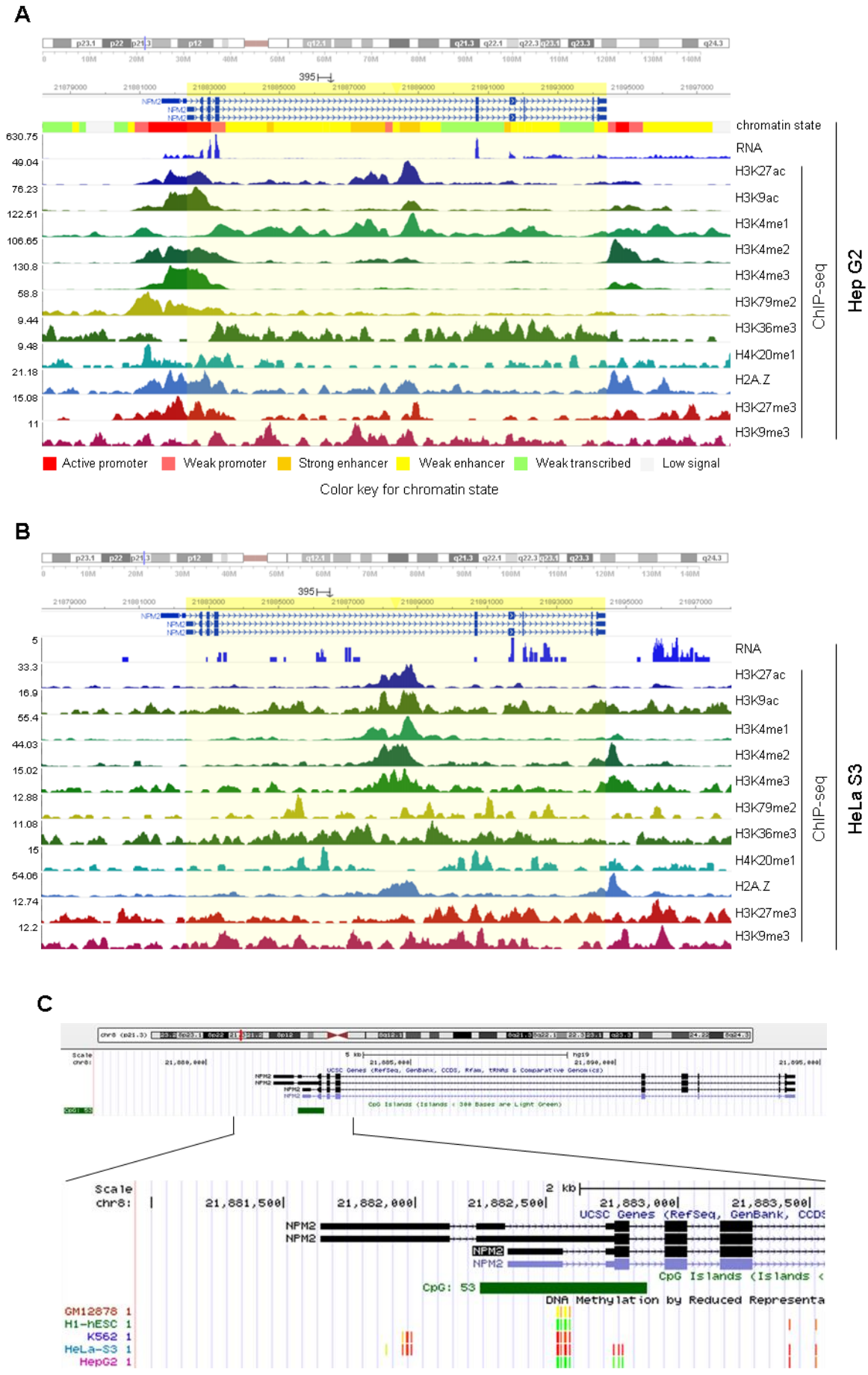

A
chromatin state
RNA
H3K27ac
H3K9ac
H3K4me1
H3K4me2
H3K4me3
H3K79me2
H3K36me3
H4K20me1
H2A.Z
H3K27me3
H3K9me3
ChIP-seq
Hep G2
Active promoter
Weak promoter
Strong enhancer
Weak enhancer
Weak transcribed
Low signal
Color key for chromatin state
B
HeLa S3
C
Scale
chr8:
21,881,500
21,882,000
21,882,500
21,883,000
21,883,500
2 kb
UCSC Genes (RefSeq, GenBank, CCDS
NPM2
CpG Islands (Islands
CpG: 53
DNA Methylation by Reduced Representa
GM12878
H1-hESC
K562
HeLa-S3
HepG2

**Figure 5.2.14. The epigenetic landscape of human NPM2 gene and promoter in cancer cell lines:** (A – B) WashU Epigenome Browser (v46.2) screenshots showing the mRNA expression levels of NPM2 (gene highlighted in yellow) and enrichment of various histone marks as indicated in the human cancer cells lines (A) Hep G2 and (B) HeLa S3 determined through RNA-seq and ChIP-seq analyses respectively, retrieved from the hg19 assembly of the human genome in ENCODE database. Note the scales of the individual tracks showing that the NPM2 transcript is expressed in higher levels in Hep G2 cells compared to HeLa S3 cells. The lower value of the scale of every track is zero. The parameters of the respective tracks for the figure (A) and (B) are the same for the sake of comparison. (A) The ChromHMM track for Hep G2 cells is shown at the top indicating the state of the chromatin while its color key is shown at the bottom. (C) UCSC Genome Browser screenshot from the ENCODE database showing the presence of 53 CpG sites (denoted in green bar) in the promoter region of human NPM2. The zoomed-in view if the NPM2 promoter region is shown below. The methylation status of the CpG sites in different cell lines is color-coded. Red denotes methylated, green denotes unmethylated and yellow denotes partially methylated CpG. Note that CpG island in NPM2 promoter in Hep G2 cell line is unmethylated while that in HeLa S3 cell line is methylated.

From the data presented so far, it appears that NPM2 is generally not expressed enough to be detected in cancer cells. However, it is expressed in certain differentiated adult tissues and cells such as muscle cells, brain tissue, adrenal glands, and so on, at least at the transcript level, which needs to be further confirmed at the protein level. The expression of NPM2 transcript at significant amounts but not the protein, in the liver cancer cell line Hep G2 could be due to transcriptionally active chromatin in the region of the NPM2 gene. It is unclear at this point the reason for the absence of the translated protein from the full-length mRNA that is transcribed in these cells. Possible explanations for this observation are the presence of a translational block that is preventing the synthesis of the NPM2 polypeptide or the generation of some non-coding RNA from the NPM2 primary transcript which might have functional roles in the physiology of the tumor cells.

## *5.3. Discussion*

Nucleoplasmin is the founder member of the family of histone chaperones which includes the multi-functional protein Nucleophosmin (NPM1) and its ortholog NPM3 (Frehlick et al. 2007). Nucleoplasmin has been mostly reported to be an oocyte-specific protein (Burns et al. 2003; Lingenfelter et al. 2011; Bouleau et al. 2014) that is important for sperm chromatin remodeling (Philpott et al. 1991), oocyte development, and egg competence (Aegerter et al. 2005; Bouleau et al. 2014). One of the roles played by Nucleoplasmin in these developmental processes is through its histone chaperone activity towards the histone

proteins present in the egg which are involved in the remodeling of the chromatin during fertilization and zygote development (Dutta et al. 2001), a function that is also governed by the different post-translational modifications (PTMs) of Nucleoplasmin such as arginine methylation (Wilczek et al. 2011), glutamylation (Onikubo et al. 2015) and more importantly, phosphorylation (Leno et al. 1996; Finn et al. 2012). Another possible mode of functioning of Nucleoplasmin is through its potential transcription regulation and chromatin organization activity. It has been found that the absence or knockdown of Nucleoplasmin results in reduced zygotic gene expression in mouse and zebrafish embryos (Burns et al. 2003; Bouleau et al. 2014). Absence of Nucleoplasmin also causes an abnormal organization of the nucleolar chromatin (Burns et al. 2003) and the improper development of the nucleolus-like bodies in mouse oocytes which are important for sperm chromatin decondensation after fertilization of the oocytes (Inoue and Aoki 2010; Inoue et al. 2011). The biochemical mechanisms behind such observations have not been substantially investigated to date. It was previously demonstrated that Nucleoplasmin could stimulate the binding of transcription factors to a nucleosomal DNA template, thereby converting the nucleosome-DNA complex to a transcription factor-DNA complex, a mechanism which can potentially activate transcription from chromatin *in vivo* (Chen et al. 1994; Walter et al. 1995). In support of this hypothesis, Nucleoplasmin was found to be localized on the lampbrush chromosomes at the sites of transcription, associated with the ribonucleoprotein (RNP) particles in the amphibian model *Pleurodeles waltl*, (Moreau et al. 1986), indicating its association with the process of transcription.

In this study, we have characterized the inherent histone chaperone activity and transcriptional activation potential of the mammalian orthologs of Nucleoplasmin, that is NPM2 (human) or Npm2 (mouse). We found that recombinant, unmodified NPM2 has minimal intrinsic histone chaperone activity (Figure 5.2.1) and transcription activation potential (Figure 5.2.4) *in vitro*, which are significantly less than those of its paralog NPM1. In a cellular system, NPM2 was found to moderately enhance the transcription of a reporter gene, which was lesser than the enhancement obtained by the activity of NPM1 (Figure 5.2.5A). Artificial overexpression of NPM2 in a system where it is naturally not expressed such as in HEK-293 cells, did not significantly alter the expression of candidate genes (Figure 5.2.5C) indicating that NPM2 might not be a regulator of expression of those genes or it could possibly manifest its transcription regulation function in its natural environment with the contribution of other factors such as its PTMs and/or interacting partners. The role

of PTMs such as phosphorylation, in chromatin remodeling and the regulation of transcription by Nucleoplasmin, is evident from the fact that hyperphosphorylated Nucleplasmin isolated from *Xenopus* eggs could unfold sperm and somatic chromatin more efficiently by removing the chromosomal proteins from linker DNA regions (Ramos et al. 2005), a mechanism which can be potentially utilized for activating transcription from the cellular chromatin. Further, human phospho-mimic NPM2 was found to regulate the expression of genes involved in the naïve stem cell stage during induced pluripotent stem cell generation (Huynh et al. 2016). In this context, we screened for specific enzyme-mediated PTMs of human NPM2 with the objective of studying the functional roles of such PTMs especially in the regulation of chromatin transcription. In our experiments, we could not detect acetylation of NPM2 by the lysine acetyltransferases (KATs) p300, PCAF, Tip60, and HAT (histone acetyltransferase) domain of MOZ, or arginine methylation by the methyltransferase PRMT4 or CARM1 (Figure 5.2.6). However, we found that NPM2 was an *in vitro* substrate of the mitotic kinases Aurora kinase A and B (Figure 5.2.7), both of which phosphorylated NPM2 majorly at its site Ser 196 which is conserved across many species (Figure 5.2.8). We found that Aurora kinase-mediated mass phosphorylated NPM2 exhibited significantly higher histone chaperone activity compared to its unmodified forms (Figure 5.2.7). However, the role of specific site-modified NPM2 in its activity and transcriptional activation needs to be further tested by the use of phospho-mimic mutants of NPM2 in the *in vitro* histone chaperone and chromatin transcription assays.

To study the physiological role of Aurora kinase-mediated phosphorylation of NPM2, we searched for cellular systems where NPM2 is naturally expressed. RNA-seq derived expression data publicly available at the ENCODE databases for human and mouse showed that besides the expected expression of Nucleoplasmin in the oocytes, it is also expressed at the transcript level in other human adult tissues such as endocrine glands, brain, spleen, testis, kidney, and prostate (Figure 5.2.9A), and in placenta and testis in mouse (Figure 5.2.9D), the latter been validated experimentally by us (Figure 5.2.9B). We also observed that the expression of Npm2 transcript increases in the course of muscle cell differentiation in the *in vitro* muscle differentiation model that is the mouse myoblast cell line C2C12 (Figure 5.2.10C). This observation was supported by previous reports that suggested that Nucleoplasmin is expressed at low levels in the striated muscle tissue of the tadpole and heart muscle tissue of the adult *Xenopus laevis* (Litvin and King 1988), as well as in zebrafish and *Xenopus tropicalis* muscle tissues (Cheung et al. 2018). The expression of

Nucleoplasmin in the mouse testis and muscle cells need to be confirmed at the protein level, in which case, sperm and muscle cell differentiation would provide interesting models, besides oocyte development and induced pluripotency, to study the functions of Nucleoplasmin and its phosphorylation mediated by the Aurora kinases which are also known to play important roles in these processes (Kimmins et al. 2007; Lee et al. 2012a; Nguyen and Schindler 2017; Johnson et al. 2018; Willems et al. 2018; Dhanasekaran et al. 2019; Karthigeyan et al. 2019). It would be interesting to study the regulation of Nucleoplasmin expression in these systems. A study characterizing the promoter of three maternal effect genes including Nucleoplasmin in mouse showed that the cis-acting elements E-box and NBE (NOBOX element) present in its promoter DNA sequence contribute to its oocyte-specific expression (Tsunemoto et al. 2008). This suggests that germ-cell specific transcription factors that bind to these DNA elements and whose absence or reduction results in female sterility, the phenotype that is also manifested by Nucleoplasmin depletion, might regulate Nucleoplasmin expression in the germ cells. Examples of such factors include FIGLA (folliculogenesis specific bHLH transcription factor) (Liang et al. 1997; Soyal et al. 2000; Zhao et al. 2008), NOBOX (newborn ovary homeobox) (Rajkovic et al. 2004; Qin et al. 2007), SOHLH1 and SOHLH2 (spermatogenesis and oogenesis specific basic helix-loop-helix) (Ballow et al. 2006; Pangas et al. 2006; Suzuki et al. 2012) and so on. Also, the myogenic factors such as MyoD, Myf5, Myogenin, and MRF4 can heterodimerize with E-protein family members to bind to E-box elements at the promoters of their target genes (Lassar et al. 1991). Although database analyses indicated that the proximal promoter of mouse Npm2 was not significantly enriched with MyoD, Myogenin or activating histone modification marks in the mouse myoblast cell line C2C12 (Figure 5.2.11A), Npm2 promoter harbors binding motifs for Myf proteins (Figure 5.2.11B), suggesting Myf-mediated basal activation of Npm2 gene transcription in these cells.

Since our studies and previous reports indicated NPM2 as a germ-cell specific protein, we studied its expression in cancerous cells of different histological origins since germ-cell specific proteins are often found to get expressed in tumor cells due to their deregulated expression regulation (Hofmann et al. 2008; Lifantseva et al. 2011; Bruggeman et al. 2018). In accordance with the expression data available from different databases such as the ENCODE, GeneCards, the Human Protein Atlas and TCGA (Figure 5.2.12), we did not observe appreciably high expression of NPM2 in most of the cancer cell lines tested except

for the hepatocellular carcinoma cell line Hep G2 (Figure 5.2.13A – B), where we could not detect the presence of NPM2 protein though by western blotting (Figure 5.2.13C). The expression of NPM2 transcript in Hep G2 cells but not in other cell lines such as HeLa S3 was probably due to a transcriptionally active environment at the NPM2 promoter in the former, such as presence of activating histone marks (H3K27ac, H3K9ac, H3K4me1, H3K4me2, H3K4me3, H3K79me2, and H2A.Z) and unmethylated CpG islands, which is absent in the latter cell type (Figure 5.2.14). The expression of the full-length transcript without a concomitant expression of the protein in these cells could be due to a translational block occurring in this scenario or the processing of the NPM2 primary transcript to generate a non-coding RNA having some role in the survival of these tumor cells. These are hypothetical explanations for our observations and it is up to further investigations to confirm if these are true.

# *Chapter 6: Summary and Discussion*

Nucleophosmin (NPM1) is a multifunctional histone chaperone that has been shown to regulate RNA Pol II-driven transcription (Swaminathan et al. 2005). Acetylation of NPM1 mediated by the lysine acetyltransferase p300 could enhance its histone chaperone activity and transcription activation potential (Swaminathan et al. 2005). However, the mechanism of regulation of RNA Pol II transcription in the nucleus of the cell by a predominantly nucleolar protein that is NPM1 was largely unclear. The observation that acetylated NPM1 (AcNPM1) was distributed to the nucleoplasm instead of the nucleolus and colocalized with the RNA Pol II foci suggested that this specific post-translational modification was important for the transcription coactivation function of NPM1 (Shandilya et al. 2009). The analysis of genome-wide occupancy of AcNPM1 by ChIP-seq assay in HeLa S3 cells showed that AcNPM1 was indeed present at the promoters and TSS of active genes along with RNA Pol II (Senapati P, Ph.D. thesis, 2014). In this study, we have further characterized the AcNPM1 ChIP-seq peaks to gain more insights into the genomic localization of AcNPM1 and understand the stage of transcription where it is involved. We found that the enrichment of AcNPM1 peaks in the genome correlated well with histone modification marks present in promoter regions such as H3K4me3, H3K9ac, H3K27ac, H3K4me2, and H3K4me1. On the other hand, there was no correlation between the occupancy of AcNPM1 and histone modifications present on the gene bodies (H3K79me2, H3K36me3, and H4K20me1) or in transcriptionally repressed regions of the genome (H3K9me3 and H3K27me3). This indicated that AcNPM1 is associated with the initiation of transcription at the promoters of active genes. However, we cannot rule out the possibility of the presence of unmodified NPM1 in the gene bodies and associated with transcription elongation since several attempts to perform NPM1 ChIP-seq have been unsuccessful due to the unavailability of ChIP-grade polyclonal antibodies against NPM1. We also observed enrichment of AcNPM1 in the enhancer regions of the genome which are characterized by the presence of DNase I hypersensitive sites and the histone modification mark H3K27ac. A significant overlap of the occupancy of AcNPM1 with RNA Pol II and the transcriptional coactivator p300 further supported the idea that AcNPM1 is associated with active transcription at gene promoters and enhancers.

In this study, we have also tried to address the mechanism of NPM1/AcNPM1-mediated transcription regulation. Biochemical approaches undertaken previously by our group suggested that the histone chaperone activity of NPM1 mediated by its N-terminal oligomerization domain is required for its transcription activation potential *in vitro* (Swaminathan et al. 2005). Point mutants of NPM1, namely L18Q NPM1 and Y17T-C21F NPM1, which were deficient in their oligomerization propensity, were found to be deficient in their histone interaction ability, histone chaperone activity, as well as transcription activation function (Senapati P, Ph.D. thesis, 2014). It can be presumed that, compared to wild-type NPM1, these oligomerization-defective mutants could be deficient in evicting the histones from a nucleosomal template, which is a prerequisite for transcription through chromatin. Our current efforts are underway to experimentally validate this hypothesis. In addition, NPM1 could also potentially activate transcription by recruiting the RNA Pol II machinery and other transcriptional activators and coactivators at the gene promoters, by interacting with these proteins. High throughput approaches previously undertaken by our group have suggested a host of cellular proteins as potential interacting partners of NPM1 (Senapati P, Ph.D. thesis, 2014). Among these proteins, we obtained numerous transcription-related protein candidates such as RNA Pol II subunits, chromatin remodelers, transcription factors, subunits of the Mediator complex and other transcriptional coactivators. Upon analysis, the AcNPM1 ChIP-seq peaks were found to be enriched in binding motifs of various transcription factors. There was also substantial overlap between the genome-wide enrichment profiles of AcNPM1 and some of the transcription enhancing proteins belonging to the class of transcription factors, RNA Pol II subunits, chromatin remodelers, among others. These analyses support the hypothesis that NPM1/AcNPM1 might assist in the recruitment of the transcription machinery at the gene promoters through its interactions with transcription factors, coactivators, and RNA Pol II subunits. It could also help in remodeling the chromatin to facilitate the movement of the transcribing RNA Pol II, through its interactions with chromatin remodelers. NPM1 could also positively affect transcription through indirect ways such as modulating the activity of the lysine acetyltransferase (KAT) and transcriptional coactivator p300. NPM1 was shown to induce the autoacetylation of p300 which enhances its KAT activity (Arif et al. 2010). The autoacetylated p300 can hyperacetylate nucleosomal histones at the gene promoter regions which can facilitate nucleosome disassembly for activating chromatin transcription. Interestingly, it was shown that oligomerization of NPM1 and its molecular chaperone activity was important for NPM1-mediated induction of p300 autoacetylation (Kaypee et

al. 2018b). This further emphasizes the fact that the oligomerization of NPM1 contributes to its transcription activation potential.

The functional characterization of NPM1 has been mostly done in the context of cancer. NPM1 was found to be more abundant in actively proliferating cells (Feuerstein and Mond 1987; Feuerstein et al. 1988a). Several functions of NPM1 such as ribosome biogenesis, cell cycle progression, inhibition of differentiation and apoptosis, stimulation of DNA synthesis, mediating DNA repair and activating gene transcription, could prove advantageous to cell growth and proliferation in cancerous conditions. In agreement with this point, it has been found that NPM1 gets overexpressed in a variety of cancer types and downregulation of NPM1 levels could result in the reduction of tumorigenic properties of the cancer cells (Grisendi et al. 2006). However, the molecular mechanisms of NPM1 overexpression in cancer have not been substantially investigated. We have previously reported that NPM1 is overexpressed in oral cancer with a concomitant increase in its acetylated levels. The levels of NPM1 were also found to be increased with the advancing grades of oral tumors (Shandilya et al. 2009). In this study, we have carried out investigations to understand the molecular mechanisms of NPM1 overexpression in cancer, focussing especially on oral cancer. After characterizing the minimal promoter of the human NPM1 gene, we identified the oncogene c-fos as a novel regulator of NPM1 promoter activity (Senapati P, Ph.D. thesis, 2014). We showed that c-fos, together with its heterodimerizing partner c-jun forming the complex AP-1, transcriptionally activates NPM1 expression by directly binding to its cognate sites on the NPM1 promoter. Our immunohistochemical analysis using oral cancer patient tissue samples also showed a positive correlation between NPM1 and c-fos expressions indicating that c-fos-mediated regulation of NPM1 expression might contribute to oral cancer manifestation. We also observed a moderate but positive correlation in the staining for NPM1 and p53 protein in the oral cancer patient tissue samples which we used for our study. The tumor suppressor p53 is frequently found to lose its tumor-suppressive functions in cancer cells due to mutations in the *TP53* gene. Certain mutations known as hot-spot gain-of-function mutations, which are mainly missense in nature and selected in cancer cells at a relatively higher frequency, confer the protein with altered structural and functional properties that are manifested as oncogenic rather than tumor-suppressive. We found that one such mutant, R175H p53, which is one of the most frequently observed mutants of p53 in cancer, could positively regulate NPM1 expression while there was no significant effect of wild-type p53

on NPM1 expression. R175H p53 could transcriptionally upregulate NPM1 expression although it is incapable of binding to p53 cognate sites in the promoter DNA by itself without the help of ancillary factors. Interestingly, our results showed that R175H p53 could interact with c-fos and get recruited at the c-fos/AP-1 binding sites on the NPM1 promoter region, resulting in a possible synergistic effect on the regulation of NPM1 expression in cancer cells. Although we have studied only three hot-spot p53 mutants namely R175H, R249S, and R273H, out of which only R175H was found to transcriptionally upregulate NPM1 expression, we cannot rule out the effects of other similar p53 mutants on the regulation of NPM1 expression. In our immunohistochemical analysis in the oral cancer patient tissue samples, we presume that the positive staining for p53 indicates the presence of mutant forms of p53 in the tumor samples. The isolation of the genomic DNA from these tumor cells and sequencing of the *TP53* gene would help us in identifying the specific *TP53* gene mutations prevalent in oral cancer patients in the Indian cohort. Further biochemical and molecular biological analysis could help us understand if the mechanisms of mutant p53-mediated regulation of NPM1 expression are specific to certain mutants or more generalized.

Besides c-fos and mutant p53, we also found that the transcription factor YY1 regulates the expression of NPM1 in oral cancer cells, the mechanisms of which are yet to be determined. We speculate that the effect of YY1 on NPM1 expression could be direct since *NPM1* promoter harbors numerous YY1-binding sites as well as it has been previously reported to be regulated by YY1 (Chan et al. 1997). However, the effect of YY1 on NPM1 expression could also be indirect or synergistic with other factors such as c-fos since c-fos is a known target of YY1 (Natesan and Gilman 1993; Zhou et al. 1995).

Our endeavors to discern the mechanisms of regulation of NPM1 expression would provide insights into the fundamental transcriptional networks operating in cells that contribute to oncogenesis. Additionally, such studies could help in developing panels of molecular markers for the detection and diagnosis of specific types of cancers in specific regional cohorts. A comprehensive analysis of the underlying molecular pathways is important to target specific proteins at the optimum degree for the purpose of cancer treatment which could have the maximum benefit and least possible adverse effects. We have shown that reducing the levels of NPM1 in oral cancer cells and orthotopic tumors in mice indeed result in a significant reduction in the tumorigenic properties of the cells as well as the size and metastasis of the oral tumors *in vivo*. This provides the first proof of principle that

targeting NPM1 levels and thereby its functions can lead to oral tumor regression. NPM1 regulates the expressions of a large number of genes most of which belong to various oncogenic pathways. Some of these genes, such as c-fos, regulate NPM1 expression as well thereby operating in feedback loops. Hence NPM1 serves as a promising target in cancer including oral cancer, which has the potential to reduce tumor progression to a clinically significant extent. Collectively, our studies have contributed to the fundamental biology of the role of overexpressed NPM1 in cancer and its clinical implications (Figure 6.1).

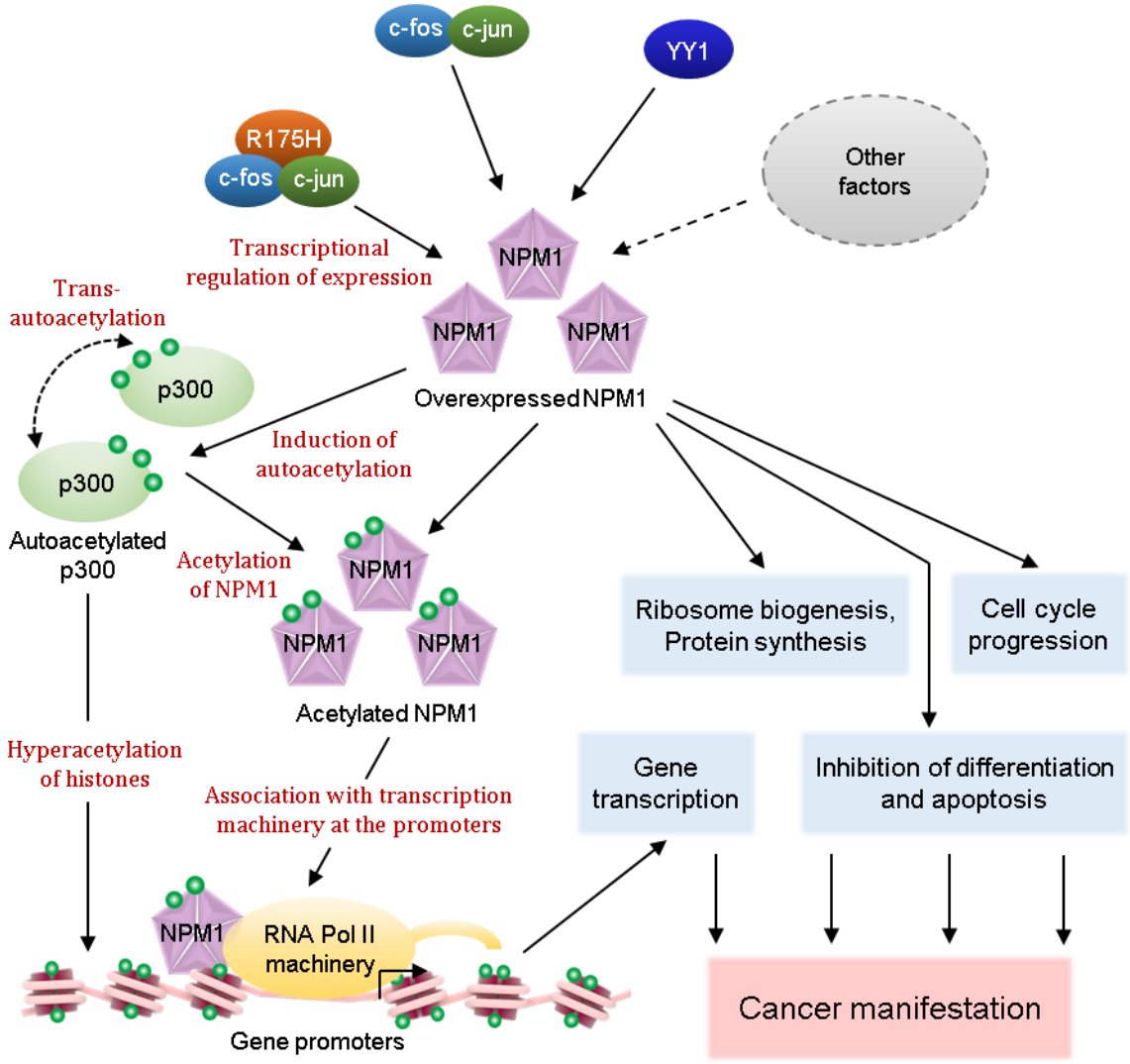



**Figure 6.1. Role of overexpressed NPM1 in cancer:** The overexpression of NPM1 in cancer can be brought about by several factors such as c-fos/AP-1, mutant p53 (Senapati et al. 2018), YY1, among others, through multiple mechanisms. The increased NPM1 levels could also lead to an elevated pool of acetylated NPM1 (AcNPM1). Acetylation of NPM1 is mediated by p300 (Swaminathan et al. 2005; Shandilya et al. 2009) whose enzymatic activity can be enhanced through its trans-autoacetylation that is induced by NPM1 itself (Arif et al. 2010; Kaypee et al. 2018b). The activated p300 can hyperacetylate nucleosomal histones at the gene promoters which facilitate their eviction to open up the chromatin during transcription. Additionally, AcNPM1 also occupies gene promoters and possibly associate with the transcription machinery along with other

transcription activating factors (not shown here), leading to the upregulation of gene expression implicated in tumorigenesis (Senapati et al. 2019). The non-transcriptional functions of NPM1 such as ribosome biogenesis, cell cycle progression, and inhibition of differentiation and apoptosis are also boosted due to its overexpression in tumor cells and further contributes to cancer manifestation.

Our studies on NPM1 conducted to date have mostly focussed on the transcription regulatory function of NPM1 and the role of specific post-translational modifications of NPM1 which have implications in the pathophysiology of cancer, especially oral cancer (Swaminathan et al. 2005; Shandilya et al. 2009; Shandilya et al. 2014a; Shandilya et al. 2014b). We have tried to understand the mechanisms of NPM1-mediated transcriptional regulation, which, as mentioned earlier, can occur in multiple ways. While NPM1 by itself has the capability to activate transcription from the chromatin template (Swaminathan et al. 2005), this activity can be further enhanced by other factors such as its acetylation and association with transcription activating proteins. Its histone chaperone activity appears to be important for its intrinsic transcription activation property. This prompted us to ask the question if the members of the Nucleoplasmin family that includes NPM1, can fundamentally act as transcriptional coactivators by virtue of their similar structures and histone chaperone activities. Our previous study showed that the Nucleoplasmin homolog NPM3 lacked intrinsic histone chaperone activity but could enhance activator-dependent transcription from a reporter plasmid (Gadad et al. 2010). However, it is uncertain if NPM3, without its histone chaperone property, would be able to activate transcription by itself from the chromatin template. The histone chaperone activity of the Nucleoplasmin homologs is essentially contributed by their ability to form strong oligomers. NPM3 is unable to form homo-pentamers but may form weak dimers or trimers. On the other hand, NPM2 is capable of forming stable pentamers and decamers and hence can be expected to have histone chaperone activity and enhance chromatin transcription like NPM1. However, in this study, we observed that both the above-mentioned functions are significantly less in human NPM2 compared to NPM1. NPM2, like NPM3, could enhance activator-dependent transcription from a reporter construct, although the fold-activation was appreciably less than that obtained by NPM1. This suggests that NPM2 is a less efficient histone chaperone and transcriptional coactivator than NPM1, probably due to the absence of certain stretches of protein sequence such as the C-terminal region and an acidic tract, which contributes towards these functional properties. Further, we found that phosphorylation of NPM2 mediated by Aurora Kinase A and B caused a moderate but significant increase in the

histone chaperone activity of NPM2 *in vitro* and might affect its transcription activation property which is yet to be confirmed. Presumably, phosphorylation compensates for the deficiency of the negative charge due to the absence of an acidic tract in NPM2. This could potentially modulate the interaction ability of NPM2 towards core histones, thereby increasing its histone chaperone activity. The functional role of Aurora Kinase-mediated phosphorylation of NPM2 at the conserved site Ser196 is yet to be determined. The systems of mammalian germ cell development and skeletal muscle differentiation where there are indications for NPM2 expression could serve as interesting models to study this aspect of NPM2 function. *In silico* and experimental analyses of NPM2 expression in various human cancer cells suggest that it might not be present in rapidly proliferating tumor cells like NPM1 and NPM3. We speculate that NPM2 performs specialized non-redundant functions in specific differentiated cell types such as oocytes. Possibly, the natural *in vivo* cellular environment provided by these cells is required for the manifestation of the transcription regulatory function of Nucleoplasmin and its orthologs. In support of this hypothesis, we found that overexpression of NPM2 in HEK-293 cells where it is naturally absent did not result in any significant effect on the expression of housekeeping or ubiquitously expressed genes. However, an artificial perturbation of the levels of Nucleoplasmin orthologs in cells where they are physiologically expressed could alter the expressions of specific genes such as zygotic and pluripotency-related genes (Burns et al. 2003; Bouleau et al. 2014; Huynh et al. 2016).

In summary, this study was designed to understand the functional aspects of the mammalian Nucleoplasmin homologs with respect to their histone chaperone and transcription regulatory functions, as well as the regulation of their differential expression patterns. Our investigations provide interesting insights into these attributes with potential implications in clinical as well as fundamental transcription biology.

# *Chapter 7: Significance and Future Perspectives*

The regulation of the dynamicity of the chromatin is a fundamental requirement for critical nuclear processes such as DNA replication, repair, and transcription, which dictate the physiological or pathophysiological state of the cell. Among various factors, histone chaperones have emerged as important regulators of the chromatin and the different chromatin-templated processes such as transcription. FACT (facilitates chromatin transcription) is a classic example of a histone chaperone which plays an important role in nucleosome assembly and disassembly during transcription initiation and elongation (Formosa 2012). Other histone chaperones such as Nap1, Chz1, Asf1, HIRA, and Spt6 have also been reported to assist in the transcription process through multiple mechanisms (Avvakumov et al. 2011; Reddy et al. 2017). The nucleolar protein NPM1, a member of the Nucleoplasmin family of histone chaperones, was reported to regulate the transcription of the rRNA gene mediated by RNA Pol I (Murano et al. 2008). Our work reported here as well as in previous studies have shown that NPM1 also acts as a coactivator of RNA Pol II-driven transcription (Swaminathan et al. 2005; Shandilya et al. 2009; Senapati et al. 2019). In this process, the acetylation of NPM1 plays an important role that results in its localization in the nucleoplasm instead of the nucleolus (Shandilya et al. 2009), and its possible association with the transcription machinery at the active regions of the chromatin such as gene promoters and enhancers (Senapati et al. 2019). Another nucleolar protein, Nucleolin (NCL), which is functionally analogous to NPM1, was also found to exhibit similar features. The unacetylated form of NCL was found to be enriched at the rDNA genes in the nucleolus, regulating their transcription mediated by RNA Pol I (Cong et al. 2012). On the other hand, the acetylated pool of NCL was found to be localized in the nucleoplasm predominantly in the nuclear speckles instead of the nucleolus, where it colocalized with the splicing factor SC35 indicating its role in the regulation of mRNA maturation and transcription by RNA Pol II (Das et al. 2013). This suggests that acetylation-dependent localization in the nucleoplasm and association with the transcription machinery to mediate RNA Pol II-driven transcription might be a general mechanism for nucleolar proteins such as NPM1 and NCL. However, upon distribution of acetylated NPM1 (AcNPM1) to the nucleoplasm, the precise downstream biochemical mechanisms of AcNPM1-mediated regulation of transcription has to be elucidated. Our

findings so far help us to frame multiple hypotheses of the possible mechanisms in this direction such as assisting in nucleosome disassembly and chromatin remodeling during transcription, and recruitment of the transcription machinery and/or other transcription-activating proteins. These possibilities need to be further tested and validated by biochemical approaches.

An emerging concept in the transcription field suggests that cellular transcription occurs in phase-separated condensates which helps to concentrate and compartmentalize biochemical reactions in a three-dimensional spatial organization of the cell. This phenomenon provides the physical means to control transcription which is in compliance with the established and recently proposed models of multi-factorial regulation of cellular transcription (Hnisz et al. 2017). The formation of these phase-separated condensates is contributed by the linker histone H1 of the chromatin as well as the intrinsically disordered regions (IDRs) in the structures of several proteins associated with transcription such as the RNA Pol II CTD and general coactivators (such as Mediator complex, BRD4) (Kwon et al. 2013; Cho et al. 2018; Gibson et al. 2019). The IDRs present in NPM1 have also been shown to contribute to liquid-liquid phase separation (LLPS) in the nucleolus through multiple mechanisms, which help in the process of ribosome biogenesis (Mitrea et al. 2018). Likewise, NPM1 could potentially facilitate the formation of such phase-separated condensates in the nucleoplasm during RNA Pol II transcription through its interactions with RNA and other coactivators and transcription factors. A conceptually similar mechanism is adopted by a nucleolar ribogenesis factor WDR43 which was recently shown to be a chromatin-associated RNA binding protein that was enriched at a large number of active gene promoters and enhancers in embryonic stem cells. WDR43 could regulate gene expression by promoting the release of paused RNA Pol II through multivalent interactions probably in phase-separated condensates (Bi et al. 2019). The C-terminal IDR of NPM1 which has been shown to contribute to LLPS also harbors sites for various PTMs such as phosphorylation, SUMOylation as well as acetylation (Okuwaki 2008; Shandilya et al. 2009). It would be interesting to study the effect of these PTMs, especially acetylation, on the formation of such phase-separated transcription hubs, and if there is any possible cross-talk among these PTMs to regulate transcription.

From the studies performed by us and other groups, it appears that the role of NPM1 as a transcriptional coactivator and its mechanistic aspects could share more similarity with functionally analogous nucleolar proteins such as NCL and WDR43 as discussed in the

preceding paragraphs, rather than its structural homologs in the Nucleoplasmin family such as NPM2. The members of the Nucleoplasmin family probably have some basic ability to regulate transcription, which has evolved to different extents among the three members, the greatest being in NPM1. It also appears that this transcription coactivation function may or may not be dependent on the histone chaperone activities of the NPM proteins but is rather dictated by the *in vivo* environment where they are physiologically expressed. The inherent transcription activation property of NPM1 and its multiple other transcription independent functions, which are important for cell growth and survival, probably justify its ubiquitous expression under the normal physiological conditions, and overexpression in cancer cells. On the other hand, NPM2 could probably have evolved to perform very specialized functions in only the few cell types where it is physiologically expressed. Such functions would include its role as a special chaperone for histones and possibly histone variants present in germ cells, specific remodeling of the germ cell chromatin, and regulation of certain genes relevant for germ cell development, naïve pluripotency, and cellular differentiation. Our preliminary data support some of these above-mentioned hypotheses. However, further experimental investigations are required to validate these theories. Nonetheless, these observations and speculations provide fundamental insights into the functional evolution of proteins – how the presence of a few additional and unique motifs in one protein (NPM1) in a family of structurally related proteins (NPM2, NPM3) could confer dramatic differences in their cellular functions. This is also reflected in the differential expression patterns of the NPM proteins and the regulation of their expressions. The ubiquitous physiological expression of NPM1, its overexpression in tumor cells which often gets further stimulated in response to various signals as we have earlier discussed, could be brought about through several mechanisms such as by transcription factors, non-coding RNAs and an active epigenetic state of the chromatin at the *NPM1* gene promoter. On the other hand, the tissue-restricted expression of NPM2 is expected to be regulated by tissue-specific factors that can alter the epigenetic state of its gene promoter into a transcriptionally permissive one only in those specific cell types instead of all tissues. The deficiency or absence of NPM2 has been found to result in fertility defects in female mice (Burns et al. 2003). Similar fertility defects in humans have been associated with mutations in the genes *FIGLA* and *NOBOX*, which code for the master transcription factors present in oocytes (Rajkovic et al. 2004; Qin et al. 2007; Zhao et al. 2008). Further, the mouse *Npm2* gene promoter has been shown to harbor cognate binding sites for such transcription factors (Tsunemoto et al. 2008). This suggests that there could be a functional relationship between

NPM2 and these transcription factors which regulate the transcriptional profile of the oocytes where they are expressed. Studying these pathways could provide valuable insights into the fundamental molecular biology of germ cell development.

Finally, in the context of NPM1 function and expression in cancer, in our studies to date, we have not considered the role of non-coding RNAs which have emerged as important regulators of protein function as well as the transcription process. Besides being a histone chaperone, NPM1 is also an RNA binding protein whose functions in cancer cells could be modulated and complicated through its associations with non-coding RNA as has been suggested in a few reports. For instance, the lnc RNA LNMICC could promote lymph node metastasis in cervical cancer by reprogramming fatty acid metabolism where it recruited NPM1 to the promoter of the fatty acid-binding protein FABP5 (Shang et al. 2018). However, the functional consequences of some other such associations are yet to be uncovered (Dumbovic et al. 2018). We have observed the presence of AcNPM1 at active enhancer regions of the genome through our ChIP-seq analysis. It would be interesting to study if nascent enhancer RNAs transcribed from these regions associate with NPM1/AcNPM1 to exert their gene regulatory functions and the role of NPM1 itself in these regulatory networks. Few non-coding RNAs, such as miRNA and lnc RNA, have also been identified to regulate NPM1 expression in different contexts. For example, NPM1 was found to be a target of miR-646, and the lnc RNA HOTAIR promoted NPM1 expression by negatively regulating miR-646 in endometrial cancer cells (Zhou et al. 2018). In another instance, the miRNA hcmv-miR-US25-1-5p was found to downregulate NPM1 expression during HCMV infection causing the inhibition of viral replication (Jiang et al. 2015). Our preliminary *in silico* analysis has identified a number of putative miRNAs which could potentially regulate NPM1 expression. Future experimental studies in this direction would help us uncover specific non-coding RNAs regulating NPM1 expression in cancer, specifically oral cancer, thereby serving as prognostic markers for this very prevalent disease in the Indian subcontinent. Our collective efforts are underway to correlate the deregulation of NPM1 to the incidence and prognosis of oral cancer in the Indian cohorts where we have discovered NPM1 to be overexpressed, to potentially bring out the translational significance of our fundamental findings.

# *Publications*

**Research Articles:**

1. Senapati P*, **Dey S***, Sudarshan D, Das S, Kumar M, Kaypee S, Mohiyuddin A, Kodaganur GS, Kundu TK. (2018) Oncogene c-fos and mutant R175H p53 regulate expression of Nucleophosmin implicating cancer manifestation. ***FEBS J***. 285(18):3503-3524.

2. Senapati P, **Dey S**, Sudarshan D, Bhattacharya A, Shyla G, Das S, Sudevan S, Maliekal TT, Kundu TK. Histone chaperone Nucleophosmin regulates transcription of key genes involved in oral tumorigenesis. ***Manuscript communicated*** and uploaded in ***bioRxiv*** (https://doi.org/10.1101/852095).

3. Bose A, Modi K*, **Dey S***, Dalvi S, Nadkarni P, Sudarshan M, Kundu TK, Venkatraman P, Dalal SN. 14-3-3γ inhibits centrosome duplication by inhibiting NPM1 function. ***Manuscript communicated***.

   * Equal Contribution.

# *Abbreviations*

| | |
|---|---|
| °C | degrees Celsius |
| ~ | approximately |
| Δ | deletion |
| Δ | heat treatment (at 90°C) |
| / | per |
| % | percent |
| 2D | two dimensional |
| 3C | chromosome conformation capture |
| 3D | three dimensional |
| 14-3-3γ/YWHAG | tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein gamma |
| [α-$^{32}$P]-UTP | uridine triphosphate (UTP) radiolabeled on the alpha phosphate group with $^{32}$P |
| β-Me | β-mercaptoethanol |
| [γ-$^{32}$P]-ATP | adenosine triphosphate (ATP), radiolabeled on the gamma phosphate group with $^{32}$P |
| ε | molar extinction coefficient (of histone at 276 nm) |
| µ | micro |
| µCi | micro Curie |
| µg | microgram(s) |
| µl | microlitre(s) |
| µm | micrometer(s) |
| µM | micromolar |
| Ψ | retrovirus psi (packaging sequence) |
| A | adenosine |
| $A_{260/230}$ | absorbance at 260 nm wavelength by absorbance at 230 nm |
| $A_{260/280}$ | absorbance at 260 nm wavelength by absorbance at 280 nm |
| Ac-CoA | acetyl coenzyme A |
| Acf1 | ATP-utilizing chromatin assembly and remodeling factor 1 |
| AcNPM1 | acetylated NPM1 |
| Actb | mouse beta-actin |
| ACTB | human beta-actin |
| ALCL | anaplastic large cell lymphoma |
| ALK | anaplastic lymphoma kinase |
| AML | acute myeloid leukemia |
| amp$^{r}$ | ampicillin resistance (marker gene) |
| Ampicillin$^{R}$ | ampicillin resistance (marker gene) |
| ANGPTL4 | angiopoietin like 4 |
| ANOVA | analysis of variance |
| ANP32E | acidic nuclear phosphoprotein 32 family member E |
| AP-1 | Activator protein 1 |
| AP-2 alpha/AP-2α (TFAP2A) | Activator Protein 2 alpha (transcription factor AP-2 alpha) |
| AP-2 gamma/AP-2γ (TFAP2C) | Activator Protein 2 gamma (transcription factor AP-2 gamma) |
| AP-4 (TFAP4) | Activator Protein 4 (transcription factor AP-4) |
| AP1BS | AP-1-binding site |
| APE1 | apurinic/apyrimidinic endonuclease 1 |
| APL | acute promyelocytic leukemia |
| APLF | Aprataxin and polynucleotide kinase like factor |
| APS | ammonium persulphate |
| ARF | alternative reading frame |
| Arg | arginine |
| Arp4 | actin related protein 4 |
| Arp7 | actin related protein 7 |
| Arp8 | actin related protein 8 |
| Arp9 | actin related protein 9 |
| Asf1 | anti-silencing factor 1 |
| ATF | activating transcription factor |

| | |
|---|---|
| ATM | ATM (Ataxia Telangiectasia Mutated) serine/threonine kinase |
| ATO | arsenic trioxide |
| ATP | adenosine triphosphate |
| ATRA | all-trans-retinoic-acid |
| ATRX | ATRX chromatin remodeler/ATP-dependent helicase ATRX/X-linked helicase II |
| AURKA | aurora kinase A |
| AURKB | aurora kinase B |
| BARD1 | BRCA1 associated RING domain 1 |
| BEGM | Bronchial Epithelial cell Growth Medium |
| BER | base excision repair |
| BGH | bovine growth hormone |
| bHLH | basic helix-loop-helix |
| BIV | bovine immunodeficiency virus |
| BOD | biochemical oxygen demand |
| bp | base pair(s) |
| BRCA1 | BRCA1 DNA repair associated/breast cancer 1 |
| BRD4 | bromodomain containing 4 |
| Bre | E3 ubiquitin-protein ligase Bre |
| BRE | TFIIB recognition element |
| $BRE_d$ | downstream TFIIB recognition element |
| $BRE_u$ | upstream TFIIB recognition element |
| BSA | bovine serum albumin |
| bZIP | basic leucine zipper |
| C | cytidine |
| c-myc | MYC proto-oncogene |
| C-terminal | carboxy-terminal |
| C/EBP-beta | CCAAT enhancer binding protein beta |
| $CaCl_2$ | calcium chloride |
| CacyBP/SIP | calcyclin-binding protein/Siah-1-interacting protein |
| CAD | caspase-activated DNase |
| CAF-1 | chromatin assembly factor 1 |
| CaMKII | Calcium/calmodulin-dependent protein kinase II |
| cAMP | cyclic adenosine monophosphate |
| CARM1 | coactivator associated arginine methyltransferase 1 |
| Cat. | catalogue |
| CBB | Coomassie Brilliant Blue |
| CBP | CREB-binding protein |
| CCNA2 | cyclin A2 |
| CCNB2 | cyclin B2 |
| CCND1 | cyclin D1 |
| CCND2 | cyclin D2 |
| CD44 | CD44 molecule (Indian blood group) |
| cdc2 | cyclin-dependent protein kinase Cdk1/Cdc (cell division control)2 |
| CDC25C | cell division cycle 25C |
| CDE | cell cycle-dependent element |
| cDNA | copy DNA |
| CDK1 | cyclin dependent kinase 1 |
| CDK2 | cyclin dependent kinase 2 |
| cdk5 | cyclin dependent kinase 5 |
| CDK6 | cyclin dependent kinase 6 |
| CDK7 | cyclin dependent kinase 7 |
| CDK8 | cyclin dependent kinase 8 |
| CDK9 | cyclin dependent kinase 9 |
| CDK12 | cyclin dependent kinase 12 |
| CDK13 | cyclin dependent kinase 13 |
| CDKN1A | cyclin dependent kinase inhibitor 1A |
| CDS | coding sequence |
| CEBPB | CCAAT enhancer binding protein beta |
| CENP-A/CENPA | centromere protein A |
| $CH_3COOK$ | potassium acetate |
| ChIP | chromatin immunoprecipitation |
| ChIP-seq | chromatin immunoprecipitation sequencing |
| CHR | cell cycle genes homology region |
| Ci | Curie |
| circRNA | circular RNA |
| CHD | chromodomain helicase DNA binding protein |

| | |
|---|---|
| CHD1 | chromodomain helicase DNA binding protein 1 |
| CHD2 | chromodomain helicase DNA binding protein 2 |
| Chz1 | chaperone for H2A.Z-H2B |
| CKI | Casein Kinase I |
| CKII/CK-2 | Casein Kinase II |
| CKM | CDK8 kinase module |
| CMV | Cytomegalovirus |
| CnBr | cyanogen bromide |
| CNS | central nervous system |
| $CO_2$ | carbon dioxide |
| COL16A1 | collagen type XVI alpha 1 chain |
| COMPASS | complex proteins associated with Set1 |
| CP | core promoter |
| cpm | counts per minute |
| CPSF | cleavage and polyadenylation specificity factor |
| CREB | cAMP responsive element binding protein |
| CRM1 | cellular export receptor 1/ chromosomal region maintenance protein 1 |
| cryo-EM | cryogenic electron microscopy |
| CSTF | cleavage stimulation factor |
| Ct | cycle threshold |
| CTCF | CCCTC-binding factor |
| CTEA | chromatin transcription-enabling activity |
| CTP | cytidine triphosphate |
| CUG2 | cancer upregulated gene 2 |
| CXCL8 | C-X-C motif chemokine ligand 8 |
| CycC | cyclin C |
| DAB | 3,3'-diaminobenzidine |
| DAXX | death domain associated protein |
| DB | Direct Blue |
| DBD | DNA binding domain |
| dCENP-A | *Drosophila* centromere protein A |
| DDR | DNA-damage response |
| DDX4 | DEAD-box helicase 4 |
| DDX31 | DEAD-box helicase 31 |
| DEGs | differentially expressed genes |
| DEK | DEK proto-oncogene |
| DEPC | diethyl pyrocarbonate |
| DHSs | DNase I hypersensitive sites |
| DMEM | Dulbecco's Modified Eagle's Medium |
| DMSO | dimethyl sulfoxide |
| DNA | deoxyribonucleic Acid |
| DNAPK | DNA-dependent protein kinase |
| dNTP | deoxyribonucleotide triphosphate |
| DPE | downstream promoter element |
| DPX | dibutylphthalate polystyrene xylene |
| DRB | 5,6-dichloro-1-β-d-ribofuranosylbenzimidazole |
| DSB | double-strand break |
| DSIF | DRB sensitivity inducing factor |
| dT | deoxythymidine |
| DTT | dithiothreitol |
| dTopo I | *Drosophila* Topoisomerase I |
| E | eluate |
| E-box | enhancer box |
| *E. coli* | *Escherichia coli* |
| E2F | E2 transcription factor |
| E2F1 | E2F transcription factor 1 |
| E2F4 | E2F transcription factor 4 |
| E2F6 | E2F transcription factor 6 |
| Eaf3 | essential Sas2-related acetyltransferase 1-associated factor 3 |
| EB | Epstein-Barr (virus) |
| EBNA-1 | Epstein-Barr nuclear antigen 1 |
| EBNA-2 | Epstein–Barr nuclear antigen 2 |
| EBP1 | ErbB3-binding protein 1 |
| EBV | Epstein-Barr virus |
| EC | elongation complex |
| ECGT | (−)-epigallocatechin-3-gallate |

| | |
|---|---|
| EDTA | ethylenediaminetetraacetic acid |
| EGFR | epidermal growth factor receptor |
| EGR1 | early growth response 1 |
| EGTA | ethylene glycol-bis(β-aminoethyl ether)-N,N,N′,N′-tetraacetic acid |
| ELK1 | ETS transcription factor ELK1, ETS Like-1 |
| ELK4 | ETS transcription factor ELK4, ETS Like-4 |
| ENCODE | Encyclopedia of DNA Elements |
| ENH | enhancer (sequence) |
| ERCC3 | ERCC excision repair 3, TFIIH core complex helicase subunit |
| ERRgamma | estrogen-related receptor gamma |
| EtBr | ethidium bromide |
| ETS | E26 transformation-specific/E-twenty-six |
| Ets-1 | ETS proto-oncogene 1, transcription factor |
| EV | empty vector |
| f1 | bacteriophage f1 |
| F3 | coagulation factor III, tissue factor |
| FABP5 | fatty acid binding protein 5 |
| FACT | facilitates chromatin transcription |
| FBS | fetal bovine serum |
| Fbw7γ | F-box and WD repeat domain containing 7 (gamma/isoform 3) |
| FIGLA | folliculogenesis specific bHLH transcription factor |
| FKBP | FK506 binding protein |
| FLAG | FLAG octapeptide |
| FN1 | fibronectin 1 |
| Fos | (mouse) FBJ osteosarcoma oncogene |
| FOS | (human) Fos proto-oncogene, AP-1 transcription factor subunit |
| FOXM1 | forkhead box M1 |
| FP | forward primer |
| FRET | fluorescence resonance energy transfer |
| g | gram(s) |
| g | units of times gravity (relative centrifugal force) |
| G | guanine |
| GABP | GA-binding protein |
| GAGE | G antigen |
| GAL4 | galactose-responsive transcription factor GAL4 |
| GAPDH | glyceraldehyde 3-phosphate dehydrogenase |
| GCN5/KAT2A | general control non-derepressible 5/lysine acetyltransferase 2A |
| GOF | gain-of-function |
| GRK5 | G protein-coupled receptor kinase 5 |
| GSEA | Gene Set Enrichment Analyses |
| GSK3 | glycogen synthase kinase 3 |
| GTF2A2 | general transcription factor IIA subunit 2 |
| GTF2E1 | general transcription factor IIE subunit 1 |
| GTF2E2 | general transcription factor IIE subunit 2 |
| GTF2F1 (TFIIF-alpha) | general transcription factor IIF subunit 1 |
| GTFs | general transcription factors |
| GTP | guanosine triphosphate |
| [$^3$H]-Ac-CoA | tritiated acetyl coenzyme A |
| h | hour(s) |
| H-score | histo-score |
| H2A.Bbd | H2A Barr body-deficient |
| H2A.XS139(129)ph | histone H2A.X serine 139 or 129 phosphorylation |
| H2BK123ub | histone H2B lysine 123 ubiquitination |
| H3ac | histone H3 acetylation |
| H3K4 | histone H3 lysine 4 |
| H3K4me1 | histone H3 lysine 4 monomethylation |
| H3K4me2 | histone H3 lysine 4 dimethylation |
| H3K4me3 | histone H3 lysine 4 trimethylation |
| H3K9ac | histone H3 lysine 9 acetylation |
| H3K9me2/3 | histone H3 lysine 9 di/trimethylation |
| H3K9me3 | histone H3 lysine 9 trimethylation |
| H3K14ac | histone H3 lysine 14 acetylation |
| H3K23ac | histone H3 lysine 23 acetylation |
| H3K27ac | histone H3 lysine 27 acetylation |
| H3K27me3 | histone H3 lysine 27 trimethylation |

| | |
|---|---|
| H3K36 | histone H3 lysine 36 |
| H3K36me2 | histone H3 lysine 36 dimethylation |
| H3K36me3 | histone H3 lysine 36 trimethylation |
| H3K56ac | histone H3 lysine 56 acetylation |
| H3K79me2 | histone H3 lysine 79 dimethylation |
| H3R2 | histone H3 arginine 2 |
| H4K5ac | histone H4 lysine 5 acetylation |
| H4K12ac | histone H4 lysine 12 acetylation |
| HAT | histone acetyltransferase |
| HBEGF | heparin binding EGF like growth factor |
| HBR | H2B repression |
| HBV | Hepatitis B virus |
| HCV | Hepatitis C virus |
| HCl | hydrochloric acid |
| HD | HAT domain |
| HDAC | histone deacetylase |
| HDAC2 | histone deacetylase 2 |
| hDEK-CAN | human DEK-CAN fusion protein |
| Hdm2/HDM2 | Mouse Double Minute 2 human homolog |
| HEK | Human embryonic kidney |
| HEXIM1 | HEXIM P-TEFb complex subunit 1/hexamethylene bisacetamide (HMBA) inducible protein 1 |
| HF | high fidelity |
| HFD | histone fold domain |
| hGH | human growth hormone |
| Hif1 | Hat1 interacting factor 1 |
| HIF-1α (HIF1A) | hypoxia inducible factor 1 subunit alpha |
| HIRA | histone cell cycle regulator |
| $His_6$ | hexahistidine |
| HIV | human immunodeficiency virus |
| HJURP | Holliday junction recognition protein |
| HLJ1/DNAJB4 | DnaJ heat shock protein family (Hsp40) member B4 |
| HM450 | HumanMethylation450 |
| HMG | high mobility group |
| HMGB1 | high mobility group box 1 |
| HMGB2 | high mobility group box 2 |
| HMM | hidden Markov model |
| HMOX1 | heme oxygenase 1 |
| HMR | Hybrid male rescue |
| HNSCC | head and neck squamous cell carcinoma |
| HP1 | heterochromatin protein 1 |
| HP1α | heterochromatin protein 1 alpha |
| HP1β | heterochromatin protein 1 beta |
| HP1γ | heterochromatin protein 1 gamma |
| hPTTG1 | human pituitary tumor-transforming gene 1 |
| HRP | Horse Radish Peroxidase |
| HSC70 | heat shock cognate protein 70 |
| HSP90A | heat shock protein 90 alpha |
| HSP90B | heat shock protein 90 beta |
| HuR (ELAVL1) | Hu antigen R (ELAV like RNA binding protein 1) |
| IDRs | intrinsically disordered regions |
| IFN-γ | interferon gamma |
| IgG | immunoglobulin G |
| IHC | immunohistochemistry |
| IKKα | IκB kinase alpha |
| INO80 | INO80 complex ATPase subunit |
| INSR | insulin receptor |
| INR | initiator motif |
| IP | immunoprecipitation |
| IPO4 | importin 4 |
| IPTG | isopropyl β- d-1-thiogalactopyranoside |
| IR | infra-red |
| IRES | internal ribosomal entry site |
| IRF | interferon regulatory factors |
| IRF1 | interferon regulatory factor 1 |
| ISWI | Imitation SWItch |

| | |
|---|---|
| ITS2 | internal transcribed spacer 2 |
| JDP2 | Jun dimerization protein 2 |
| JUN | Jun proto-oncogene, AP-1 transcription factor subunit |
| JUND | JunD proto-oncogene, AP-1 transcription factor subunit |
| KAT | lysine acetyltransferase |
| kb | kilobases |
| KCl | potassium chloride |
| kDa | kilodalton(s) |
| KDM5/Lid | lysine demethylase 5/little imaginal discs |
| KDM6A/UTX | lysine demethylase 6A/ubiquitously transcribed tetratricopeptide repeat, X chromosome |
| $KH_2PO_4$ | potassium dihydrogen phosphate |
| Ki67 (MKI67) | marker of proliferation Ki-67 |
| KLH | Keyhole Limpet Hemocyanin |
| KSRP (KHSRP) | K-homology splicing regulatory protein (KH-type splicing regulatory protein) |
| l | litre(s) |
| LB | Lysogeny broth/Luria broth/Luria Bertani Miller broth |
| LLPS | liquid-liquid phase separation |
| LTR | long terminal repeat |
| Luc | Luciferase |
| Lys | lysine |
| m | meter(s) |
| M | molar |
| M | molecular weight marker or ladder |
| MACS | Model-based Analysis of ChIP-Seq |
| MAGE | melanoma antigen |
| MALDI-TOF | matrix-assisted laser desorption/ionization – time-of-flight analysis |
| MAX | MYC-associated factor X |
| MAZ | MYC associated zinc finger protein |
| MCM2 | minichromosome maintenance complex component 2 |
| MCS | multiple cloning site |
| MDa | megadalton(s) |
| MDM2 | Mouse Double Minute 2 |
| MDS | myelodysplastic syndrome |
| MeCP2 | methyl CpG binding protein 2 |
| MED | mediator complex subunit |
| MED1 | mediator complex subunit 1 |
| MED2 | mediator complex subunit 2 |
| MED3 | mediator complex subunit 3 |
| MED4 | mediator complex subunit 4 |
| MED5 | mediator complex subunit 5 |
| MED6 | mediator complex subunit 6 |
| MED7 | mediator complex subunit 7 |
| MED8 | mediator complex subunit 8 |
| MED10 | mediator complex subunit 10 |
| MED11 | mediator complex subunit 11 |
| MED12 | mediator complex subunit 12 |
| MED13 | mediator complex subunit 13 |
| MED15 | mediator complex subunit 15 |
| MED16 | mediator complex subunit 16 |
| MED17 | mediator complex subunit 17 |
| MED18 | mediator complex subunit 18 |
| MED19 | mediator complex subunit 19 |
| MED20 | mediator complex subunit 20 |
| MED21 | mediator complex subunit 21 |
| MED22 | mediator complex subunit 22 |
| MED23 | mediator complex subunit 23 |
| MED25 | mediator complex subunit 25 |
| MED26 | mediator complex subunit 26 |
| MED28 | mediator complex subunit 28 |
| MED30 | mediator complex subunit 30 |
| MED31 | mediator complex subunit 31 |
| MEF | mouse embryonic fibroblast |
| MEF2 | myocyte enhancer factor 2 |
| MEM | Minimal Essential Medium |
| MENT | methylated in normal thymocytes |

| | |
|---|---|
| Mep50 | methylosome protein 50 |
| MET | (start codon for amino acid residue) methionine |
| mg | milligram(s) |
| $MgCl_2$ | magnesium chloride |
| $MgSO_4$ | magnesium sulfate |
| miR/miRNA | micro RNA |
| min | minute(s) |
| ml | millilitre(s) |
| ML | Major Late |
| MLF1 | myeloid leukemia factor 1 |
| MLL/KMT2A | mixed lineage leukemia/ lysine methyltransferase 2A |
| mM | millimolar |
| MMLV-RT | Moloney Murine Leukemia Virus reverse transcriptase |
| MMP1 | matrix metallopeptidase 1 |
| mmol | milli mole(s) |
| $MnCl_2$ | manganese chloride |
| Mot1 | DNA-binding ATPase |
| MRF4/Myf6 | myogenic regulatory factor 4/myogenic factor 6 |
| mRNA | messenger RNA |
| mSin3A | SIN3 transcription regulator family member A |
| mTOR | mechanistic target of rapamycin |
| mut | mutant |
| MXI1 | MAX interactor 1, dimerization protein |
| MYC | MYC proto-oncogene, bHLH transcription factor |
| MyoD/Myod1 | myogenic differentiation/myogenic differentiation 1 |
| Myf5 | myogenic factor 5 |
| Myog | myogenin |
| $N_2$ | nitrogen |
| N6L | NucAnt 6L |
| N.B. | *Nota bene* (note well) |
| N-terminal | amino-terminal |
| NA | not available |
| Na-deoxycholate | sodium deoxycholate |
| Na-MOPS | 4-morpholinepropanesulfonic acid sodium salt |
| $Na_2HPO_4$ | sodium hydrogen phosphate |
| NaCl | sodium chloride |
| $NaHCO_3$ | sodium bicarbonate |
| Nanog | (mouse) Nanog homeobox |
| NANOG | (human) Nanog homeobox |
| NaOAc | sodium acetate |
| Nap1/NAP1 | nucleosome assembly protein |
| NASP (N1/N2) | nuclear autoantigenic sperm protein |
| NBE | NOBOX binding element |
| NC2 | negative cofactor 2 |
| NCBI | National Center for Biotechnology Information |
| NCI | National Cancer Institute |
| NCL | Nucleolin |
| NCP | nucleosome core particle |
| NDF | nucleolus-derived foci |
| NDR | nucleosome-depleted region |
| NELF | negative elongation factor |
| $neo^r$ | neomycin resistance marker gene |
| NER | nucleotide excision repair |
| NES | nuclear export signal |
| NES | normalized enrichment score |
| NF-κB | nuclear factor kappa B |
| NF-κB2 | nuclear factor kappa B subunit 2 |
| NF-Y | nuclear transcription factor Y |
| NFR | nucleosome-free region |
| NFYA | nuclear transcription factor Y subunit alpha |
| NFYB | nuclear transcription factor Y subunit beta |
| ng | nanogram(s) |
| NGF | nerve growth factor |
| Ni-NTA | Nickel-Nitrilotriacetic acid |
| NLP | Nucleoplasmin like protein |
| NLS | nuclear localization signal |

| | |
|---|---|
| nm | nanometre(s) |
| nM | nanomolar |
| No. | number |
| NOBOX | NOBOX oogenesis homeobox, newborn ovary homeobox |
| NoLS | nucleolar localization signal |
| nos. | numbers |
| NP | *Xenopus* Nucleoplasmin |
| NPH | *Drosophila* Nucleophosmin |
| Npm1 | mouse Nucleophosmin |
| NPM1 | human Nucleophosmin |
| Npm2 | mouse Nucleoplasmin |
| NPM2 | human Nucleoplasmin |
| NPM3 | human Nucleoplasmin/Nucleophosmin 3 |
| NR | nuclear receptor |
| NR4A1 | nuclear receptor subfamily 4 group A member 1 |
| Nrd1 | nardilysin, N-arginine dibasic convertase, NRD convertase 1 |
| NRF | nuclear respiratory factor |
| ns | non-significant |
| NSC348884 | (di-[((6-methyl-1H -benzo[d]imidazol-2-yl)methyl)((5-methyl-3-oxo-3H -indol-2-yl)methyl)]) aminoethane |
| NuMA/NUMA | nuclear mitotic apparatus protein |
| NUP214 (CAN) | nucleoporin 214 |
| NuRD | nucleosome remodeling deacetylase |
| Nut-3a | Nutlin-3a |
| $O.D._{600}$ | optical density at 600 nm wavelength |
| Oct4/Pou5f1 | POU domain, class 5, transcription factor 1 |
| OP | operator (sequence) |
| ori | origin of replication |
| OSCC | oral squamous cell carcinoma |
| P-TEFb | positive transcription elongation factor b |
| P/BAF | Polybromo-associated BAF complex |
| p21 | protein $p21^{Cip1}$ (or $p21^{Waf1}$ or CDKN1A) |
| p38MAPK | p38 mitogen-activated protein kinase |
| p53 | tumor protein 53 |
| p300 | E1A binding protein p300 |
| PADI4 | peptidyl arginine deiminase 4 |
| Paf1 | RNA polymerase II-associated factor 1 |
| Paf1C | PAF1 complex |
| PAGE | polyacrylamide gel electrophoresis |
| PanNETs | pancreatic neuroendocrine tumors |
| PARP1 | poly(ADP-ribose) polymerase 1 |
| PARP2 | poly(ADP-ribose) polymerase 2 |
| PARylation | poly(ADP-ribosyl)ation |
| pBR322 | plasmid “Bolivar” and “Rodriguez” 322 |
| PBS | phosphate-buffered saline |
| pc | postcoitus |
| PC4 (SUB1) | positive coactivator 4 (SUB1 regulator of transcription) |
| PCA | principal component analysis |
| Pcf11 | PCF11 cleavage and polyadenylation factor subunit |
| PCNA | proliferating cell nuclear antigen |
| PCR | Polymerase Chain Reaction |
| PDB | Protein Data Bank |
| PDGFA | platelet derived growth factor subunit A |
| PDGFB | platelet derived growth factor subunit B |
| pH | pouvoir hydrogène / power of hydrogen |
| PHD | plant homeodomain |
| PI(3,4,5)P3 | phosphatidyl inositol (3,4,5)-triphosphate |
| PI3K | phosphoinositide 3-kinase |
| PIC | pre-initiation complex |
| PIPES | piperazine-N,N′-bis(2-ethanesulfonic acid) |
| piRNA | Piwi-interacting RNA |
| PKA | protein kinase A |
| PKB/Akt | protein kinase B |
| PKC | protein kinase C |
| PKCζ | protein kinase C zeta |
| PKG | protein kinase G |

| | |
|---|---|
| PKR | protein kinase R/protein kinase RNA-activated |
| Plk1/PLK1 | polo like kinase 1 |
| Plk2/PLK2 | polo like kinase 2 |
| PML | promyelocytic leukemia |
| pmol | pico moles |
| PMSF | phenylmethylsulfonyl fluoride |
| PNBs | prenucleolar bodies |
| POLR2A/RPB1 | RNA polymerase II subunit A |
| POLR2B/RPB2 | RNA polymerase II subunit B |
| POLR2C/RPB3/ Rpb3 | RNA polymerase II subunit C |
| POLR2E | RNA polymerase II subunit E |
| POLR2F/RPB6 | RNA polymerase II subunit F |
| POLR2I | RNA polymerase II subunit I |
| POLR2J/POLR2J1/ RPB11/RPB11-a | RNA polymerase II subunit J |
| POLR2K/RPB10α | RNA polymerase II subunit K |
| POLR2M | RNA polymerase II subunit M |
| poly A | polyadenylation |
| poly(A) | polyadenylation |
| POPOP | 1,4-bis(5-phenyloxazol-2-yl) benzene |
| PP1β/ PPP1CB | protein phosphatase 1 catalytic subunit beta |
| PPO | 2,5-diphenyloxazole |
| PRC2 | polycomb repressor complex 2 |
| PRMT1 | protein arginine methyltransferase 1 |
| PRMT4 | protein arginine methyltransferase 4 |
| PRMT5 | protein arginine methyltransferase 5 |
| Pro | proline |
| pUC | plasmid 'University of California' |
| $Puro^R$ | puromycin resistance (marker gene) |
| $Puromycin^R$ | puromycin resistance (marker gene) |
| PVDF | polyvinylidene difluoride |
| qPCR | quantitative PCR |
| RA | retinoic acid |
| RAD21 | RAD21 cohesin complex component |
| Ran | Ras-related nuclear protein |
| RAR-α | retinoic acid receptor alpha |
| Rat1 | RNA-trafficking protein 1 |
| Rb | retinoblastoma protein |
| RbAp46 | pRB-associated proteins p46 |
| RbAp48 | pRB-associated proteins p48 |
| RCSB | Research Collaboratory for Structural Bioinformatics |
| rDNA | ribosomal DNA |
| RefSeq | NCBI Reference Sequences |
| RFP | red fluorescent protein |
| RFX | regulatory factor X |
| RHD | Rel homology domain |
| RIPA | radioimmunoprecipitation assay |
| Rmix | ribonucleotide master mix |
| RNA | ribonucleic acid |
| RNAi | RNA interference |
| RNA Pol II | RNA Polymerase II |
| RNF8 | ring finger protein 8 |
| RNF20 | ring finger protein 20 |
| RNF40 | ring finger protein 40 |
| RP | reverse primer |
| RPB | RNA Polymerase B/RNA Polymerase II subunit |
| Rpd3S | histone deacetylase RPD3 small |
| RPL5 | ribosomal protein L5 |
| RPL23 | ribosomal protein L23 |
| RPS9 | ribosomal protein S9 |
| rpm | revolutions per minute |
| RRE | rev response element |
| rRNA | ribosomal RNA |
| Rsf1 | remodeling and spacing factor 1 |
| RSK | ribosomal s6 kinase |

| | |
|---|---|
| RSV | Rous sarcoma virus |
| RT | room temperature |
| RT | reverse transcriptase |
| Rtt106 | regulator of Ty1 transposition |
| rtTA3 | reverse tetracycline transactivator 3 |
| s | second(s) |
| S | Svedberg unit |
| S/MARs | scaffold/matrix attachment regions |
| SAGA | Spt-Ada-Gcn5 acetyltransferase |
| SAT1 | spermidine/spermine N1-acetyltransferase 1 |
| SC35/SRSF2 | serine and arginine rich splicing factor 2 |
| SCC | squamous cell carcinoma |
| scr | scrambled |
| SDS | sodium dodecyl sulfate |
| SEC | super elongation complex |
| SEM | standard error of mean |
| Sen1 | senataxin |
| SENP3 | SUMO specific peptidase 3 |
| Ser | serine |
| SERPINE1 | serpin family E member 1 |
| SET | SET nuclear proto-oncogene/Patient SE translocation |
| Set1 | histone lysine methyltransferase Set1 |
| Set2 | histone lysine methyltransferase Set2 |
| shRNA | short hairpin RNA |
| shNPM1 | shRNA against NPM1 |
| shYY1 | shRNA against YY1 |
| si | silencing |
| si-RNA | silencing RNA |
| SIAH1 | siah E3 ubiquitin protein ligase 1 |
| SIN | self-inactivating |
| SIRT1 | sirtuin 1 |
| SIRT6 | sirtuin 6 |
| SIRT7 | sirtuin 7 |
| Sl. | serial |
| SMC3 | structural maintenance of chromosomes 3 |
| SNAI1 | snail family transcriptional repressor 1 |
| snRNA | small nuclear RNA |
| snoRNA | small nucleolar RNA |
| SOD2 | superoxide dismutase 2 |
| SOHLH1 | spermatogenesis and oogenesis specific basic helix-loop-helix 1 |
| SOHLH2 | spermatogenesis and oogenesis specific basic helix-loop-helix 2 |
| Sox2 | SRY (sex determining region Y)-box 2 |
| Sp1 | Sp1 transcription factor |
| Spt2/SPTY2D1 | SPT2 chromatin protein domain containing 1 |
| SPT6 | chromatin-remodeling histone chaperone SPT6 |
| SPT16 | chromatin-remodeling protein SPT16 |
| SRC | SRC proto-oncogene, non-receptor tyrosine kinase |
| SRE | serum response element |
| SRF | serum response factor |
| SREBP | Sterol regulatory element binding protein |
| SSBPs | sperm-specific basic proteins |
| SSRP1 | structure specific recognition protein 1 |
| STAT | signal transducers and activators of transcription |
| STAT1 | signal transducer and activator of transcription 1 |
| STAT3 | signal transducer and activator of transcription 3 |
| STAT5 | signal transducer and activator of transcription 5 |
| STAT5A | signal transducer and activator of transcription 5A |
| SUMO | small ubiquitin-like modifier |
| SV40 | Simian virus 40 |
| SWI/SNF | SWItch/Sucrose Non-Fermentable |
| SWR1 | Swi2/snf2-related 1 |
| T | thymidine |
| TAD | topologically associating domain |
| TAFs | TBP-associated factors |
| Taf1 | TATA-box binding protein associated factor 1 |
| TAF1 (TAFII-250) | TATA-box binding protein associated factor 1 |

| | |
|---|---|
| TAF1A/TAF(I)48 | TATA-box binding protein associated factor, RNA polymerase I subunit A |
| TAF3 | TATA-box binding protein associated factor 3 |
| TAF5L | TATA-box binding protein associated factor 5 like |
| TAF6 | TATA-box binding protein associated factor 6 |
| TAF7L | TATA-box binding protein associated factor 7 like |
| TAF8 | TATA-box binding protein associated factor 8 |
| TAF9 | TATA-box binding protein associated factor 9 |
| Tat | trans-activator of transcription |
| TATA | TATA box (Goldberg-Hogness box) |
| TBE | tris/borate/EDTA |
| TBP | TATA-box binding protein |
| TBS | tris-buffered saline |
| TCA | trichloroacetic acid |
| TCGA | The Cancer Genome Atlas |
| TEMED | N,N,N',N'-tetramethylethane-1,2-diamine |
| Tet | tetracycline |
| TetR | tetracycline repressor |
| TEX19 | testis expressed 19 |
| TF | transcription factor |
| TFIIA | transcription factor II A |
| TFIIB | transcription factor II B |
| TFIID | transcription factor II D |
| TFIIE | transcription factor II E |
| TFIIF | transcription factor II F |
| TFIIF-alpha | transcription factor II F alpha |
| TFIIH | transcription factor II H |
| TFIIS/ SII | transcription factor II S |
| TGF-β | transforming growth factor beta |
| TH2A/H2AC1 | H2A clustered histone 1 |
| TH2B/H2BC1 | H2B clustered histone 1 |
| THAP | THAP domain containing |
| Thr | threonine |
| TIF1β | transcriptional intermediary factor 1 beta |
| Tip60 | Tat interactive protein 60 |
| TIPs | transcription initiation platforms |
| tiRNA | transcription initiation associated RNA |
| TLS | translesion DNA synthesis |
| Tm | melting temperature |
| Tmix | transcription master mix |
| TmPyP4 | tetra-N-methyl-pyridyl porphyrin |
| TNF | tumor necrosis factor |
| TNF-α | tumor necrosis factor alpha |
| TONSL | tonsoku like, DNA repair protein |
| TopBP1 | DNA topoisomerase II binding protein 1 |
| TP2 | transition protein 2 |
| TRE | tetracycline response element |
| TRFs | TBP-related factors |
| tRFP | TurboRFP reporter |
| Tris | trisaminomethane |
| tRNA | transfer RNA |
| TSS | transcription start site |
| TSSmiRNA | transcription start site associated miRNA |
| TTF1 | transcription termination factor 1 |
| TTLL4 | tubulin tyrosine ligase like 4 or tubulin polyglutamylase |
| U | units |
| U | uracil |
| UBC | (human) ubiquitin C promoter |
| UbcH6/ UBE2E1 | ubiquitin conjugating enzyme E2 E1 |
| UCSC | University of California Santa Cruz |
| UM | University of Michigan |
| UPCI | University of Pittsburgh Cancer Institute |
| USF | upstream stimulatory factor |
| USP36 | ubiquitin specific peptidase 36 |
| UT | untreated/untransfected |
| UTP | uridine triphosphate |
| UTR | untranslated region |

| | |
|---|---|
| V | volts |
| v/v | volume by volume |
| VDR | vitamin D receptor |
| VEGFA | vascular endothelial growth factor A |
| Vps75 | vacuolar protein sorting 75 |
| w/v | weight by volume |
| WashU | Washington University |
| WHD | winged-helix domain |
| WPRE | woodchuck (hepatitis virus) post-transcriptional regulatory element |
| WT | wild-type |
| XCPE1 | X core promoter element 1 |
| XCPE2 | X core promoter element 2 |
| WDR43 | WD repeat domain 43 |
| XPB | xeroderma pigmentosum type B |
| Xrn2 | 5'-3' exoribonuclease 2 |
| YAP1 | Yes associated protein 1 |
| YTR107 | (Z) −5 - ((N-benzyl-1H-indol-3-yl) methylene) pyrimidine-2, 4, 6 (1H, 3H, 5H) trione |
| YY1 | YY1 transcription factor (or Yin Yang 1) |
| ZC3H12A | zinc finger CCCH-type containing 12A |
| Znf | zinc finger |

# *Appendix*

**Table A.1. Previously generated/requested plasmids and constructs used in the study**.

| Sl. No. | Construct | Comments | Source/Reference |
|---|---|---|---|
| 1 | NPM1-$His_6$ | Bacterial expression clone of C-terminal $His_6$-tagged WT human NPM1 in pET28b vector (Novagen). | (Swaminathan et al. 2005). |
| 2 | $His_6$-NPM2 | Bacterial expression clone of N-terminal $His_6$-tagged WT human NPM2 in pET14b vector (Novagen). | Kind gift from Dr. Mitsuru Okuwaki, University of Tsukuba, Japan (Okuwaki et al. 2012). |
| 3 | NPM3-$His_6$ | Bacterial expression clone of C-terminal $His_6$-tagged WT human NPM3 in pET28b vector (Novagen). | (Gadad et al. 2010). |
| 4 | NP | Bacterial expression clone of untagged *Xenopus* Nucleoplasmin (NP) in pET11b vector. | Kind gift from Prof. Arturo Muga of Biofisika, Spain (Hierro et al. 2001). |
| 5 | $His_6$-TFAP4 | Bacterial expression clone of N-terminal $His_6$-tagged human TFAP4 in pET28b vector (Novagen). | (Senapati P, Ph.D. thesis, 2014). |
| 6 | POLR2K-$His_6$ | Bacterial expression clone of C-terminal $His_6$-tagged human POLR2K in pET28b vector (Novagen). | (Senapati P, Ph.D. thesis, 2014). |
| 7 | PC4-$His_6$ | Bacterial expression clone of C-terminal $His_6$-tagged human PC4. | (Batta and Kundu 2007). |
| 8 | SNAI1-$His_6$ | Bacterial expression clone of C-terminal $His_6$-tagged human SNAI1 in pET28b vector (Novagen). | (Senapati P, Ph.D. thesis, 2014). |
| 9 | H2A<br>H2B<br>H3<br>H4 | Bacterial expression clones of untagged *Xenopus* core histones in pET3d vector. | Kind gifts from Dr. K. Luger (University of Colorado, USA) (Luger et al. 1999). |
| 10 | FLAG-c-fos | Mammalian expression clone of N-terminal FLAG-tagged human c-fos clone in the pFLAG-CMV2 vector (Sigma). | (Senapati P, Ph.D. thesis, 2014). |

| | | | |
|---|---|---|---|
| 11 | c-jun | Mammalian expression construct of untagged c-jun. | Kind gift from Dr. Sagar Sengupta (National Institute of Immunology, New Delhi, India). |
| 12 | HA-YY1 | Mammalian expression clone of HA-tagged human YY1 in the pcDNA3 vector. | Kind gift from Prof. Purnima Dubey (Wake Forest School of Medicine, North Carolina, USA). |
| 13 | WT p53 | Human wild-type p53 (pCMV-wtp53) expression plasmid. | Kind gift from Prof. Bert Vogelstein (Johns Hopkins University, Baltimore, USA). |
| 14 | FLAG-WT p53, FLAG-mutant p53 | Mammalian expression clones of N-terminal FLAG-tagged WT and mutant p53 in pFLAG-CMV2 and p3xFLAG-CMV10 vectors (Sigma). Specific cloning and site-directed mutagenesis primers are listed in Table A.2. | (Kaypee S, Ph.D. thesis, 2017), (Kaypee et al. 2018a). |
| 15 | 3xFLAG-R175H p53 | Mammalian expression, doxycycline-inducible, N-terminal 3xFLAG-tagged R175H p53 clone in pEBTetD vector generated by sub-cloning from pEBTetD SLC22A1 construct. The primers used for this cloning (Table A.2) were the same as the subcloning of WT p53 insert into the pEBTetD vector previously done in the laboratory. | pEBTetD SLC22A1 vector was a kind gift from Dr. Dirk Gründemann (University of Cologne, Germany)(Bach et al. 2007). (Kaypee et al. 2018a), (Kaypee S, Ph.D. thesis, 2017). |
| 16 | NPM1 promoter-Luc | Mammalian expression clones of different promoter fragments of the NPM1 gene in pGL3 basic vector (Promega). | (Senapati P, Ph.D. thesis, 2014). |

**Table A.2. List of cloning primer sequences** used for generating different expression constructs as indicated. Restriction sites have been underlined.

| Sl. No. | Construct | Forward primer (5' – 3') | Reverse primer (5' – 3') |
|---|---|---|---|
| 1 | NPM2-$His_6$ | CATGCCATGGGAAACCTGAGCTCCG | CCGCTCGAGCTTTTTGAAGCCCGG |
| 2 | FLAG-NPM2 | CCCAAGCTTATGAACCTGAGCTCC | CCGCTCGAGCTACTTTTTGAAGC |
| 3 | $His_6$-NPM2 fragment | GGAATTCCATATGCTGGAGGGGAAG | CCCAAGCTTTCAGGCCTGGAG |
| 4 | $His_6$-NP | GGGAATTCCATATGATGGCCTCTAC | CCCAAGCTTTCACTTCTTAGCAGC |

| | | | |
|---|---|---|---|
| 5 | $His_6$-Npm2 | GGAATTCCATATGATGAGTCGCCAC | CCCAAGCTTTCATTTCTTGGTCACTG |
| 6 | 3xFLAG-NPM2 | CCCAAGCTTATGAATCTCAGTAGCGC | CGGGATCCTCATTTCTTGAATCCTGG |
| 7 | 3xFLAG-WT p53 (in p3xFLAG-CMV10 vector) | CCCAAGCTTATGGAGGAGCCGCAGTC | CGCGGATCCTCAGTCTGAGTCAGGC |

**Table A.3. List of primers used for site-directed mutagenesis:** Serial nos. 1 – 6 are related to the mutants generated in $His_6$-NPM2 bacterial expression construct. Serial nos. 7 – 10 are related to the mutations generated at the c-fos/AP-1 binding sites in the NPM1 promoter construct Luc 2 (Senapati P, Ph.D. thesis, 2014). Serial nos. 11 – 12 are related to the mutations generated in the FLAG-c-fos mammalian expression construct (Senapati P, Ph.D. thesis, 2014).

| Sl. No. | Construct | Forward primer (5' – 3') | Reverse primer (5' – 3') |
|---|---|---|---|
| 1 | S174D $His_6$-NPM2 | CAGTTTCTTCTTTTTGGCTACATCCGCTTGTTTCTGCGGAACTAAC | GTTAGTTCCGCAGAAACAAGCGGATGTAGCCAAAAAGAAGAAACTG |
| 2 | S196D $His_6$-NPM2 | CGCTTTCTTGACTGGATCTTTATCGCGAACTGATGCACGGATTTCCT | AGGAAATCCGTGCATCAGTTCGCGATAAAGATCCAGTCAAGAAAGCG |
| 3 | S174E $His_6$-NPM2 | CCAGTTTCTTCTTTTTGGCTACCTCCGCTTGTTTCTGCGGAACTAAC | GTTAGTTCCGCAGAAACAAGCGGAGGTAGCCAAAAAGAAGAAACTGG |
| 4 | S196E $His_6$-NPM2 | CTTTCTTGACTGGCTCTTTATCGCGAACTGATGCACGGATTTC | GAAATCCGTGCATCAGTTCGCGATAAAGAGCCAGTCAAGAAAG |
| 5 | S174A $His_6$-NPM2 | CTTCTTTTTGGCTACAGCCGCTTGTTTCTGCGGAA | TTCCGCAGAAACAAGCGGCTGTAGCCAAAAAGAAG |
| 6 | S196A $His_6$-NPM2 | CGTGCATCAGTTCGCGATAAAGCGCCAGTCAAG | CTTGACTGGCGCTTTATCGCGAACTGATGCACG |
| 7 | AP1BS1mut | GTCCTTGCTAATTTGGAGACAGTTGCCGTCCCCTTTTGGCCCCCAAG | CTTGGGGGCCAAAAGGGGACGGCAACTGTCTCCAAATTAGCAAGGAC |
| 8 | AP1BS2mut | GAATCGAGGTGCTCTCGGCCACTTTCGCAGCCGGCTAAC | GTTAGCCGGCTGCGAAAGTGGCCGAGAGCACCTCGATTC |
| 9 | AP1BS3mut | GTCTTCCTTTCTGAGGCTATCATTTGTATACCACTCTTCTTAAATTTGTTTGATATGT | ACATATCAAACAAATTTAAGAAGAGTGGTATACAAATGATAGCCTCAGAAAGGAAGAC |
| 10 | AP1BS4mut | GTCCTTTCCCTGGTGAGTTGCCGTCCTGCGCGGTT | AACCGCGCAGGACGGCAACTCACCAGGGAAAGGAC |
| 11 | Cfosmut1 (K153Q, R155Q, R157Q, R158Q and R159Q) | GCAGCCCAATGCCAAAACCAGCAGCAGGAGCTGACTGATACACTCCAAG | GCTCCTGCTGCTGGTTTTGGCATTGGGCTGCAGCCATCTTATTCCTTTC |
| 12 | Cfosmut2 (L179V, L186A and L193V) | GCTGTGCAGACCGAGATTGCCAACGCGCTGAAGGAGAAGGAAAAAGTAGAGTTCATCCTG | TACTTTTTCCTTCTCCTTCAGCGCGTTGGCAATCTCGGTCTGCACAGCAGACTTCTCATC |

**Table A.4. Mammalian cell lines used in the study.**

| Sl. No. | Cell line | Culture medium requirements | Comments/Source |
|---|---|---|---|
| 1 | HEK-293 (Cat. No. CRL-1573) | DMEM (Sigma)+ 35 mM NaHCO3 (Sigma) | American Type Culture Collection (ATCC, Manassas, VA, USA) |
| 2 | HeLa S3 (Cat. No. CCL-2.2) | | |
| 3 | HepG2 (Cat. No. HB-8065) | MEM (Gibco/HiMedia) + 1X Glutamax (Gibco)/HiGlutaXL (HiMedia) | |
| 4 | MCF7 (Cat. No. HTB-22) | MEM (Gibco/HiMedia) + 1X Glutamax (Gibco)/HiGlutaXL (HiMedia) + 0.01 mg/ml human recombinant insulin (Sigma) | |
| 5 | NCI-H1299 (Cat. No. CRL-5803) | RPMI-1640 (HiMedia) + HiGlutaXL (HiMedia) | |
| 6 | AU565 (Cat. No. CRL-2351) | | |
| 7 | HCT 116p53$^{+/+}$ (Cat. No. CCL-247) | Mc-Coy's-5a medium (HiMedia) | |
| 8 | UPCI:SCC-29B | MEM (Gibco/HiMedia) + 1X Glutamax (Gibco)/HiGlutaXL (HiMedia) | Human oral cancer cell lines; kind gifts from Prof. Susanne M. Gollin (University of Pittsburgh, USA) |
| 9 | UPCI:SCC-40 | | |
| 10 | UPCI:SCC-122 | | |
| 11 | AW13516 | | Human oral cancer cell lines; kind gifts from Dr. Amit Dutt (ACTREC, Mumbai, India). |
| 12 | NT8e | | |
| 13 | OT9 | | |
| 14 | AW8507 | | |
| | | DMEM (Sigma)+ 35 mM NaHCO3 (Sigma) | |
| 15 | esophagus normal Het-1A | Bronchial Epithelial Cell Growth Medium (BEGM) (Lonza/Clonetics Corporation, Basel, Switzerland) ***Note**: No serum used in the culture medium* | |
| 16 | UM-SCC-1 | DMEM (Sigma) + 35 mM NaHCO3 | HNSCC cell line; kind gift from Dr. Gautam Sethi (National University of Singapore). |
| 17 | A2780 (cisplatin resistant) | RPMI-1640 (HiMedia) + HiGlutaXL (HiMedia) | Kind gift from Dr. Alan Prem Kumar (Cancer Science Institute, National University of Singapore). |
| 18 | C2C12 | DMEM + 35 mM NaHCO3 | Mouse myoblast cell line; kind gift from Dr. Reshma Taneja (National University of Singapore). |

| | | | |
|---|---|---|---|
| 19 | Hep3B | MEM (Gibco/HiMedia) + 1X Glutamax (Gibco)/HiGlutaXL (HiMedia) | Cell Repository (National Centre for Cell Science, Pune, India). |
| 20 | SK-Hep-1 | | |
| 21 | PLC/PRF/5 | | |
| 22 | Huh1 | DMEM (Sigma)+ 35 mM NaHCO3 (Sigma) | Kind gift from Prof. Saumitra Das (Indian Institute of Science, Bangalore, India). |

**Table A.5. List of primers used in this study for mRNA expression analysis by RT-qPCR.**

| Sl. No. | Gene Target | Forward primer (5' – 3') | Reverse primer (5' – 3') |
|---|---|---|---|
| 1 | *Npm2* (mouse) | TCAGTGGCCTGGAATGTTATG | GCCTCACTACTGTTTCATCCTC |
| 2 | *Npm1* (mouse) | GAGGCTATTCAAGATCTCTGGC | CCAAGTAAAGGGTGGAGTTCC |
| 3 | *Actb* (mouse) (Arun et al. 2012) | AGGTCATCACTATTGGCAACG | TACTCCTGCTTGCTGATCCAC |
| 4 | *MyoD* (mouse) | ATCCGCTACATCGAAGGTCT | CGCTGTAATCCATCATGCCA |
| 5 | *Myogenin* (mouse) | CAGTGAATGCAACTCCCACA | CGAGCAAATGATCTCCTGGG |
| 6 | *Fos* (mouse) | CTGAAGAGGAAGAGAAACGGAG | GGCTGCCAAAATAAACTCCAG |
| 7 | *NPM2* | ATGAAGCATCAGACCTAACCTG | GCTGGCTCTTATTTCCTCTTCT |
| 8 | *ACTB* (β-actin) | AGATGTGGATCAGCAAGCAGGAGT | TCCTCGGCCACATTGTGAACTTTG |
| 9 | *NPM1* primer # 1 | GTGAAGAAATCTATACGAGATACTCCAGCC | CTTCCACTTTGGGAAGAGAACCACC |
| 10 | *NPM1* primer # 2 | GTTCAGGGCCAGTGCATATTA | TTTCTGTGGAACCTTGCTACC |
| 11 | *FOS* (c-fos) | GACTGATACACTCCAAGCGG | CATCAGGGATCTTGCAGGC |
| 12 | *JUN* (c-jun) | TCCTTAAGAACACAAAGCGGG | ACACAGTTAACGAAGGCAGG |
| 13 | *CCNA2* | CGAAAGACTGGATATACCCTGG | CATCTTAGAAAACAAAGGCAGTCT |
| 14 | *CCNB2* | GGCTGGTACAAGTCCACTCC | GAAGCCAAGAGCAGAGCAGT |
| 15 | *CDC25C* | TTCTGGTGAAGGACATGAGC | GGCCTGGATACAAGTTGGTAG |
| 16 | *MYC* (c-myc) | ATGTCCTGAGCAATCACCTATG | CAAAGTCCAATTTGAGGCAGTT |
| 17 | *CDK1* | AACTTGGATGAAAATGGCTTGG | AAGAGTTAACAATAAAAACACAACTATCTG |
| 18 | *CDKN1A* (p21) | GAACTTCGACTTTGTCACCG | TGGAGTGGTAGAAATCTGTC |
| 19 | *PDGFA* | GATACCTCGCCCATGTTCTG | GTCCAAAGAATCCTCACTCCC |
| 20 | *PDGFB* | CACCGGAAATTCAAGCACAC | CGAATCAGGCATCGAGACAG |
| 21 | *VEGFA* | AGTCCAACATCACCATGCAG | CCTTTCCCTTTCCTCGAACTG |
| 22 | *CCND1* | CATCTACACCGACAACTCCATC | GTTCAATGAAATCGTGCGGG |
| 23 | *CCND2* | GAGCTGCTGGCTAAGATCAC | ATATCCCGCACGTCTGTAGG |
| 24 | *CD44* | CAGATGGCATGAGGGATATCG | GGGTGTGAGATTGGGTTGAAG |
| 25 | *MMP1* | AAGACAGATTCTACATGCGCACAA | CTGTCCCTGAACAGCCCAGT |
| 26 | *CXCL8* | ATACTCCAAACCTTTCCACCC | AAACTTCTCCACAACCCTCTG |
| 27 | *SERPINE1* | TGCAGAAAGTGAAGATCGAGG | CACAAAGAGGAAGGGTCTGTC |

| | | | |
|---|---|---|---|
| 28 | *SAT1* | CTGATCAAGGAGCTGGCTAAA | ATCCACGGGTCATAGGTAAAATAG |
| 29 | *HMOX1* | ACTGCGTTCCTGCTCAACAT | GGGGCAGAATCTTGCACTTT |
| 30 | *F3* | AACACTTTCCTAAGCCTCCG | TACTCTTCCGGTTAACTGTTCG |
| 31 | *FN1* | GTGGCAGAAGGAATATCTCGG | GCATGAAGCACTCAATTGGG |
| 32 | *TNF* | TAGCCCATGTTGTAGCAAACC | ATGAGGTACAGGCCCTCTGAT |
| 33 | *HBEGF* | GAGAAAGTGAAGTTGGGCATG | GGCAAAGCAATTATGGGAGG |
| 34 | *COL16A1* | GGGACATTGGTATTGGCATTG | GCAGTCAGAGGGATTACAGTG |
| 35 | *ANGPTL4* | AGACACAACTCAAGGCTCAG | CTCATGGTCTAGGTGCTTGTG |
| 36 | *ZC3H12A* | TTGTGAAGCTGGCCTACGAG | TCAGGGGGCATAAACTTGTCA |
| 37 | *NR4A1* | CTTGTCCTCATCACCGACCG | TGCCCAACAGACGTGACAGG |
| 38 | *SOD2* | GGGTTGGCTTGGTTTCAATAAG | TGCTCCCACACATCAATCC |
| 39 | *EGR1* | ACAGCAACCTTTTCTCCCAG | CCAATAGACCTTCCACTCCA |
| 40 | *18S* | GTAACCCGTTGAACCCCATT | CCATCCAATCGGTAGTAGCG |
| 41 | *YY1* | TGGTCCTCAGATGAAAAAAAAGATATTGAC | GGCTTCATTCTAGCAAATTCTGCC |
| 42 | *NANOG* | CACGGAGACTGTCTCTCCTC | GAGAGTTCTTGCATCTGCTG |
| 43 | *GAPDH* | TCCACCTTTGACGCTGGGGCTGGC | TGGCAGGGACTCCCCAGCAGTGAG |

**Table A.6. List of commercial antibodies used in this study.**

| Sl. No. | Name of the antibody | Brand | Cat. No. | Comments |
|---|---|---|---|---|
| 1 | anti-NPM2 | Sigma | SAB1400381 | Mouse polyclonal antibody. |
| 2 | anti-FLAG M2 | Sigma | F1804 | Mouse monoclonal antibody; has been used for ChIP assay. |
| 3 | anti-polyHistidine | Sigma | H1029 | Mouse monoclonal antibody. |
| 4 | anti-α-tubulin (DM1A) | Calbiochem | CP06 | Mouse monoclonal antibody. |
| 5 | anti-c-fos | Santa Cruz Biotechnology | sc-52 | Rabbit polyclonal antibody; has been used for ChIP assay. |
| 6 | anti-c-jun | Abcam | ab7964 | Rabbit polyclonal antibody; kind gift from Dr. Sagar Sengupta (National Institute of Immunology, New Delhi, India). |
| 7 | anti-p53 | Calbiochem | OP43 | Mouse monoclonal antibody; pantropic |
| 8 | anti-YY1 | Abcam | ab12132 | Rabbit polyclonal antibody. |
| 9 | Secondary HRP-conjugated antibodies | Genei, Sigma, Abcam | ab97051, ab97023 | Goat anti-rabbit and anti-mouse IgG secondary antibodies used for western blotting. |
| 10 | F(ab')2-goat anti-rabbit IgG (H + L) cross-adsorbed secondary | Invitrogen | A11070 | Goat anti-rabbit and anti-mouse IgG secondary |

| | | | | |
|---|---|---|---|---|
| | antibody, Alexa Fluor 488 | | | antibodies used for immunofluorescence. |
| 11 | goat anti-mouse IgG (H + L) highly cross-adsorbed secondary antibody, Alexa Fluor 633 | Invitrogen | A21052 | |
| 12 | F(ab')2-goat anti-mouse IgG (H + L) cross-adsorbed secondary antibody, Alexa Fluor 488 | Invitrogen | A11017 | |

**Table A.7. List of primers used in this study for ChIP-qPCR analysis.**

| Sl. No. | Target region | Forward primer (5' – 3') | Reverse primer (5' – 3') |
|---|---|---|---|
| 1 | AP1BS1 (NPM1 promoter) | TCTTACAAGTCACCCGCTTTC | TGTAGTTACCGGCCAGACTTA |
| 2 | AP1BS3 (NPM1 promoter) | ACTGTTCATTCCTCTCTTGATAGAC | AGCTACACCTTGACAAACTCC |
| 3 | AP1BS4 (NPM1 promoter) | AGGACGGCTACGGTACG | ACGCACTTAGGTAGGAGAGAA |
| 4 | cfos_P1 (NPM1 promoter) | GAGACGGGATTTCTCCATGTT | CATTACCTTGAGTCATGTTGTCATT |
| 5 | cfos_P2 (NPM1 promoter) | ACCCAGGCTAGAGAGTAGTG | AAATTAGCCGGGTGTGGTAG |
| 6 | cfos_P3 (NPM1 promoter) | CGTGCTTCGGCCAGTTA | GCCTCTTAACATTTCCCACTTC |
| 7 | CCNA2 (CCNA2 promoter) | GAGTCAGCCTTCGGACAGCC | CCAGAGATGCAGCGAGCAGC |
| 8 | CDC25C (CDC25C promoter) | GAATGGACATCACTAGTAAGGCGCG | GCAGGCGTTGACCATTCAAACCTTC |
| 9 | CDK1 (CDK1 promoter) | GAACTGTGCCAATGCTGGGA | GCAGTTTCAAACTCACCGCG |
| 10 | Negative control | CAGAAAGGAAGGAGCCACAA | TAGCAGGGTGGGAACTCTAA |
| 11 | NR4A1 (NR4A1 promoter) | ACCAAGTTCAGCTTGTGGAG | AGTTAGAGCCCTCGCTTAGT |
| 12 | VEGFA (VEGFA promoter) | CTACTGTCTCCAGACCCTACC | CTTCTCCGCAGAGAGAAATGAA |
| 13 | SERPINE1 (SERPINE1 promoter) | CAGACAAGGTTGTTGACACAAG | GACTCTTGGTCTTTCCCTCATC |
| 14 | FOS (c-fos promoter) | CTCAGCGCAGATTTGAGTTC | CGTGCGAAGACATTTGAAGG |
| 15 | PDGFB (PDGFA promoter) | GGCGAAGGTAATGAATGAAGAAC | CCGGAGTCGGCATGAATC |
| 16 | JUN (c-jun promoter) | CAACACGCACACGCTTAAC | CGAGTACTACTGCGTGACTTTAT |
| 17 | EGR1 (EGR1 promoter) | GGAACAGCCTTTCGGTT | GCTGGGAAATTGAGGATAGGAA |

| | | | |
|---|---|---|---|
| 18 | ANGPTL4 (ANGPTL4 promoter) | AATGTGGTCCAGCCCTTTAG | TTCTCAGGCAGGTGGAGATA |
| 19 | SOD2 (SOD2 promoter) | GGCAGGATTTGCACACATTTAT | AGTGAGGAAGGTGGGAACTA |

**Table A.8. Clinico-pathological information of oral cancer patients.** (NA: Not available)

| Sl. No. | Patient ID | Gender | Age (at the time of surgery) (years) | State of differentiation | Grade of tumor |
|---|---|---|---|---|---|
| 1 | 66911 | Male | 40 | Moderately differentiated | II |
| 2 | 66951 | Male | NA | Well-differentiated | I |
| 3 | 66970 | NA | NA | NA | NA |
| 4 | 70531 | Male | 87 | Well-differentiated | I |
| 5 | 72990 | NA | NA | NA | NA |
| 6 | 74597 | Male | 51 | Poorly differentiated | III |
| 7 | 74805 | Female | 37 | Poorly differentiated | III |
| 8 | 75556 | NA | NA | NA | NA |
| 9 | 79928 | NA | NA | NA | NA |
| 10 | 79962 | Female | 50 | Moderately differentiated | II |
| 11 | 80206 | Male | 43 | Well-differentiated | I |
| 12 | 81056 | Male | 54 | Moderately differentiated | II |
| 13 | 81406 | NA | NA | NA | NA |
| 14 | 907559 | Female | 78 | NA | NA |
| 15 | 908855 | Female | 72 | Well-differentiated | I |
| 16 | 917986 | NA | NA | NA | NA |
| 17 | 920669 | Female | 75 | Moderately differentiated | II |
| 18 | 922107 | Male | 48 | Well-differentiated | I |
| 19 | 56249 | NA | NA | NA | NA |
| 20 | 56271 | NA | NA | NA | NA |
| 21 | 56273 | Female | NA | NA | NA |
| 22 | 5841 | NA | NA | NA | NA |
| 23 | 58075 | Male | NA | NA | NA |
| 24 | 58457 | NA | NA | NA | NA |
| 25 | 81239 | NA | NA | NA | NA |
| 26 | 81673 | Female | 60 | Well-differentiated | I |
| 27 | 81908 | Female | 55 | Well-differentiated | I |
| 28 | 84429 | Female | 60 | Well-differentiated | I |
| 29 | 85576 | NA | NA | NA | NA |
| 30 | 924073 | Female | 60 | Well-differentiated | I |
| 31 | 924178 | Female | 65 | Moderately differentiated | II |
| 32 | 930713 | Female | 60 | Well-differentiated | I |
| 33 | 934927 | Female | 66 | Moderately differentiated | II |
| 34 | 998486 | Male | 58 | Well-differentiated | I |
| 35 | 1165 | Female | 40 | NA | NA |
| 36 | 62265 | Female | 65 | Moderately differentiated | II |
| 37 | 77147 | Male | 52 | NA | NA |
| 38 | 89014 | Female | 45 | NA | NA |
| 39 | 1015497 | Female | 65 | Well-differentiated | I |

| | | | | | |
|---|---|---|---|---|---|
| 40 | 955865 | Male | 70 | Well-differentiated | I |
| 41 | 957783 | Female | 60 | Well-differentiated | I |
| 42 | 961721 | Male | 86 | Well-differentiated | I |
| 43 | 962579 | Male | 45 | Well-differentiated | I |
| 44 | 1004604 | Female | 65 | Well-differentiated | I |
| 45 | 1009225 | Female | 55 | NA | NA |
| 46 | 1018358 | Female | 40 | Moderately differentiated | II |
| 47 | 74288 | NA | NA | NA | NA |

**Table A.9. Types of cancers with available transcriptomics data in TCGA which have been considered for analyses in the present study.**

| Sl. No. | Cancer type (abbreviation) | Histological origin |
|---|---|---|
| 1 | Acute Myeloid Leukemia (LAML) | leukocytes (myeloid lineage) |
| 2 | Adrenocortical Carcinoma (ACC) | adrenal gland and paraganglia |
| 3 | Bladder Urothelial Carcinoma (BLCA) | urinary bladder |
| 4 | Brain Lower Grade Glioma (LGG) | glial cells of CNS |
| 5 | Breast Invasive Carcinoma (BRCA) | breast |
| 6 | Cervical Squamous Cell Carcinoma and Endocervical Adenocarcinoma (CESC) | cervix |
| 7 | Cholangiocarcinoma (CHOL) | liver and intrahepatic bile ducts |
| 8 | Colorectal Adenocarcinoma (COAD) | colon |
| 9 | Esophageal Carcinoma (ESO) | esophagus |
| 10 | Glioblastoma Multiforme (GBM) | glial cells of the brain |
| 11 | Head and Neck Squamous Cell Carcinoma (HNSC) | epithelium lining the sinonasal tract, oral cavity, pharynx, and larynx |
| 12 | Kidney Chromophobe (KICH) | distal convoluted tubules and cortical collecting ducts of the kidney |
| 13 | Kidney Renal Cell Clear Carcinoma (KIRC) | the lining of the proximal convoluted tubule of the kidney |
| 14 | Kidney Renal Papillary Cell Carcinoma (KIRP) | proximal convoluted tubules of kidney nephrons |
| 15 | Liver Hepatocellular Carcinoma (LIHC) | liver |
| 16 | Lung Adenocarcinoma (LUAD) | lung |
| 17 | Lung Squamous Cell Carcinoma (LUSC) | lung |
| 18 | Lymphoid Neoplasm Diffuse Large B-cell Lymphoma (DLBC) | B-lymphocytes |
| 19 | Mesothelioma (MESO) | the mesothelial lining of pleura, peritoneum, pericardium and tunica vaginalis |
| 20 | Ovarian Serous Cystadenocarcinoma(OV) | ovary |
| 21 | Pancreatic Adenocarcinoma (PAAD) | pancreas |
| 22 | Pheochromocytoma and Paraganglioma (PCPG) | the autonomic nervous system, derived from chromaffin tissue in the adrenal medulla or extra-adrenal ganglia |
| 23 | Prostate Adenocarcinoma (PRAD) | prostate gland |
| 24 | Sarcoma (SARC) | mesenchyme (connective tissue) |
| 25 | Skin Cutaneous Melanoma (SKCM) | non-glabrous skin |

| | | |
|---|---|---|
| 26 | Stomach Adenocarcinoma (STAC) | stomach |
| 27 | Testicular Germ Cell Cancer (TGCT) | testis |
| 28 | Thymoma (THYM) | epithelial cells of the thymus gland |
| 29 | Thyroid Carcinoma (THCA) | thyroid |
| 30 | Uterine Carcinosarcoma (UCS) | uterus |
| 31 | Uterine Corpus Endometrial Carcinoma (UCEC) | uterine endometrium |
| 32 | Uveal Melanoma (UVM) | melanocytes in uvea (within the eye) |

**Table A.10. List of transcription factor-binding motifs predicted using the TRANSFAC database.** Only a single high-confidence site for each transcription factor is listed.

| Sl. No. | Matrix | Factor name | Position (strand) | Core score | Matrix score | Sequence |
|---|---|---|---|---|---|---|
| 1 | V$AP1_Q6_02 | AP-1 | 499 (+) | 0.936 | 0.942 | TGATTcag |
| 2 | V$ARID5A_03 | Arid5a | 354 (+) | 1 | 0.96 | caAATATtcgaatt |
| 3 | V$CREBP1_01 | ATF-2 group | 526 (+) | 1 | 1 | TTACGtaa |
| 4 | V$BCL6_Q3_01 | BCL-6 factors | 619 (+) | 0.984 | 0.937 | cTTCCTaaca |
| 5 | V$BEN_01 | BEN | 978 (+) | 1 | 0.983 | CAGCGgag |
| 6 | V$BLIMP1_Q4 | Blimp-1 | 1205 (-) | 1 | 0.957 | CTTTCcctggtg |
| 7 | V$BRCA_01 | BRCA1 | 231 (+) | 1 | 0.978 | ataTGTTG |
| 8 | V$CEBPA_Q6 | C/EBP group | 523 (+) | 0.975 | 0.97 | aagttaCGTAAag |
| 9 | V$CP2_Q6 | CP2-related factors | 650 (+) | 1 | 0.981 | ctcTCTGGca |
| 10 | V$E2F_Q6_01 | E2F related factors | 875 (-) | 1 | 0.928 | tctTGGCGggag |
| 11 | V$EGR1_Q6 | EGR1 group | 1047 (+) | 1 | 0.953 | gtGGGGGcga |
| 12 | V$ETS_Q6 | Ets-related factors | 560 (-) | 1 | 0.987 | caGGAAGg |
| 13 | V$HNF3B_Q6 | FOX factors | 224 (+) | 1 | 0.967 | tTGTTTgat |
| 14 | V$GEN_INI_B | general initiator | 209 (+) | 0.99 | 0.99 | cctCACTT |
| 15 | V$GLI_Q3 | GLI group | 450 (-) | 1 | 0.945 | gcCCACCctc |
| 16 | V$GMEB2_04 | GMEB | 526 (-) | 1 | 1 | tTACGTaa |
| 17 | V$HES1_Q6 | HES-1 group | 1126 (+) | 1 | 0.958 | agCACGCgcg |
| 18 | V$HIF1A_Q5 | HIF1 | 909 (+) | 1 | 0.991 | ggACGTGga |
| 19 | V$RUSH1A_02 | hltf | 534 (-) | 1 | 0.998 | agATAAGgac |
| 20 | V$HMGIY_Q3 | HMGA factors | 708 (-) | 1 | 0.949 | gtaaaAAATTcctga |

| | | | | | | |
|---|---|---|---|---|---|---|
| 21 | V$HNF4A_Q3 | HNF-4 group | 661 (-) | 1 | 0.931 | ctgaaCTTTGgggt |
| 22 | V$HNF1A_Q4 | HNF1-like factors | 671 (+) | 1 | 0.965 | gggtaacgATTAActg |
| 23 | V$HSF1_01 | HSF dimer | 1307 (+) | 0.974 | 0.977 | GGAAGattcg |
| 24 | V$EKLF_Q5_01 | KLF1 group | 985 (-) | 1 | 0.99 | gGGGTGgggc |
| 25 | V$LRH1_Q5_01 | LRH-1 group | 478 (-) | 1 | 0.976 | tgtCCTTGcta |
| 26 | V$MAF_Q4 | MAF group | 498 (-) | 0.93 | 0.912 | cTGATTcagt |
| 27 | V$MAZ_Q6_01 | MAZ | 979 (+) | 1 | 0.966 | agcGGAGGgtggg |
| 28 | V$MAZR_01 | MAZ group | 984 (+) | 0.93 | 0.952 | aggggTGGGGcca |
| 29 | V$MUSCLEINI_B | muscle initiator | 282 (-) | 1 | 0.947 | gcccgcgGGGTGctgggg tgc |
| 30 | V$MYB_Q4 | Myb-like factors | 17 (+) | 1 | 0.985 | ggaaCAGTTaaa |
| 31 | V$NF1_Q6 | NF-1 factors | 875 (+) | 1 | 0.984 | tcTTGGCgggaggccggc |
| 32 | V$NKX25_Q6 | Nkx group | 718 (-) | 1 | 0.969 | cctgAAGTGat |
| 33 | V$P53_Q3 | p53 related factors | 741 (-) | 0.993 | 0.964 | gagGCTTGcag |
| 34 | V$REST_Q5 | REST | 286 (+) | 1 | 0.916 | gcggggTGCTGgg |
| 35 | V$SF1_Q5_01 | SF-1 group | 478 (+) | 1 | 0.971 | tgtCCTTGc |
| 36 | V$SOX2_Q3_01 | Sox-related factors | 575 (+) | 1 | 0.99 | tgaaaaACAAAgttc |
| 37 | V$SOX10_Q3 | Sox10 | 626 (+) | 0.992 | 0.985 | ACAAAgaa |
| 38 | V$SP1_Q6_01 | Sp1 group | 985 (+) | 0.945 | 0.959 | ggGGTGGggc |
| 39 | V$SREBP_Q6 | SREBP factors | 1209 (-) | 0.992 | 0.985 | ccctggTGTGAttcc |
| 40 | V$STAT1_Q6 | STAT factors | 145 (-) | 0.991 | 0.98 | tTTCCTgtaa |
| 41 | V$TATA_01 | TBP-related factors | 1179 (+) | 0.936 | 0.939 | atATATAagcgcggg |
| 42 | V$TBX5_01 | TBX5 | 315 (+) | 1 | 0.963 | caaGGTGTagct |
| 43 | V$TEF1_Q6_04 | Tef-1-related factors | 407 (-) | 1 | 0.996 | aggaGGAATgt |
| 44 | V$XVENT1_01 | Xvent-1 | 351 (-) | 1 | 0.902 | tccCAAATattcg |
| 45 | V$DELTAEF1_01 | ZEB1 group | 655 (+) | 1 | 0.987 | tggCACCTgaa |

| | | | | | | |
|---|---|---|---|---|---|---|
| 46 | V$ZFP206_01 | Zfp206 | 1081 (-) | 1 | 0.988 | gcgctTGCGCa |

**Table A.11. List of transcription factor-binding motifs predicted using the Consite database.** Only a single high-confidence site for each transcription factor is listed.

| Sl. No. | Transcription factor | Sequence | From | To | Score | Strand |
|---|---|---|---|---|---|---|
| 1 | AGL3 | CTATATATAA | 1177 | 1186 | 8.008 | - |
| 2 | AML-1 | AGCCGCAAT | 636 | 644 | 8.12 | - |
| 3 | ARNT | CACGCG | 956 | 961 | 7.103 | - |
| 4 | Athb-1 | CTATCATT | 195 | 202 | 8.654 | + |
| 5 | Broad-complex_1 | AATTTGTTTGATAT | 221 | 234 | 9.396 | - |
| 6 | Broad-complex_4 | AAGTAAAAAAT | 706 | 716 | 9.84 | + |
| 7 | bZIP911 | GAGGACGTGGAA | 907 | 918 | 10.5 | + |
| 8 | c-FOS | CTGATTCA | 498 | 505 | 8.378 | + |
| 9 | c-REL | GGAAAGCACG | 950 | 959 | 9.177 | - |
| 10 | CF2-II | CTATATATAA | 1177 | 1186 | 11.6 | + |
| 11 | CFI-USP | GGGGTAACGA | 670 | 679 | 9.784 | + |
| 12 | COUP-TF | TGAACTTTGGGGTA | 662 | 675 | 9.029 | + |
| 13 | CREB | CAAGTCACCCGC | 387 | 398 | 8.421 | - |
| 14 | Dorsal_1 | GGGGTGTCTTCC | 174 | 185 | 8.362 | + |
| 15 | Dorsal_2 | GGTGTCTTCC | 176 | 185 | 8.658 | + |
| 16 | E2F | TTTGGCCC | 513 | 520 | 8.832 | + |
| 17 | E4BP4 | AAGTTACGTAA | 523 | 533 | 13.66 | - |
| 18 | E74A | CAGGAAG | 560 | 566 | 9.644 | + |
| 19 | FREAC-4 | GTAAAAAA | 708 | 715 | 9.167 | + |
| 20 | Hen-1 | CAGCAGCGGAGG | 975 | 986 | 8.185 | - |
| 21 | HFH-1 | AACTGTTCATT | 99 | 109 | 8.706 | + |
| 22 | HFH-2 | AAAAAAAGATTA | 248 | 259 | 9.784 | - |
| 23 | HFH-3 | TTAAATTTGTTT | 218 | 229 | 9.008 | + |
| 24 | HLF | AGTTACGTAAAG | 524 | 535 | 9.539 | + |
| 25 | HMG-IY | GTGTCTTCCTTTCTGA | 177 | 192 | 8.019 | - |
| 26 | HNF-1 | GGTAACGATTAACT | 672 | 685 | 13.36 | - |
| 27 | HNF-3beta | AAGTAAAAAATT | 706 | 717 | 10.36 | - |
| 28 | Max | GAGGACGTGG | 907 | 916 | 6.75 | + |
| 29 | MEF2 | CTATATATAA | 1177 | 1186 | 12.3 | - |
| 30 | Myc-Max | GAGGACGTGGA | 907 | 917 | 9.902 | + |
| 31 | Myf | CGACAGCAGCGG | 972 | 983 | 10.48 | + |
| 32 | n-MYC | CACGCG | 956 | 961 | 7.79 | + |
| 33 | NF-kappaB | GTTAATTCCC | 333 | 342 | 9.145 | - |
| 34 | NRF-2 | GTCTTCCTTT | 179 | 188 | 9.24 | - |
| 35 | p65 | GGAAAGCACG | 950 | 959 | 7.822 | - |
| 36 | Pbx | TTTGTTTGATAT | 223 | 234 | 11.02 | - |
| 37 | Snail | CACCTG | 658 | 663 | 10.74 | - |
| 38 | Sox-5 | AAATAAT | 156 | 162 | 7.499 | + |
| 39 | SOX17 | GACACTGAA | 122 | 130 | 7.165 | - |
| 40 | Spz1 | AGGGTTGGAGG | 918 | 928 | 9.182 | + |
| 41 | SQUA | CTTCTTAAATTTGT | 214 | 227 | 7.345 | - |
| 42 | Staf | GCAGGGCACTAGGGGATGGG | 748 | 767 | 11.64 | - |

| | | | | | | |
|---|---|---|---|---|---|---|
| 43 | TBP | ATATATAAGCGCGGG | 1179 | 1193 | 13.06 | + |
| 44 | TEF-1 | AGGAGGAATGTT | 407 | 418 | 12.54 | - |
| 45 | Thing1-E47 | CGACTGGAAA | 945 | 954 | 7.556 | + |
| 46 | USF | GACGTGG | 910 | 916 | 6.973 | + |

**Table A.12. Expression analysis of NPM1 in OSCC tissues:** H-score data for NPM1 expression in oral tumor vs adjacent normal patient tissue samples.

| Sl. No. | Patient ID | H-score (adjacent normal) | H-score (tumor) |
|---|---|---|---|
| 1 | 66911 | 93.98 | 224.58 |
| 2 | 66951 | 63.51 | 176.26 |
| 3 | 66970 | 70.28 | 176.17 |
| 4 | 70531 | 62.64 | 233.60 |
| 5 | 72990 | 100.00 | 264.55 |
| 6 | 74597 | 38.72 | 131.58 |
| 7 | 74805 | 181.71 | 283.74 |
| 8 | 75556 | 97.25 | 153.76 |
| 9 | 79928 | 95.45 | 158.94 |
| 10 | 79962 | 152.96 | 189.18 |
| 11 | 80206 | 52.78 | 290.42 |
| 12 | 81056 | 154.45 | 279.22 |
| 13 | 81406 | 136.49 | 255.40 |
| 14 | 907559 | 110.87 | 101.59 |
| 15 | 908855 | 31.58 | 101.46 |
| 16 | 917986 | 80.71 | 100.00 |
| 17 | 920669 | 26.97 | 185.84 |
| 18 | 922107 | 100.00 | 123.72 |
| 19 | 56249 | 69.80 | 139.16 |
| 20 | 56271 | 120.55 | 159.87 |
| 21 | 56273 | 77.01 | 100.00 |
| 22 | 5841 | 115.98 | 156.44 |
| 23 | 58075 | 102.88 | 217.05 |
| 24 | 58457 | 100.00 | 139.58 |
| 25 | 81239 | 100.00 | 166.57 |
| 26 | 81673 | 100.00 | 110.51 |
| 27 | 81908 | 100.00 | 126.88 |
| 28 | 84429 | 115.67 | 138.41 |
| 29 | 85576 | 16.13 | 101.57 |
| 30 | 924073 | 100.00 | 150.62 |
| 31 | 924178 | 100.00 | 163.23 |
| 32 | 930713 | 76.92 | 110.94 |
| 33 | 934927 | 100.00 | 124.15 |
| 34 | 998486 | 100.00 | 144.42 |
| 35 | 1165 | 112.70 | 142.09 |
| 36 | 62265 | 23.28 | 153.12 |
| 37 | 77147 | 106.69 | 162.58 |
| 38 | 89014 | 122.96 | 141.30 |
| 39 | 1015497 | 3.04 | 208.98 |

| | | | |
|---|---|---|---|
| 40 | 955865 | 51.57 | 199.55 |
| 41 | 957783 | 100.00 | 258.58 |
| 42 | 961721 | 35.81 | 126.18 |
| 43 | 962579 | 123.70 | 262.88 |
| 44 | 1004604 | 82.97 | 115.35 |
| 45 | 1009225 | 100.00 | 116.67 |
| 46 | 1018358 | 101.74 | 151.50 |

**Table A.13. Expression analysis of c-fos in OSCC tissues:** H-score data for c-fos expression in oral tumor vs adjacent normal patient tissue samples.

| Sl. No. | Patient ID | H-score (adjacent normal) | H-score (tumor) |
|---|---|---|---|
| 1 | 66911 | 15.14 | 110.19 |
| 2 | 66951 | 15.06 | 62.35 |
| 3 | 70531 | 46.43 | 141.06 |
| 4 | 72990 | 11.86 | 182.55 |
| 5 | 74597 | 100.00 | 103.19 |
| 6 | 74805 | 0.00 | 117.24 |
| 7 | 75556 | 0.00 | 51.09 |
| 8 | 79928 | 66.21 | 109.23 |
| 9 | 80206 | 0.00 | 157.52 |
| 10 | 81056 | 0.00 | 143.00 |
| 11 | 907559 | 0.00 | 100.00 |
| 12 | 920669 | 0.00 | 161.94 |
| 13 | 922107 | 0.00 | 100.00 |
| 14 | 56249 | 10.09 | 100.00 |
| 15 | 56271 | 1.80 | 50.95 |
| 16 | 56273 | 3.02 | 111.34 |
| 17 | 5841 | 23.33 | 86.17 |
| 18 | 58075 | 0.00 | 92.42 |
| 19 | 58457 | 0.00 | 38.07 |
| 20 | 81239 | 0.00 | 48.77 |
| 21 | 84429 | 0.00 | 55.98 |
| 22 | 85576 | 0.00 | 45.05 |
| 23 | 924073 | 0.00 | 113.48 |
| 24 | 924178 | 9.19 | 116.49 |
| 25 | 930713 | 0.00 | 85.66 |
| 26 | 934927 | 0.00 | 116.72 |
| 27 | 998486 | 0.00 | 123.53 |
| 28 | 89014 | 5.54 | 152.97 |
| 29 | 1015497 | 0.00 | 159.46 |
| 30 | 955865 | 0.00 | 137.89 |
| 31 | 957783 | 22.64 | 133.23 |
| 32 | 961721 | 0.00 | 134.08 |
| 33 | 962579 | 0.00 | 120.44 |
| 34 | 1018358 | 0.00 | 116.71 |
| 35 | 74288 | 24.34 | 130.24 |

**Table A.14. Correlation analysis for the expression of YY1 and NPM1 transcript in various human cancers** (Table A.9)**:** Analyses have been carried out using transcriptome data publicly available in the TCGA (Methods, Section 2.13.3).

| Cancer type | No. of samples | Pearson's correlation value | *P* value |
|---|---|---|---|
| **Positive correlation** | | | |
| ACC | 79 | 0.4084 | 0.0002 |
| BLCA | 408 | 0.1059 | 0.0323 |
| BRCA | 1100 | 0.1859 | 5.16E-010 |
| COAD | 382 | 0.1666 | 0.001 |
| GBM | 166 | 0.2197 | 0.0044 |
| KIRC | 534 | 0.2564 | 1.84E-009 |
| KIRP | 291 | 0.4018 | 1.03E-012 |
| LAML | 173 | 0.3113 | 3.05E-005 |
| LUAD | 517 | 0.2235 | 2.84E-007 |
| PAAD | 179 | 0.3345 | 4.71E-006 |
| PRAD | 498 | 0.2766 | 3.41E-010 |
| SKCM | 472 | 0.2846 | 3.06E-010 |
| STAC | 415 | 0.104 | 0.0341 |
| TGCT | 156 | 0.3237 | 3.76E-005 |
| THCA | 509 | 0.1564 | 4.00E-004 |
| THYM | 120 | 0.4022 | 5.26E-006 |
| UCEC | 177 | 0.1539 | 0.0408 |
| UVM | 79 | 0.5323 | 4.44E-007 |
| **No correlation** | | | |
| CESC | 306 | 0.0207 | 0.7171 |
| CHOL | 36 | -0.2823 | 0.0953 |
| DLBC | 48 | -0.0718 | 0.6278 |
| ESO | 185 | 0.115 | 0.1191 |
| HNSC | 522 | -0.039 | 0.3742 |
| KICH | 66 | 0.1977 | 0.1116 |
| LGG | 530 | 0.0391 | 0.3695 |
| LIHC | 373 | -0.0644 | 0.2149 |
| LUSC | 501 | 0.0338 | 0.4506 |
| MESO | 87 | 0.1049 | 0.3336 |
| OV | 307 | -0.0205 | 0.7203 |
| PCPG | 184 | 0.0408 | 0.5828 |
| SARC | 263 | -0.136 | 0.8265 |
| UCS | 57 | 0.0844 | 0.5326 |

# *References*


2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**: 57-74.

Abe M, Lin J, Nagata K, Okuwaki M. 2018. Selective regulation of type II interferon-inducible genes by NPM1/nucleophosmin. *FEBS Lett* **592**: 244-255.

Abraham R, Singh S, Nair SR, Hulyalkar NV, Surendran A, Jaleel A, Sreekumar E. 2017. Nucleophosmin (NPM1)/B23 in the Proteome of Human Astrocytic Cells Restricts Chikungunya Virus Replication. *J Proteome Res* **16**: 4144-4155.

Acin S, Li Z, Mejia O, Roop DR, El-Naggar AK, Caulin C. 2011. Gain-of-function mutant p53 but not p53 deletion promotes head and neck cancer progression in response to oncogenic K-ras. *J Pathol* **225**: 479-489.

Adachi Y, Copeland TD, Hatanaka M, Oroszlan S. 1993. Nucleolar targeting signal of Rex protein of human T-cell leukemia virus type I specifically binds to nucleolar shuttle protein B-23. *J Biol Chem* **268**: 13930-13934.

Adkins MW, Carson JJ, English CM, Ramey CJ, Tyler JK. 2007. The histone chaperone anti-silencing function 1 stimulates the acetylation of newly synthesized histone H3 in S-phase. *J Biol Chem* **282**: 1334-1340.

Adkins MW, Howar SR, Tyler JK. 2004. Chromatin disassembly mediated by the histone chaperone Asf1 is essential for transcriptional activation of the yeast PHO5 and PHO8 genes. *Mol Cell* **14**: 657-666.

Aegerter S, Jalabert B, Bobe J. 2005. Large scale real-time PCR analysis of mRNA abundance in rainbow trout eggs in relationship with egg quality and post-ovulatory ageing. *Mol Reprod Dev* **72**: 377-385.

Ahmad K, Henikoff S. 2002. The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol Cell* **9**: 1191-1200.

Ahn JY, Liu X, Cheng D, Peng J, Chan PK, Wade PA, Ye K. 2005. Nucleophosmin/B23, a nuclear PI(3,4,5)P(3) receptor, mediates the antiapoptotic actions of NGF by inhibiting CAD. *Mol Cell* **18**: 435-445.

Akey CW, Luger K. 2003. Histone chaperones and nucleosome assembly. *Curr Opin Struct Biol* **13**: 6-14.

Akoulitchev S, Chuikov S, Reinberg D. 2000. TFIIH is negatively regulated by cdk8-containing mediator complexes. *Nature* **407**: 102-106.

Allan J, Hartman PG, Crane-Robinson C, Aviles FX. 1980. The structure of histone H1 and its location in chromatin. *Nature* **288**: 675-679.

Allen BL, Taatjes DJ. 2015. The Mediator complex: a central integrator of transcription. *Nat Rev Mol Cell Biol* **16**: 155-166.

Amemiya HM, Kundaje A, Boyle AP. 2019. The ENCODE Blacklist: Identification of Problematic Regions of the Genome. *Sci Rep* **9**: 9354.

Amor DJ, Bentley K, Ryan J, Perry J, Wong L, Slater H, Choo KH. 2004. Human centromere repositioning "in progress". *Proc Natl Acad Sci U S A* **101**: 6542-6547.

Ando K, Parsons MJ, Shah RB, Charendoff CI, Paris SL, Liu PH, Fassio SR, Rohrman BA, Thompson R, Oberst A et al. 2017. NPM1 directs PIDDosome-dependent caspase-2 activation in the nucleolus. *J Cell Biol* **216**: 1795-1810.

Angelov D, Bondarenko VA, Almagro S, Menoni H, Mongelard F, Hans F, Mietton F, Studitsky VM, Hamiche A, Dimitrov S et al. 2006. Nucleolin is a histone chaperone with FACT-like activity and assists remodeling of nucleosomes. *Embo j* **25**: 1669-1679.

Anish R, Hossain MB, Jacobson RH, Takada S. 2009. Characterization of transcription from TATA-less promoters: identification of a new core promoter element XCPE2 and analysis of factor requirements. *PLoS One* **4**: e5103.

Annunziato AT. 2013. Assembling chromatin: the long and winding road. *Biochim Biophys Acta* **1819**: 196-210.

Anselm E, Thomae AW, Jeyaprakash AA, Heun P. 2018. Oligomerization of Drosophila Nucleoplasmin-Like Protein is required for its centromere localization. *Nucleic Acids Res* **46**: 11274-11286.

Apta-Smith MJ, Hernandez-Fernaud JR, Bowman AJ. 2018. Evidence for the nuclear import of histones H3.1 and H4 as monomers. *Embo j* **37**.

Arents G, Moudrianakis EN. 1995. The histone fold: a ubiquitous architectural motif utilized in DNA compaction and protein dimerization. *Proc Natl Acad Sci U S A* **92**: 11170-11174.

Arif M, Vedamurthy BM, Choudhari R, Ostwal YB, Mantelingu K, Kodaganur GS, Kundu TK. 2010. Nitric oxide-mediated histone hyperacetylation in oral cancer: target for a water-soluble HAT inhibitor, CTK7A. *Chem Biol* **17**: 903-913.

Armache KJ, Mitterweger S, Meinhart A, Cramer P. 2005. Structures of complete RNA polymerase II and its subcomplex, Rpb4/7. *J Biol Chem* **280**: 7131-7134.

Arndt KM, Reines D. 2015. Termination of Transcription of Short Noncoding RNAs by RNA Polymerase II. *Annu Rev Biochem* **84**: 381-404.

Arnoldo L, Sgarra R, Chiefari E, Iiritano S, Arcidiacono B, Pegoraro S, Pellarin I, Brunetti A, Manfioletti G. 2015. A novel mechanism of post-translational modulation of HMGA functions by the histone chaperone nucleophosmin. *Sci Rep* **5**: 8552.

Arun G, Akhade VS, Donakonda S, Rao MR. 2012. mrhl RNA, a long noncoding RNA, negatively regulates Wnt signaling through its protein partner Ddx5/p68 in mouse spermatogonial cells. *Mol Cell Biol* **32**: 3140-3152.

Ashraf K, Nabeel-Shah S, Garg J, Saettone A, Derynck J, Gingras AC, Lambert JP, Pearlman RE, Fillingham J. 2019. Proteomic Analysis of Histones H2A/H2B and Variant Hv1 in Tetrahymena thermophila Reveals an Ancient Network of Chaperones. *Mol Biol Evol* **36**: 1037-1055.

Atchison ML. 2014. Function of YY1 in Long-Distance DNA Interactions. *Front Immunol* **5**: 45.

Auble DT, Hansen KE, Mueller CG, Lane WS, Thorner J, Hahn S. 1994. Mot1, a global repressor of RNA polymerase II transcription, inhibits TBP binding to DNA by an ATP-dependent mechanism. *Genes Dev* **8**: 1920-1934.

Avvakumov N, Nourani A, Cote J. 2011. Histone chaperones: modulators of chromatin marks. *Mol Cell* **41**: 502-514.

Bach M, Grigat S, Pawlik B, Fork C, Utermohlen O, Pal S, Banczyk D, Lazar A, Schomig E, Grundemann D. 2007. Fast set-up of doxycycline-inducible protein expression in human cell lines with a single plasmid based on Epstein-Barr virus replication and the simple tetracycline repressor. *Febs j* **274**: 783-790.

Bakiri L, Lallemand D, Bossy-Wetzel E, Yaniv M. 2000. Cell cycle-dependent variations in c-Jun and JunB phosphorylation: a role in the control of cyclin D1 expression. *Embo j* **19**: 2056-2068.

Ballow DJ, Xin Y, Choi Y, Pangas SA, Rajkovic A. 2006. Sohlh2 is a germ cell-specific bHLH transcription factor. *Gene Expr Patterns* **6**: 1014-1018.

Balusu R, Fiskus W, Rao R, Chong DG, Nalluri S, Mudunuru U, Ma H, Chen L, Venkannagari S, Ha K et al. 2011. Targeting levels or oligomerization of nucleophosmin 1 induces differentiation and loss of survival of human AML cells with mutant NPM1. *Blood* **118**: 3096-3106.

Bamberger AM, Milde-Langosch K, Rossing E, Goemann C, Loning T. 2001. Expression pattern of the AP-1 family in endometrial cancer: correlations with cell cycle regulators. *J Cancer Res Clin Oncol* **127**: 545-550.

Banaszynski LA, Allis CD, Lewis PW. 2010. Histone variants in metazoan development. *Dev Cell* **19**: 662-674.

Banaszynski LA, Wen D, Dewell S, Whitcomb SJ, Lin M, Diaz N, Elsasser SJ, Chapgier A, Goldberg AD, Canaani E et al. 2013. Hira-dependent histone H3.3 deposition facilitates PRC2 recruitment at developmental loci in ES cells. *Cell* **155**: 107-120.

Bannister AJ, Kouzarides T. 2011. Regulation of chromatin by histone modifications. *Cell Res* **21**: 381-395.

Bannister AJ, Schneider R, Myers FA, Thorne AW, Crane-Robinson C, Kouzarides T. 2005. Spatial distribution of di- and tri-methyl lysine 36 of histone H3 at active genes. *J Biol Chem* **280**: 17732-17736.

Banuelos S, Omaetxebarria MJ, Ramos I, Larsen MR, Arregi I, Jensen ON, Arizmendi JM, Prado A, Muga A. 2007. Phosphorylation of both nucleoplasmin domains is required for activation of its chromatin decondensation activity. *J Biol Chem* **282**: 21213-21221.

Bao Y, Konesky K, Park YJ, Rosu S, Dyer PN, Rangasamy D, Tremethick DJ, Laybourn PJ, Luger K. 2004. Nucleosomes containing the histone variant H2A.Bbd organize only 118 base pairs of DNA. *Embo j* **23**: 3314-3324.

Barnes CE, English DM, Cowley SM. 2019. Acetylation & Co: an expanding repertoire of histone acylations regulates chromatin and transcription. *Essays Biochem* **63**: 97-107.

Barnes CO, Calero M, Malik I, Graham BW, Spahr H, Lin G, Cohen AE, Brown IS, Zhang Q, Pullara F et al. 2015. Crystal Structure of a Transcribing RNA Polymerase II Complex Reveals a Complete Transcription Bubble. *Mol Cell* **59**: 258-269.

Barrero MJ, Malik S. 2013. The RNA polymerase II transcriptional machinery and its epigenetic context. *Subcell Biochem* **61**: 237-259.

Bartkowiak B, Liu P, Phatnani HP, Fuda NJ, Cooper JJ, Price DH, Adelman K, Lis JT, Greenleaf AL. 2010. CDK12 is a transcription elongation-associated CTD kinase, the metazoan ortholog of yeast Ctk1. *Genes Dev* **24**: 2303-2316.

Batta K, Kundu TK. 2007. Activation of p53 function by human transcriptional coactivator PC4: role of protein-protein interaction, DNA bending, and posttranslational modifications. *Mol Cell Biol* **27**: 7603-7614.

Beaudry M, Hidalgo M, Launay T, Bello V, Darribere T. 2016. Regulation of myogenesis by environmental hypoxia. *J Cell Sci* **129**: 2887-2896.

Becker PB, Workman JL. 2013. Nucleosome remodeling and epigenetics. *Cold Spring Harb Perspect Biol* **5**.

Bednar J, Garcia-Saez I, Boopathi R, Cutter AR, Papai G, Reymer A, Syed SH, Lone IN, Tonchev O, Crucifix C et al. 2017. Structure and Dynamics of a 197 bp Nucleosome in Complex with Linker Histone H1. *Mol Cell* **66**: 384-397.e388.

Behera AK, Kumar M, Shanmugam MK, Bhattacharya A, Rao VJ, Bhat A, Vasudevan M, Gopinath KS, Mohiyuddin A, Chatterjee A et al. 2019. Functional interplay between YY1 and CARM1 promotes oral carcinogenesis. *Oncotarget* **10**: 3709-3724.

Belotserkovskaya R, Oh S, Bondarenko VA, Orphanides G, Studitsky VM, Reinberg D. 2003. FACT facilitates transcription-dependent nucleosome alteration. *Science* **301**: 1090-1093.

Beneke S. 2012. Regulation of chromatin structure by poly(ADP-ribosyl)ation. *Front Genet* **3**: 169.

Bergstralh DT, Conti BJ, Moore CB, Brickey WJ, Taxman DJ, Ting JP. 2007. Global functional analysis of nucleophosmin in Taxol response, cancer, chromatin regulation, and ribosomal DNA transcription. *Exp Cell Res* **313**: 65-76.

Bernecky C, Herzog F, Baumeister W, Plitzko JM, Cramer P. 2016. Structure of transcribing mammalian RNA polymerase II. *Nature* **529**: 551-554.

Bernstein BE, Humphrey EL, Erlich RL, Schneider R, Bouman P, Liu JS, Kouzarides T, Schreiber SL. 2002. Methylation of histone H3 Lys 4 in coding regions of active genes. *Proc Natl Acad Sci U S A* **99**: 8695-8700.

Bertwistle D, Sugimoto M, Sherr CJ. 2004. Physical and functional interactions of the Arf tumor suppressor protein with nucleophosmin/B23. *Mol Cell Biol* **24**: 985-996.

Bevington JM, Needham PG, Verrill KC, Collaco RF, Basrur V, Trempe JP. 2007. Adeno-associated virus interactions with B23/Nucleophosmin: identification of sub-nucleolar virion regions. *Virology* **357**: 102-113.

Bi X, Xu Y, Li T, Li X, Li W, Shao W, Wang K, Zhan G, Wu Z, Liu W et al. 2019. RNA Targets Ribogenesis Factor WDR43 to Chromatin for Transcription and Pluripotency Control. *Mol Cell* **75**: 102-116.e109.

Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, Delano D, Zhang L, Schroth GP, Gunderson KL et al. 2011. High density DNA methylation array with single CpG site resolution. *Genomics* **98**: 288-295.

Billon P, Cote J. 2013. Precise deposition of histone H2A.Z in chromatin for genome expression and maintenance. *Biochim Biophys Acta* **1819**: 290-302.

Birse CE, Minvielle-Sebastia L, Lee BA, Keller W, Proudfoot NJ. 1998. Coupling termination of transcription to messenger RNA maturation in yeast. *Science* **280**: 298-301.

Bischof D, Pulford K, Mason DY, Morris SW. 1997. Role of the nucleophosmin (NPM) portion of the non-Hodgkin's lymphoma-associated NPM-anaplastic lymphoma kinase fusion protein in oncogenesis. *Mol Cell Biol* **17**: 2312-2325.

Bland KI, Konstadoulakis MM, Vezeridis MP, Wanebo HJ. 1995. Oncogene protein co-expression. Value of Ha-ras, c-myc, c-fos, and p53 as prognostic discriminants for breast carcinoma. *Ann Surg* **221**: 706-718; discussion 718-720.

Blazek E, Mittler G, Meisterernst M. 2005. The mediator of RNA polymerase II. *Chromosoma* **113**: 399-408.

Blom N, Gammeltoft S, Brunak S. 1999. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J Mol Biol* **294**: 1351-1362.

Boehning M, Dugast-Darzacq C, Rankovic M, Hansen AS, Yu T, Marie-Nelly H, McSwiggen DT, Kokic G, Dailey GM, Cramer P et al. 2018. RNA polymerase II clustering through carboxy-terminal domain phase separation. *Nat Struct Mol Biol* **25**: 833-840.

Bonetti P, Davoli T, Sironi C, Amati B, Pelicci PG, Colombo E. 2008. Nucleophosmin and its AML-associated mutant regulate c-Myc turnover through Fbw7 gamma. *J Cell Biol* **182**: 19-26.

Bonnefoy E, Orsi GA, Couble P, Loppin B. 2007. The essential role of Drosophila HIRA for de novo assembly of paternal chromatin at fertilization. *PLoS Genet* **3**: 1991-2006.

Borer RA, Lehner CF, Eppenberger HM, Nigg EA. 1989. Major nucleolar proteins shuttle between nucleus and cytoplasm. *Cell* **56**: 379-390.

Boudra R, Lagrafeuille R, Lours-Calet C, de Joussineau C, Loubeau-Legros G, Chaveroux C, Saru JP, Baron S, Morel L, Beaudoin C. 2016. mTOR transcriptionally and post-transcriptionally regulates Npm1 gene expression to contribute to enhanced proliferation in cells with Pten inactivation. *Cell Cycle* **15**: 1352-1362.

Bouleau A, Desvignes T, Traverso JM, Nguyen T, Chesnel F, Fauvel C, Bobe J. 2014. Maternally inherited npm2 mRNA is crucial for egg developmental competence in zebrafish. *Biol Reprod* **91**: 43.

Bowman A, Ward R, Wiechens N, Singh V, El-Mkami H, Norman DG, Owen-Hughes T. 2011. The histone chaperones Nap1 and Vps75 bind histones H3 and H4 in a tetrameric conformation. *Mol Cell* **41**: 398-408.

Bowman EA, Kelly WG. 2014. RNA polymerase II transcription elongation and Pol II CTD Ser2 phosphorylation: A tail of two kinases. *Nucleus* **5**: 224-236.

Box JK, Paquet N, Adams MN, Boucher D, Bolderson E, O'Byrne KJ, Richard DJ. 2016. Nucleophosmin: from structure and function to disease development. *BMC Mol Biol* **17**: 19.

Boyer TG, Martin ME, Lees E, Ricciardi RP, Berk AJ. 1999. Mammalian Srb/Mediator complex is targeted by adenovirus E1A protein. *Nature* **399**: 276-279.

Braccioli L, de Wit E. 2019. CTCF: a Swiss-army knife for genome organization and transcription regulation. *Essays Biochem* **63**: 157-165.

Brady SN, Yu Y, Maggi LB, Jr., Weber JD. 2004. ARF impedes NPM/B23 shuttling in an Mdm2-sensitive tumor suppressor pathway. *Mol Cell Biol* **24**: 9327-9338.

Braun T, Gautel M. 2011. Transcriptional mechanisms regulating skeletal muscle differentiation, growth and homeostasis. *Nat Rev Mol Cell Biol* **12**: 349-361.

Brivanlou AH, Darnell JE, Jr. 2002. Signal transduction and the control of gene expression. *Science* **295**: 813-818.

Bruggeman JW, Koster J, Lodder P, Repping S, Hamer G. 2018. Massive expression of germ cell-specific genes is a hallmark of cancer and a potential target for novel treatment development. *Oncogene* **37**: 5694-5700.

Buratowski S. 1994. The basics of basal transcription by RNA polymerase II. *Cell* **77**: 1-3.

Buratowski S. 2005. Connections between mRNA 3' end processing and transcription termination. *Curr Opin Cell Biol* **17**: 257-261.

Buratowski S, Hahn S, Guarente L, Sharp PA. 1989. Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell* **56**: 549-561.

Burgess RJ, Zhang Z. 2013. Histone chaperones in nucleosome assembly and human disease. *Nat Struct Mol Biol* **20**: 14-22.

Burglin TR, Mattaj IW, Newmeyer DD, Zeller R, De Robertis EM. 1987. Cloning of nucleoplasmin from Xenopus laevis oocytes and analysis of its developmental expression. *Genes Dev* **1**: 97-107.

Burke TW, Kadonaga JT. 1997. The downstream core promoter element, DPE, is conserved from Drosophila to humans and is recognized by TAFII60 of Drosophila. *Genes Dev* **11**: 3020-3031.

Burley SK, Roeder RG. 1996. Biochemistry and structural biology of transcription factor IID (TFIID). *Annu Rev Biochem* **65**: 769-799.

Burns KH, Viveiros MM, Ren Y, Wang P, DeMayo FJ, Frail DE, Eppig JJ, Matzuk MM. 2003. Roles of NPM2 in chromatin and nucleolar organization in oocytes and embryos. *Science* **300**: 633-636.

Bushnell DA, Westover KD, Davis RE, Kornberg RD. 2004. Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 Angstroms. *Science* **303**: 983-988.

Bönisch C, Hake SB. 2012. Histone H2A variants in nucleosomes and chromatin: more or less stable? *Nucleic Acids Res* **40**: 10719-10741.

Cai XZ, Zeng WQ, Xiang Y, Liu Y, Zhang HM, Li H, She S, Yang M, Xia K, Peng SF. 2015. iTRAQ-Based Quantitative Proteomic Analysis of Nasopharyngeal Carcinoma. *J Cell Biochem* **116**: 1431-1441.

Cakiroglu A, Clapier CR, Ehrensberger AH, Darbo E, Cairns BR, Luscombe NM, Svejstrup JQ. 2019. Genome-wide reconstitution of chromatin transactions reveals that RSC preferentially disrupts H2AZ-containing nucleosomes. *Genome Res* **29**: 988-998.

Cammas A, Sanchez BJ, Lian XJ, Dormoy-Raclet V, van der Giessen K, Lopez de Silanes I, Ma J, Wilusz C, Richardson J, Gorospe M et al. 2014. Destabilization of nucleophosmin mRNA by the HuR/KSRP complex is required for muscle fibre formation. *Nat Commun* **5**: 4190.

Campos EI, Fillingham J, Li G, Zheng H, Voigt P, Kuo WH, Seepany H, Gao Z, Day LA, Greenblatt JF et al. 2010. The program for processing newly synthesized histones H3.1 and H4. *Nat Struct Mol Biol* **17**: 1343-1351.

Catalanotto C, Cogoni C, Zardo G. 2016. MicroRNA in Control of Gene Expression: An Overview of Nuclear Functions. *Int J Mol Sci* **17**.

Cecconi D, Carbonare LD, Mori A, Cheri S, Deiana M, Brandi J, Degaetano V, Masiero V, Innamorati G, Mottes M et al. 2018. An integrated approach identifies new oncotargets in melanoma. *Oncotarget* **9**: 11489-11502.

Chadwick BP, Willard HF. 2001. A novel chromatin protein, distantly related to histone H2A, is largely excluded from the inactive X chromosome. *J Cell Biol* **152**: 375-384.

Chan HJ, Weng JJ, Yung BY. 2005. Nucleophosmin/B23-binding peptide inhibits tumor growth and up-regulates transcriptional activity of p53. *Biochem Biophys Res Commun* **333**: 396-403.

Chan PK. 1992. Characterization and cellular localization of nucleophosmin/B23 in HeLa cells treated with selected cytotoxic agents (studies of B23-translocation mechanism). *Exp Cell Res* **203**: 174-181.

Chan PK, Chan FY, Morris SW, Xie Z. 1997. Isolation and characterization of the human nucleophosmin/B23 (NPM) gene: identification of the YY1 binding site at the 5' enhancer region. *Nucleic Acids Res* **25**: 1225-1232.

Chang JH, Lin JY, Wu MH, Yung BY. 1998. Evidence for the ability of nucleophosmin/B23 to bind ATP. *Biochem J* **329 ( Pt 3)**: 539-544.

Chang JH, Olson MO. 1989. A single gene codes for two forms of rat nucleolar protein B23 mRNA. *J Biol Chem* **264**: 11732-11737.

Chang TP, Yu SL, Lin SY, Hsiao YJ, Chang GC, Yang PC, Chen JJ. 2010. Tumor suppressor HLJ1 binds and functionally alters nucleophosmin via activating enhancer binding protein 2alpha complex formation. *Cancer Res* **70**: 1656-1667.

Chang WH, Kornberg RD. 2000. Electron crystal structure of the transcription factor and DNA repair complex, core TFIIH. *Cell* **102**: 609-613.

Chapman RD, Heidemann M, Hintermair C, Eick D. 2008. Molecular evolution of the RNA polymerase II CTD. *Trends Genet* **24**: 289-296.

Chen H, Li B, Workman JL. 1994. A histone-binding protein, nucleoplasmin, stimulates transcription factor binding to nucleosomes and factor-induced nucleosome disassembly. *Embo j* **13**: 380-390.

Chen P, Tomschik M, Nelson KM, Oakey J, Gatlin JC, Levy DL. 2019. Nucleoplasmin is a limiting component in the scaling of nuclear size with cytoplasmic volume. *J Cell Biol*.

Chen R, Wang Y, Liu Y, Zhang Q, Zhang X, Zhang F, Shieh CH, Yang D, Zhang N. 2013. Quantitative study of the interactome of PKCzeta involved in the EGF-induced tumor cell chemotaxis. *J Proteome Res* **12**: 1478-1486.

Chen ZA, Jawhari A, Fischer L, Buchen C, Tahir S, Kamenski T, Rasmussen M, Lariviere L, Bukowski-Wills JC, Nilges M et al. 2010. Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. *Embo j* **29**: 717-726.

Cheung AC, Cramer P. 2012. A movie of RNA polymerase II transcription. *Cell* **149**: 1431-1437.

Cheung CT, Pasquier J, Bouleau A, Nguyen T, Chesnel F, Guiguen Y, Bobe J. 2018. Double maternal-effect: duplicated nucleoplasmin 2 genes, npm2a and npm2b, with essential but distinct functions are shared by fish and tetrapods. *BMC Evol Biol* **18**: 167.

Chicas A, Molina P, Bargonetti J. 2000. Mutant p53 forms a complex with Sp1 on HIV-LTR DNA. *Biochem Biophys Res Commun* **279**: 383-390.

Chiu R, Boyle WJ, Meek J, Smeal T, Hunter T, Karin M. 1988. The c-Fos protein interacts with c-Jun/AP-1 to stimulate transcription of AP-1 responsive genes. *Cell* **54**: 541-552.

Cho WK, Spille JH, Hecht M, Lee C, Li C, Grube V, Cisse, II. 2018. Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* **361**: 412-415.

Choi JW, Lee SB, Kim CK, Lee KH, Cho SW, Ahn JY. 2008. Lysine 263 residue of NPM/B23 is essential for regulating ATP binding and B23 stability. *FEBS Lett* **582**: 1073-1080.

Choy JS, Wei S, Lee JY, Tan S, Chu S, Lee TH. 2010. DNA methylation increases nucleosome compaction and rigidity. *J Am Chem Soc* **132**: 1782-1783.

Chujo M, Tarumoto Y, Miyatake K, Nishida E, Ishikawa F. 2012. HIRA, a conserved histone chaperone, plays an essential role in low-dose stress response via transcriptional stimulation in fission yeast. *J Biol Chem* **287**: 23440-23450.

Clausell J, Happel N, Hale TK, Doenecke D, Beato M. 2009. Histone H1 subtypes differentially modulate chromatin condensation without preventing ATP-dependent remodeling by SWI/SNF or NURF. *PLoS One* **4**: e0007243.

Colombo E, Alcalay M, Pelicci PG. 2011. Nucleophosmin and its complex network: a possible therapeutic target in hematological diseases. *Oncogene* **30**: 2595-2609.

Colombo E, Bonetti P, Lazzerini Denchi E, Martinelli P, Zamponi R, Marine JC, Helin K, Falini B, Pelicci PG. 2005. Nucleophosmin is required for DNA integrity and p19Arf protein stability. *Mol Cell Biol* **25**: 8874-8886.

Colombo E, Marine JC, Danovi D, Falini B, Pelicci PG. 2002. Nucleophosmin regulates the stability and transcriptional activity of p53. *Nat Cell Biol* **4**: 529-533.

Colombo E, Martinelli P, Zamponi R, Shing DC, Bonetti P, Luzi L, Volorio S, Bernard L, Pruneri G, Alcalay M et al. 2006. Delocalization and destabilization of the Arf tumor suppressor by the leukemia-associated NPM mutant. *Cancer Res* **66**: 3044-3050.

Comai G, Tajbakhsh S. 2014. Molecular and cellular regulation of skeletal myogenesis. *Curr Top Dev Biol* **110**: 1-73.

Cong R, Das S, Douet J, Wong J, Buschbeck M, Mongelard F, Bouvet P. 2014. macroH2A1 histone variant represses rDNA transcription. *Nucleic Acids Res* **42**: 181-192.

Cong R, Das S, Ugrinova I, Kumar S, Mongelard F, Wong J, Bouvet P. 2012. Interaction of nucleolin with ribosomal RNA genes and its role in RNA polymerase I transcription. *Nucleic Acids Res* **40**: 9441-9454.

Core L, Adelman K. 2019. Promoter-proximal pausing of RNA polymerase II: a nexus of gene regulation. *Genes Dev* **33**: 960-982.

Corpet A, De Koning L, Toedling J, Savignoni A, Berger F, Lemaitre C, O'Sullivan RJ, Karlseder J, Barillot E, Asselain B et al. 2011. Asf1b, the necessary Asf1 isoform for proliferation, is predictive of outcome in breast cancer. *Embo j* **30**: 480-493.

Costanzi C, Pehrson JR. 1998. Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. *Nature* **393**: 599-601.

Cotten M, Sealy L, Chalkley R. 1986. Massive phosphorylation distinguishes Xenopus laevis nucleoplasmin isolated from oocytes or unfertilized eggs. *Biochemistry* **25**: 5063-5069.

Coutinho-Camillo CM, Lourenco SV, Nishimoto IN, Kowalski LP, Soares FA. 2010. Nucleophosmin, p53, and Ki-67 expression patterns on an oral squamous cell carcinoma tissue microarray. *Hum Pathol* **41**: 1079-1086.

Cramer P. 2002. Multisubunit RNA polymerases. *Curr Opin Struct Biol* **12**: 89-97.

Cramer P. 2019. Organization and regulation of gene transcription. *Nature* **573**: 45-54.

Cramer P, Armache KJ, Baumli S, Benkert S, Brueckner F, Buchen C, Damsma GE, Dengl S, Geiger SR, Jasiak AJ et al. 2008. Structure of eukaryotic RNA polymerases. *Annu Rev Biophys* **37**: 337-352.

Cramer P, Bushnell DA, Fu J, Gnatt AL, Maier-Davis B, Thompson NE, Burgess RR, Edwards AM, David PR, Kornberg RD. 2000. Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* **288**: 640-649.

Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA et al. 2010. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci U S A* **107**: 21931-21936.

Croston GE, Laybourn PJ, Paranjape SM, Kadonaga JT. 1992. Mechanism of transcriptional antirepression by GAL4-VP16. *Genes Dev* **6**: 2270-2281.

Cuomo ME, Knebel A, Morrice N, Paterson H, Cohen P, Mittnacht S. 2008. p53-Driven apoptosis limits centrosome amplification and genomic instability downstream of NPM1 phosphorylation. *Nat Cell Biol* **10**: 723-730.

Curran T, Van Beveren C, Verma IM. 1985. Viral and cellular fos proteins are complexed with a 39,000-dalton cellular protein. *Mol Cell Biol* **5**: 167-172.

Czudnochowski N, Bosken CA, Geyer M. 2012. Serine-7 but not serine-5 phosphorylation primes RNA polymerase II CTD for P-TEFb recognition. *Nat Commun* **3**: 842.

Dai L, Li J, Ortega R, Qian W, Casiano CA, Zhang JY. 2014. Preferential autoimmune response in prostate cancer to cyclin B1 in a panel of tumor-associated antigens. *J Immunol Res* **2014**: 827827.

Dai L, Li J, Xing M, Sanchez TW, Casiano CA, Zhang JY. 2016a. Using Serological Proteome Analysis to Identify Serum Anti-Nucleophosmin 1 Autoantibody as a Potential Biomarker in European-American and African-American Patients With Prostate Cancer. *Prostate* **76**: 1375-1386.

Dai L, Tsay JC, Li J, Yie TA, Munger JS, Pass H, Rom WN, Zhang Y, Tan EM, Zhang JY. 2016b. Autoantibodies against tumor-associated antigens in the early detection of lung cancer. *Lung Cancer* **99**: 172-179.

Dai X, Bai Y, Zhao L, Dou X, Liu Y, Wang L, Li Y, Li W, Hui Y, Huang X et al. 2017. H2A.Z Represses Gene Expression by Modulating Promoter Nucleosome Structure and Enhancer Histone Modifications in Arabidopsis. *Mol Plant* **10**: 1274-1292.

Dalenc F, Drouet J, Ader I, Delmas C, Rochaix P, Favre G, Cohen-Jonathan E, Toulas C. 2002. Increased expression of a COOH-truncated nucleophosmin resulting from alternative splicing is associated with cellular resistance to ionizing radiation in HeLa cells. *Int J Cancer* **100**: 662-668.

Darracq A, Pak H, Bourgoin V, Zmiri F, Dellaire G, Affar EB, Milot E. 2019. NPM and NPM-MLF1 interact with chromatin remodeling complexes and influence their recruitment to specific genes. *PLoS Genet* **15**: e1008463.

Das C, Gadad SS, Kundu TK. 2010a. Human positive coactivator 4 controls heterochromatinization and silencing of neural gene expression by interacting with REST/NRSF and CoREST. *J Mol Biol* **397**: 1-12.

Das C, Hizume K, Batta K, Kumar BR, Gadad SS, Ganguly S, Lorain S, Verreault A, Sadhale PP, Takeyasu K et al. 2006. Transcriptional coactivator PC4, a chromatin-associated protein, induces chromatin condensation. *Mol Cell Biol* **26**: 8303-8315.

Das C, Tyler JK, Churchill ME. 2010b. The histone shuffle: histone chaperones in an energetic dance. *Trends Biochem Sci* **35**: 476-489.

Das S, Cong R, Shandilya J, Senapati P, Moindrot B, Monier K, Delage H, Mongelard F, Kumar S, Kundu TK et al. 2013. Characterization of nucleolin K88 acetylation defines a new pool of nucleolin colocalizing with pre-mRNA splicing factors. *FEBS Lett* **587**: 417-424.

Dawson MA, Gudgin EJ, Horton SJ, Giotopoulos G, Meduri E, Robson S, Cannizzaro E, Osaki H, Wiese M, Putwain S et al. 2014. Recurrent mutations, including NPM1c, activate a BRD4-dependent core transcriptional program in acute myeloid leukemia. *Leukemia* **28**: 311-320.

Day PM, Thompson CD, Pang YY, Lowy DR, Schiller JT. 2015. Involvement of Nucleophosmin (NPM1/B23) in Assembly of Infectious HPV16 Capsids. *Papillomavirus Res* **1**: 74-89.

De Koning L, Corpet A, Haber JE, Almouzni G. 2007. Histone chaperones: an escort network regulating histone traffic. *Nat Struct Mol Biol* **14**: 997-1007.

De La Fuente R, Viveiros MM, Burns KH, Adashi EY, Matzuk MM, Eppig JJ. 2004. Major chromatin remodeling in the germinal vesicle (GV) of mammalian oocytes is dispensable for global transcriptional silencing but required for centromeric heterochromatin function. *Dev Biol* **275**: 447-458.

Decock A, Ongenaert M, Cannoodt R, Verniers K, De Wilde B, Laureys G, Van Roy N, Berbegall AP, Bienertova-Vasku J, Bown N et al. 2016. Methyl-CpG-binding domain sequencing reveals a prognostic methylation signature in neuroblastoma. *Oncotarget* **7**: 1960-1972.

Dekker J, Heard E. 2015. Structural and functional diversity of Topologically Associating Domains. *FEBS Lett* **589**: 2877-2884.

Del Rosario BC, Pemberton LF. 2008. Nap1 links transcription elongation, chromatin assembly, and messenger RNP complex biogenesis. *Mol Cell Biol* **28**: 2113-2124.

den Besten W, Kuo ML, Williams RT, Sherr CJ. 2005. Myeloid leukemia-associated nucleophosmin mutants perturb p53-dependent and independent activities of the Arf tumor suppressor protein. *Cell Cycle* **4**: 1593-1598.

Destouches D, Sader M, Terry S, Marchand C, Maille P, Soyeux P, Carpentier G, Semprez F, Ceraline J, Allory Y et al. 2016. Implication of NPM1 phosphorylation and preclinical evaluation of the nucleoprotein antagonist N6L in prostate cancer. *Oncotarget* **7**: 69397-69411.

Dhanasekaran K, Bose A, Rao VJ, Boopathi R, Shankar SR, Rao VK, Swaminathan A, Vasudevan M, Taneja R, Kundu TK. 2019. Unraveling the role of aurora A beyond centrosomes and spindle assembly: implications in muscle differentiation. *Faseb j* **33**: 219-230.

Dhar SK, Lynn BC, Daosukho C, St Clair DK. 2004. Identification of nucleophosmin as an NF-kappaB co-activator for the induction of the human SOD2 gene. *J Biol Chem* **279**: 28209-28219.

Di Agostino S, Strano S, Emiliozzi V, Zerbini V, Mottolese M, Sacchi A, Blandino G, Piaggio G. 2006. Gain of function of mutant p53: the mutant p53/NF-Y protein complex reveals an aberrant transcriptional mechanism of cell cycle regulation. *Cancer Cell* **10**: 191-202.

Di Matteo A, Franceschini M, Chiarella S, Rocchio S, Travaglini-Allocatelli C, Federici L. 2016. Molecules that target nucleophosmin for cancer treatment: an update. *Oncotarget* **7**: 44821-44840.

Ding N, Zhou H, Esteve PO, Chin HG, Kim S, Xu X, Joseph SM, Friez MJ, Schwartz CE, Pradhan S et al. 2008. Mediator links epigenetic silencing of neuronal gene expression with x-linked mental retardation. *Mol Cell* **31**: 347-359.

Dion MF, Altschuler SJ, Wu LF, Rando OJ. 2005. Genomic characterization reveals a simple histone H4 acetylation code. *Proc Natl Acad Sci U S A* **102**: 5501-5506.

Douet J, Corujo D, Malinverni R, Renauld J, Sansoni V, Posavec Marjanovic M, Cantarino N, Valero V, Mongelard F, Bouvet P et al. 2017. MacroH2A histone variants maintain nuclear organization and heterochromatin architecture. *J Cell Sci* **130**: 1570-1582.

Doyen CM, Montel F, Gautier T, Menoni H, Claudet C, Delacour-Larose M, Angelov D, Hamiche A, Bednar J, Faivre-Moskalenko C et al. 2006. Dissection of the unusual structural and functional properties of the variant H2A.Bbd nucleosome. *Embo j* **25**: 4234-4244.

Drabent B, Kardalinou E, Doenecke D. 1991. Structure and expression of the human gene encoding testicular H1 histone (H1t). *Gene* **103**: 263-268.

Draker R, Cheung P. 2009. Transcriptional and epigenetic functions of histone variant H2A.Z. *Biochem Cell Biol* **87**: 19-25.

Draker R, Ng MK, Sarcinella E, Ignatchenko V, Kislinger T, Cheung P. 2012. A combination of H2A.Z and H4 acetylation recruits Brd2 to chromatin during transcriptional activation. *PLoS Genet* **8**: e1003047.

Drygin D, Rice WG, Grummt I. 2010. The RNA polymerase I transcription machinery: an emerging target for the treatment of cancer. *Annu Rev Pharmacol Toxicol* **50**: 131-156.

Dumbar TS, Gentry GA, Olson MO. 1989. Interaction of nucleolar phosphoprotein B23 with nucleic acids. *Biochemistry* **28**: 9495-9501.

Dumbovic G, Biayna J, Banus J, Samuelsson J, Roth A, Diederichs S, Alonso S, Buschbeck M, Perucho M, Forcales SV. 2018. A novel long non-coding RNA from NBL2 pericentromeric macrosatellite forms a perinucleolar aggregate structure in colon cancer. *Nucleic Acids Res* **46**: 5504-5524.

Dundr M, Olson MO. 1998. Partially processed pre-rRNA is preserved in association with processing components in nucleolus-derived foci during mitosis. *Mol Biol Cell* **9**: 2407-2422.

Dutta S, Akey IV, Dingwall C, Hartman KL, Laue T, Nolte RT, Head JF, Akey CW. 2001. The crystal structure of nucleoplasmin-core: implications for histone binding and nucleosome assembly. *Mol Cell* **8**: 841-853.

Dvir A. 2002. Promoter escape by RNA polymerase II. *Biochim Biophys Acta* **1577**: 208-223.

Dvir A, Conaway JW, Conaway RC. 2001. Mechanism of transcription initiation and promoter escape by RNA polymerase II. *Curr Opin Genet Dev* **11**: 209-214.

Dvir A, Tan S, Conaway JW, Conaway RC. 1997. Promoter escape by RNA polymerase II. Formation of an escape-competent transcriptional intermediate is a prerequisite for exit of polymerase from the promoter. *J Biol Chem* **272**: 28175-28178.

Earnshaw WC, Honda BM, Laskey RA, Thomas JO. 1980. Assembly of nucleosomes: the reaction involving X. laevis nucleoplasmin. *Cell* **21**: 373-383.

Ebright RH. 2000. RNA polymerase: structural similarities between bacterial RNA polymerase and eukaryotic RNA polymerase II. *J Mol Biol* **304**: 687-698.

Edlich-Muth C, Artero JB, Callow P, Przewloka MR, Watson AA, Zhang W, Glover DM, Debski J, Dadlez M, Round AR et al. 2015. The pentameric nucleoplasmin fold is present in Drosophila FKBP39 and a large number of chromatin-related proteins. *J Mol Biol* **427**: 1949-1963.

Eferl R, Wagner EF. 2003. AP-1: a double-edged sword in tumorigenesis. *Nat Rev Cancer* **3**: 859-868.

Egloff S, Murphy S. 2008. Cracking the RNA polymerase II CTD code. *Trends Genet* **24**: 280-288.

Eirin-Lopez JM, Frehlick LJ, Ausio J. 2006. Long-term evolution and functional diversification in the members of the nucleophosmin/nucleoplasmin family of nuclear chaperones. *Genetics* **173**: 1835-1850.

Eissenberg JC, Elgin SC. 2014. HP1a: a structural chromosomal protein regulating transcription. *Trends Genet* **30**: 103-110.

Eitoku M, Sato L, Senda T, Horikoshi M. 2008. Histone chaperones: 30 years from isolation to elucidation of the mechanisms of nucleosome assembly and disassembly. *Cell Mol Life Sci* **65**: 414-444.

Elkeles A, Juven-Gershon T, Israeli D, Wilder S, Zalcenstein A, Oren M. 1999. The c-fos proto-oncogene is a target for transactivation by the p53 tumor suppressor. *Mol Cell Biol* **19**: 2594-2600.

Elsasser SJ, D'Arcy S. 2013. Towards a mechanism for histone chaperones. *Biochim Biophys Acta* **1819**: 211-221.

Emelyanov AV, Rabbani J, Mehta M, Vershilova E, Keogh MC, Fyodorov DV. 2014. Drosophila TAP/p32 is a core histone chaperone that cooperates with NAP-1, NLP, and nucleophosmin in sperm chromatin remodeling during fertilization. *Genes Dev* **28**: 2027-2040.

Endo A, Matsumoto M, Inada T, Yamamoto A, Nakayama KI, Kitamura N, Komada M. 2009. Nucleolar structure and function are regulated by the deubiquitylating enzyme USP36. *J Cell Sci* **122**: 678-686.

Enomoto T, Lindstrom MS, Jin A, Ke H, Zhang Y. 2006. Essential role of the B23/NPM core domain in regulating ARF binding and B23 stability. *J Biol Chem* **281**: 18463-18472.

Fabrega C, Hausmann S, Shen V, Shuman S, Lima CD. 2004. Structure and mechanism of mRNA cap (guanine-N7) methyltransferase. *Mol Cell* **13**: 77-89.

Fabrega C, Shen V, Shuman S, Lima CD. 2003. Structure of an mRNA capping enzyme bound to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Mol Cell* **11**: 1549-1561.

Falini B, Bolli N, Liso A, Martelli MP, Mannucci R, Pileri S, Nicoletti I. 2009. Altered nucleophosmin transport in acute myeloid leukaemia with mutated NPM1: molecular basis and clinical implications. *Leukemia* **23**: 1731-1743.

Falini B, Bolli N, Shan J, Martelli MP, Liso A, Pucciarini A, Bigerna B, Pasqualucci L, Mannucci R, Rosati R et al. 2006a. Both carboxy-terminus NES motif and mutated tryptophan(s) are crucial for aberrant nuclear export of nucleophosmin leukemic mutants in NPMc+ AML. *Blood* **107**: 4514-4523.

Falini B, Martelli MP, Bolli N, Bonasso R, Ghia E, Pallotta MT, Diverio D, Nicoletti I, Pacini R, Tabarrini A et al. 2006b. Immunohistochemistry predicts nucleophosmin (NPM) mutations in acute myeloid leukemia. *Blood* **108**: 1999-2005.

Falini B, Mecucci C, Tiacci E, Alcalay M, Rosati R, Pasqualucci L, La Starza R, Diverio D, Colombo E, Santucci A et al. 2005. Cytoplasmic nucleophosmin in acute myelogenous leukemia with a normal karyotype. *N Engl J Med* **352**: 254-266.

Falini B, Nicoletti I, Bolli N, Martelli MP, Liso A, Gorello P, Mandelli F, Mecucci C, Martelli MF. 2007. Translocations and mutations involving the nucleophosmin (NPM1) gene in lymphomas and leukemias. *Haematologica* **92**: 519-532.

Fan JY, Rangasamy D, Luger K, Tremethick DJ. 2004. H2A.Z alters the nucleosome surface to promote HP1alpha-mediated chromatin fiber folding. *Mol Cell* **16**: 655-661.

Fan X, Wen L, Li Y, Lou L, Liu W, Zhang J. 2017. The expression profile and prognostic value of APE/Ref-1 and NPM1 in high-grade serous ovarian adenocarcinoma. *Apmis* **125**: 857-862.

Fang HT, El Farran CA, Xing QR, Zhang LF, Li H, Lim B, Loh YH. 2018. Global H3.3 dynamic deposition defines its bimodal role in cell fate transition. *Nat Commun* **9**: 1537.

Fankhauser C, Izaurralde E, Adachi Y, Wingfield P, Laemmli UK. 1991. Specific complex of human immunodeficiency virus type 1 rev and nucleolar B23 proteins: dissociation by the Rev response element. *Mol Cell Biol* **11**: 2567-2575.

Farnung L, Vos SM, Cramer P. 2018. Structure of transcribing RNA polymerase II-nucleosome complex. *Nat Commun* **9**: 5432.

Fernandes JCR, Acuña SM, Aoki JI, Floeter-Winter LM, Muxel SM. 2019. Long Non-Coding RNAs in the Regulation of Gene Expression: Physiology and Disease. *Noncoding RNA* **5**.

Fernandez-Capetillo O, Mahadevaiah SK, Celeste A, Romanienko PJ, Camerini-Otero RD, Bonner WM, Manova K, Burgoyne P, Nussenzweig A. 2003. H2AX is required for chromatin remodeling and inactivation of sex chromosomes in male mouse meiosis. *Dev Cell* **4**: 497-508.

Feuerstein N, Chan PK, Mond JJ. 1988a. Identification of numatrin, the nuclear matrix protein associated with induction of mitogenesis, as the nucleolar protein B23. Implication for the role of the nucleolus in early transduction of mitogenic signals. *J Biol Chem* **263**: 10608-10612.

Feuerstein N, Mond JJ. 1987. "Numatrin," a nuclear matrix protein associated with induction of proliferation in B lymphocytes. *J Biol Chem* **262**: 11389-11397.

Feuerstein N, Mond JJ, Kinchington PR, Hickey R, Karjalainen Lindsberg ML, Hay I, Ruyechan WT. 1990. Evidence for DNA binding activity of numatrin (B23), a cell cycle-regulated nuclear matrix protein. *Biochim Biophys Acta* **1087**: 127-136.

Feuerstein N, Spiegel S, Mond JJ. 1988b. The nuclear matrix protein, numatrin (B23), is associated with growth factor-induced mitogenesis in Swiss 3T3 fibroblasts and with T lymphocyte proliferation stimulated by lectins and anti-T cell antigen receptor antibody. *J Cell Biol* **107**: 1629-1642.

Fierz B, Poirier MG. 2019. Biophysics of Chromatin Dynamics. *Annu Rev Biophys* **48**: 321-345.

Filipescu D, Szenker E, Almouzni G. 2013. Developmental roles of histone H3 variants and their chaperones. *Trends Genet* **29**: 630-640.

Finn RM, Ellard K, Eirin-Lopez JM, Ausio J. 2012. Vertebrate nucleoplasmin and NASP: egg histone storage proteins with multiple chaperone activities. *Faseb j* **26**: 4788-4804.

Finsterbusch T, Steinfeldt T, Doberstein K, Rodner C, Mankertz A. 2009. Interaction of the replication proteins and the capsid protein of porcine circovirus type 1 and 2 with host proteins. *Virology* **386**: 122-131.

Fleming AB, Kao CF, Hillyer C, Pikaart M, Osley MA. 2008. H2B ubiquitylation plays a role in nucleosome dynamics during transcription elongation. *Mol Cell* **31**: 57-66.

Flores O, Lu H, Reinberg D. 1992. Factors involved in specific transcription by mammalian RNA polymerase II. Identification and characterization of factor IIH. *J Biol Chem* **267**: 2786-2793.

Fon Tacer K, Montoya MC, Oatley MJ, Lord T, Oatley JM, Klein J, Ravichandran R, Tillman H, Kim M, Connelly JP et al. 2019. MAGE cancer-testis antigens protect the mammalian germline under environmental stress. *Sci Adv* **5**: eaav4832.

Fondell JD, Ge H, Roeder RG. 1996. Ligand induction of a transcriptionally active thyroid hormone receptor coactivator complex. *Proc Natl Acad Sci U S A* **93**: 8329-8333.

Formosa T. 2012. The role of FACT in making and breaking nucleosomes. *Biochim Biophys Acta* **1819**: 247-255.

Francis NJ, Kingston RE, Woodcock CL. 2004. Chromatin compaction by a polycomb group protein complex. *Science* **306**: 1574-1577.

Frazier MW, He X, Wang J, Gu Z, Cleveland JL, Zambetti GP. 1998. Activation of c-myc gene expression by tumor-derived p53 mutants requires a discrete C-terminal domain. *Mol Cell Biol* **18**: 3735-3743.

Freed-Pastor WA, Mizuno H, Zhao X, Langerod A, Moon SH, Rodriguez-Barrueco R, Barsotti A, Chicas A, Li W, Polotskaia A et al. 2012. Mutant p53 disrupts mammary tissue architecture via the mevalonate pathway. *Cell* **148**: 244-258.

Freed-Pastor WA, Prives C. 2012. Mutant p53: one name, many proteins. *Genes Dev* **26**: 1268-1286.

Frehlick LJ, Eirin-Lopez JM, Ausio J. 2007. New insights into the nucleophosmin/nucleoplasmin family of nuclear chaperones. *Bioessays* **29**: 49-59.

Frehlick LJ, Eirin-Lopez JM, Jeffery ED, Hunt DF, Ausio J. 2006. The characterization of amphibian nucleoplasmins yields new insight into their role in sperm chromatin remodeling. *BMC Genomics* **7**: 99.

Fujiwara S, Nagai H, Jimbo H, Jimbo N, Tanaka T, Inoie M, Nishigori C. 2018. Gene Expression and Methylation Analysis in Melanomas and Melanocytes From the Same Patient: Loss of NPM2 Expression Is a Potential Immunohistochemical Marker for Melanoma. *Front Oncol* **8**: 675.

Fukawa T, Ono M, Matsuo T, Uehara H, Miki T, Nakamura Y, Kanayama HO, Katagiri T. 2012. DDX31 regulates the p53-HDM2 pathway and rRNA gene transcription through its interaction with NPM1 in renal cell carcinomas. *Cancer Res* **72**: 5867-5877.

Fusby B, Kim S, Erickson B, Kim H, Peterson ML, Bentley DL. 2016. Coordination of RNA Polymerase II Pausing and 3' End Processing Factor Recruitment with Alternative Polyadenylation. *Mol Cell Biol* **36**: 295-303.

Gadad SS, Rajan RE, Senapati P, Chatterjee S, Shandilya J, Dash PK, Ranga U, Kundu TK. 2011a. HIV-1 infection induces acetylation of NPM1 that facilitates Tat localization and enhances viral transactivation. *J Mol Biol* **410**: 997-1007.

Gadad SS, Senapati P, Syed SH, Rajan RE, Shandilya J, Swaminathan V, Chatterjee S, Colombo E, Dimitrov S, Pelicci PG et al. 2011b. The multifunctional protein nucleophosmin (NPM1) is a human linker histone H1 chaperone. *Biochemistry* **50**: 2780-2789.

Gadad SS, Shandilya J, Kishore AH, Kundu TK. 2010. NPM3, a member of the nucleophosmin/nucleoplasmin family, enhances activator-dependent transcription. *Biochemistry* **49**: 1355-1357.

Galande S, Kohwi-Shigematsu T. 2000. Caught in the act: binding of Ku and PARP to MARs reveals novel aspects of their functional interaction. *Crit Rev Eukaryot Gene Expr* **10**: 63-72.

Gamble MJ, Frizzell KM, Yang C, Krishnakumar R, Kraus WL. 2010. The histone variant macroH2A1 marks repressed autosomal chromatin, but protects a subset of its target genes from silencing. *Genes Dev* **24**: 21-32.

Garcia-Luis J, Lazar-Stefanita L, Gutierrez-Escribano P, Thierry A, Cournac A, Garcia A, Gonzalez S, Sanchez M, Jarmuz A, Montoya A et al. 2019. FACT mediates cohesin function on chromatin. *Nat Struct Mol Biol* **26**: 970-979.

Garcia-Saez I, Menoni H, Boopathi R, Shukla MS, Soueidan L, Noirclerc-Savoye M, Le Roy A, Skoufias DA, Bednar J, Hamiche A et al. 2018. Structure of an H1-Bound 6-Nucleosome Array Reveals an Untwisted Two-Start Chromatin Fiber Conformation. *Mol Cell* **72**: 902-915.e907.

Garrick D, Sharpe JA, Arkell R, Dobbie L, Smith AJ, Wood WG, Higgs DR, Gibbons RJ. 2006. Loss of Atrx affects trophoblast development and the pattern of X-inactivation in extraembryonic tissues. *PLoS Genet* **2**: e58.

Gates LA, Shi J, Rohira AD, Feng Q, Zhu B, Bedford MT, Sagum CA, Jung SY, Qin J, Tsai MJ et al. 2017. Acetylation on histone H3 lysine 9 mediates a switch from transcription initiation to elongation. *J Biol Chem* **292**: 14456-14472.

Gautier T, Abbott DW, Molla A, Verdel A, Ausio J, Dimitrov S. 2004. Histone variant H2ABbd confers lower stability to the nucleosome. *EMBO Rep* **5**: 715-720.

Geiger JH, Hahn S, Lee S, Sigler PB. 1996. Crystal structure of the yeast TFIIA/TBP/DNA complex. *Science* **272**: 830-836.

Georgel PT, Horowitz-Scherer RA, Adkins N, Woodcock CL, Wade PA, Hansen JC. 2003. Chromatin compaction by human MeCP2. Assembly of novel secondary chromatin structures in the absence of DNA methylation. *J Biol Chem* **278**: 32181-32188.

Giaimo BD, Ferrante F, Herchenrother A, Hake SB, Borggrefe T. 2019. The histone variant H2A.Z in gene regulation. *Epigenetics Chromatin* **12**: 37.

Gibbs-Seymour I, Fontana P, Rack JGM, Ahel I. 2016. HPF1/C4orf27 Is a PARP-1-Interacting Protein that Regulates PARP-1 ADP-Ribosylation Activity. *Mol Cell* **62**: 432-442.

Gibson BA, Doolittle LK, Schneider MWG, Jensen LE, Gamarra N, Henry L, Gerlich DW, Redding S, Rosen MK. 2019. Organization of Chromatin by Intrinsic and Regulated Phase Separation. *Cell* **179**: 470-484.e421.

Gimenez M, Marie SK, Oba-Shinjo SM, Uno M, da Silva R, Laure HJ, Izumi C, Otake A, Chammas R, Rosa JC. 2012. Quantitative proteomic analysis and functional studies reveal that nucleophosmin is involved in cell death in glioblastoma cell line transfected with siRNA. *Proteomics* **12**: 2632-2640.

Gimenez M, Souza VC, Izumi C, Barbieri MR, Chammas R, Oba-Shinjo SM, Uno M, Marie SK, Rosa JC. 2010. Proteomic analysis of low- to high-grade astrocytomas reveals an alteration of the expression level of raf kinase inhibitor protein and nucleophosmin. *Proteomics* **10**: 2812-2821.

Gjerstorff MF, Ditzel HJ. 2008. An overview of the GAGE cancer/testis antigen family with the inclusion of newly identified members. *Tissue Antigens* **71**: 187-192.

Gnatt AL, Cramer P, Fu J, Bushnell DA, Kornberg RD. 2001. Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. *Science* **292**: 1876-1882.

Goldberg AD, Banaszynski LA, Noh KM, Lewis PW, Elsaesser SJ, Stadler S, Dewell S, Law M, Guo X, Li X et al. 2010. Distinct factors control histone variant H3.3 localization at specific genomic regions. *Cell* **140**: 678-691.

Gordon S, Akopyan G, Garban H, Bonavida B. 2006. Transcription factor YY1: structure, function, and therapeutic implications in cancer biology. *Oncogene* **25**: 1125-1142.

Greber BJ, Toso DB, Fang J, Nogales E. 2019. The complete structure of the human TFIIH core complex. *Elife* **8**.

Grinstein E, Wernet P, Snijders PJ, Rosl F, Weinert I, Jia W, Kraft R, Schewe C, Schwabe M, Hauptmann S et al. 2002. Nucleolin as activator of human papillomavirus type 18 oncogene transcription in cervical cancer. *J Exp Med* **196**: 1067-1078.

Grisendi S, Bernardi R, Rossi M, Cheng K, Khandker L, Manova K, Pandolfi PP. 2005. Role of nucleophosmin in embryonic development and tumorigenesis. *Nature* **437**: 147-153.

Grisendi S, Mecucci C, Falini B, Pandolfi PP. 2006. Nucleophosmin and cancer. *Nat Rev Cancer* **6**: 493-505.

Gronroos E, Terentiev AA, Punga T, Ericsson J. 2004. YY1 inhibits the activation of the p53 tumor suppressor in response to genotoxic stress. *Proc Natl Acad Sci U S A* **101**: 12165-12170.

Grugan KD, Vega ME, Wong GS, Diehl JA, Bass AJ, Wong KK, Nakagawa H, Rustgi AK. 2013. A common p53 mutation (R175H) activates c-Met receptor tyrosine kinase to enhance tumor cell invasion. *Cancer Biol Ther* **14**: 853-859.

Grummitt CG, Townsley FM, Johnson CM, Warren AJ, Bycroft M. 2008. Structural consequences of nucleophosmin mutations in acute myeloid leukemia. *J Biol Chem* **283**: 23326-23332.

Gualberto A, Baldwin AS, Jr. 1995. p53 and Sp1 interact and cooperate in the tumor necrosis factor-induced transcriptional activation of the HIV-1 long terminal repeat. *J Biol Chem* **270**: 19680-19683.

Gudipati RK, Villa T, Boulay J, Libri D. 2008. Phosphorylation of the RNA polymerase II C-terminal domain dictates transcription termination choice. *Nat Struct Mol Biol* **15**: 786-794.

Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA. 2007. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* **130**: 77-88.

Guermah M, Palhan VB, Tackett AJ, Chait BT, Roeder RG. 2006. Synergistic functions of SII and p300 in productive activator-dependent transcription of chromatin templates. *Cell* **125**: 275-286.

Guery L, Benikhlef N, Gautier T, Paul C, Jego G, Dufour E, Jacquel A, Cally R, Manoury B, Vanden Berghe T et al. 2011. Fine-tuning nucleophosmin in macrophage differentiation and activation. *Blood* **118**: 4694-4704.

Guo Y, Liu S, Wang P, Zhang H, Wang F, Bing L, Gao J, Yang J, Hao A. 2014. Granulocyte colony-stimulating factor improves neuron survival in experimental spinal cord injury by regulating nucleophosmin-1 expression. *J Neurosci Res* **92**: 751-760.

Gurard-Levin ZA, Quivy JP, Almouzni G. 2014. Histone chaperones: assisting histone traffic and nucleosome dynamics. *Annu Rev Biochem* **83**: 487-517.

Gurumurthy M, Tan CH, Ng R, Zeiger L, Lau J, Lee J, Dey A, Philp R, Li Q, Lim TM et al. 2008. Nucleophosmin interacts with HEXIM1 and regulates RNA polymerase II transcription. *J Mol Biol* **378**: 302-317.

Haberle V, Stark A. 2018. Eukaryotic core promoters and the functional basis of transcription initiation. *Nat Rev Mol Cell Biol* **19**: 621-637.

Hahn S. 2004. Structure and mechanism of the RNA polymerase II transcription machinery. *Nat Struct Mol Biol* **11**: 394-403.

Haindl M, Harasim T, Eick D, Muller S. 2008. The nucleolar SUMO-specific protease SENP3 reverses SUMO modification of nucleophosmin and is required for rRNA processing. *EMBO Rep* **9**: 273-279.

Halazonetis TD, Georgopoulos K, Greenberg ME, Leder P. 1988. c-Jun dimerizes with itself and with c-Fos, forming complexes of different DNA binding affinities. *Cell* **55**: 917-924.

Hallor KH, Sciot R, Staaf J, Heidenblad M, Rydholm A, Bauer HC, Astrom K, Domanski HA, Meis JM, Kindblom LG et al. 2009. Two genetic pathways, t(1;10) and amplification of 3p11-12, in myxoinflammatory fibroblastic sarcoma, haemosiderotic fibrolipomatous tumour, and morphologically similar lesions. *J Pathol* **217**: 716-727.

Hammond CM, Stromme CB, Huang H, Patel DJ, Groth A. 2017. Histone chaperone networks shaping chromatin function. *Nat Rev Mol Cell Biol* **18**: 141-158.

Hansen SK, Takada S, Jacobson RH, Lis JT, Tjian R. 1997. Transcription properties of a cell type-specific TATA-binding protein, TRF. *Cell* **91**: 71-83.

Happel N, Schulze E, Doenecke D. 2005. Characterisation of human histone H1x. *Biol Chem* **386**: 541-551.

Haque M, Li J, Huang YH, Almowaled M, Barger CJ, Karpf AR, Wang P, Chen W, Turner SD, Lai R. 2019. NPM-ALK Is a Key Regulator of the Oncoprotein FOXM1 in ALK-Positive Anaplastic Large Cell Lymphoma. *Cancers (Basel)* **11**.

Hardy S, Jacques PE, Gevry N, Forest A, Fortin ME, Laflamme L, Gaudreau L, Robert F. 2009. The euchromatic and heterochromatic landscapes are shaped by antagonizing effects of transcription on H2A.Z deposition. *PLoS Genet* **5**: e1000687.

Harper TM, Taatjes DJ. 2018. The complex structure and function of Mediator. *J Biol Chem* **293**: 13778-13785.

Hartford SA, Luo Y, Southard TL, Min IM, Lis JT, Schimenti JC. 2011. Minichromosome maintenance helicase paralog MCM9 is dispensible for DNA replication but functions in germ-line stem cells and tumor suppression. *Proc Natl Acad Sci U S A* **108**: 17702-17707.

Hashimoto H, Sudo T, Mikami Y, Otani M, Takano M, Tsuda H, Itamochi H, Katabuchi H, Ito M, Nishimura R. 2008. Germ cell specific protein VASA is over-expressed in epithelial ovarian cancer and disrupts DNA damage-induced G2 checkpoint. *Gynecol Oncol* **111**: 312-319.

He J, Xiang Z, Xiao J, Xiao H, Liu L. 2016. [The poor chemotherapeutic efficacy in lung adenocarcinoma overexpressing c-Src and nucleophosmin/B23(NPM1)]. *Xi Bao Yu Fen Zi Mian Yi Xue Za Zhi* **32**: 1378-1381.

He Y, Fang J, Taatjes DJ, Nogales E. 2013. Structural visualization of key steps in human transcription initiation. *Nature* **495**: 481-486.

Healy S, Khan P, Davie JR. 2013. Immediate early response genes and cell transformation. *Pharmacol Ther* **137**: 64-77.

Heaphy CM, de Wilde RF, Jiao Y, Klein AP, Edil BH, Shi C, Bettegowda C, Rodriguez FJ, Eberhart CG, Hebbar S et al. 2011. Altered telomeres in tumors with ATRX and DAXX mutations. *Science* **333**: 425.

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**: 576-589.

Henikoff S, Smith MM. 2015. Histone variants and epigenetics. *Cold Spring Harb Perspect Biol* **7**: a019364.

Hergeth SP, Schneider R. 2015. The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle. *EMBO Rep* **16**: 1439-1453.

Hernandez-Verdun D, Gautier T. 1994. The chromosome periphery during mitosis. *Bioessays* **16**: 179-185.

Herrera JE, Savkur R, Olson MO. 1995. The ribonuclease activity of nucleolar protein B23. *Nucleic Acids Res* **23**: 3974-3979.

Herschman HR. 1991. Primary response genes induced by growth factors and tumor promoters. *Annu Rev Biochem* **60**: 281-319.

Hierro A, Arizmendi JM, Banuelos S, Prado A, Muga A. 2002. Electrostatic interactions at the C-terminal domain of nucleoplasmin modulate its chromatin decondensation activity. *Biochemistry* **41**: 6408-6413.

Hierro A, Arizmendi JM, De Las Rivas J, Urbaneja MA, Prado A, Muga A. 2001. Structural and functional properties of Escherichia coli-derived nucleoplasmin. A comparative study of recombinant and natural proteins. *Eur J Biochem* **268**: 1739-1748.

Hingorani K, Szebeni A, Olson MO. 2000. Mapping the functional domains of nucleolar protein B23. *J Biol Chem* **275**: 24451-24457.

Hirose Y, Manley JL. 2000. RNA polymerase II and the integration of nuclear events. *Genes Dev* **14**: 1415-1429.

Hirsch S, Blatte TJ, Grasedieck S, Cocciardi S, Rouhi A, Jongen-Lavrencic M, Paschka P, Kronke J, Gaidzik VI, Dohner H et al. 2017. Circular RNAs of the nucleophosmin (NPM1) gene in acute myeloid leukemia. *Haematologica* **102**: 2039-2047.

Hisaoka M, Ueshima S, Murano K, Nagata K, Okuwaki M. 2010. Regulation of nucleolar chromatin by B23/nucleophosmin jointly depends upon its RNA binding activity and transcription factor UBF. *Mol Cell Biol* **30**: 4952-4964.

Hnisz D, Shrinivas K, Young RA, Chakraborty AK, Sharp PA. 2017. A Phase Separation Model for Transcriptional Control. *Cell* **169**: 13-23.

Ho CK, Sriskanda V, McCracken S, Bentley D, Schwer B, Shuman S. 1998. The guanylyltransferase domain of mammalian mRNA capping enzyme binds to the phosphorylated carboxyl-terminal domain of RNA polymerase II. *J Biol Chem* **273**: 9577-9585.

Ho ST, Jin R, Cheung DH, Huang JJ, Shaw PC. 2019. The PinX1/NPM interaction associates with hTERT in early-S phase and facilitates telomerase activation. *Cell Biosci* **9**: 47.

Hoffman MM, Ernst J, Wilder SP, Kundaje A, Harris RS, Libbrecht M, Giardine B, Ellenbogen PM, Bilmes JA, Birney E et al. 2013. Integrative annotation of chromatin elements from ENCODE data. *Nucleic Acids Res* **41**: 827-841.

Hofmann O, Caballero OL, Stevenson BJ, Chen YT, Cohen T, Chua R, Maher CA, Panji S, Schaefer U, Kruger A et al. 2008. Genome-wide analysis of cancer/testis gene expression. *Proc Natl Acad Sci U S A* **105**: 20422-20427.

Holmberg Olausson K, Elsir T, Moazemi Goudarzi K, Nister M, Lindstrom MS. 2015. NPM1 histone chaperone is upregulated in glioblastoma to promote cell survival and maintain nucleolar shape. *Sci Rep* **5**: 16495.

Holmberg Olausson K, Nister M, Lindstrom MS. 2014. Loss of nucleolar histone chaperone NPM1 triggers rearrangement of heterochromatin and synergizes with a deficiency in DNA methyltransferase DNMT3A to drive ribosomal DNA transcription. *J Biol Chem* **289**: 34601-34619.

Holstege FC, Fiedler U, Timmers HT. 1997. Three transitions in the RNA polymerase II transcription complex during initiation. *Embo j* **16**: 7468-7480.

Honda R, Yasuda H. 1999. Association of p19(ARF) with Mdm2 inhibits ubiquitin ligase activity of Mdm2 for tumor suppressor p53. *Embo j* **18**: 22-27.

Hondele M, Ladurner AG. 2011. The chaperone-histone partnership: for the greater good of histone traffic and chromatin plasticity. *Curr Opin Struct Biol* **21**: 698-708.

Houlard M, Berlivet S, Probst AV, Quivy JP, Hery P, Almouzni G, Gerard M. 2006. CAF-1 is essential for heterochromatin organization in pluripotent embryonic cells. *PLoS Genet* **2**: e181.

Hsin JP, Manley JL. 2012. The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes Dev* **26**: 2119-2137.

Hu E, Mueller E, Oliviero S, Papaioannou VE, Johnson R, Spiegelman BM. 1994. Targeted disruption of the c-fos gene demonstrates c-fos-dependent and -independent pathways for gene expression stimulated by growth factors or oncogenes. *Embo j* **13**: 3094-3103.

Hu Z, Huang G, Sadanandam A, Gu S, Lenburg ME, Pai M, Bayani N, Blakely EA, Gray JW, Mao JH. 2010. The expression level of HJURP has an independent prognostic impact and predicts the sensitivity to radiotherapy in breast cancer. *Breast Cancer Res* **12**: R18.

Huang H, Stromme CB, Saredi G, Hodl M, Strandsby A, Gonzalez-Aguilera C, Chen S, Groth A, Patel DJ. 2015. A unique binding mode enables MCM2 to chaperone histones H3-H4 at replication forks. *Nat Struct Mol Biol* **22**: 618-626.

Huang N, Negi S, Szebeni A, Olson MO. 2005. Protein NPM3 interacts with the multifunctional nucleolar protein B23/nucleophosmin and inhibits ribosome biogenesis. *J Biol Chem* **280**: 5496-5502.

Huang WH, Yung BY, Syu WJ, Lee YH. 2001. The nucleolar phosphoprotein B23 interacts with hepatitis delta antigens and modulates the hepatitis delta virus RNA replication. *J Biol Chem* **276**: 25166-25175.

Hughes CM, Rozenblatt-Rosen O, Milne TA, Copeland TD, Levine SS, Lee JC, Hayes DN, Shanmugam KS, Bhattacharjee A, Biondi CA et al. 2004. Menin associates with a trithorax family histone methyltransferase complex and with the hoxc8 locus. *Mol Cell* **13**: 587-597.

Huo XM, Meng LF, Jiang T, Li M, Sun FZ, Sun B, Li JK. 2018. Real-time observation of nucleoplasmin-mediated DNA decondensation and condensation reveals its specific functions as a chaperone. *Biochim Biophys Acta Gene Regul Mech* **1861**: 743-751.

Huynh LM, Shinagawa T, Ishii S. 2016. Two Histone Variants TH2A and TH2B Enhance Human Induced Pluripotent Stem Cell Generation. *Stem Cells Dev* **25**: 251-258.

Imbalzano AN, Zaret KS, Kingston RE. 1994. Transcription factor (TF) IIB and TFIIA can independently increase the affinity of the TATA-binding protein for DNA. *J Biol Chem* **269**: 8280-8286.

Inoue A, Aoki F. 2010. Role of the nucleoplasmin 2 C-terminal domain in the formation of nucleolus-like bodies in mouse oocytes. *Faseb j* **24**: 485-494.

Inoue A, Ogushi S, Saitou M, Suzuki MG, Aoki F. 2011. Involvement of mouse nucleoplasmin 2 in the decondensation of sperm chromatin after fertilization. *Biol Reprod* **85**: 70-77.

Inouye CJ, Seto E. 1994. Relief of YY1-induced transcriptional repression by protein-protein interaction with the nucleolar phosphoprotein B23. *J Biol Chem* **269**: 6506-6510.

Ishibashi T, Li A, Eirin-Lopez JM, Zhao M, Missiaen K, Abbott DW, Meistrich M, Hendzel MJ, Ausio J. 2010. H2A.Bbd: an X-chromosome-encoded histone involved in mammalian spermiogenesis. *Nucleic Acids Res* **38**: 1780-1789.

Itahana K, Bhat KP, Jin A, Itahana Y, Hawke D, Kobayashi R, Zhang Y. 2003. Tumor suppressor ARF degrades B23, a nucleolar protein involved in ribosome biogenesis and cell proliferation. *Mol Cell* **12**: 1151-1164.

Ito M, Yuan CX, Okano HJ, Darnell RB, Roeder RG. 2000. Involvement of the TRAP220 component of the TRAP/SMCC coactivator complex in embryonic development and thyroid hormone action. *Mol Cell* **5**: 683-693.

Ito T, Tyler JK, Bulger M, Kobayashi R, Kadonaga JT. 1996. ATP-facilitated chromatin assembly with a nucleoplasmin-like protein from Drosophila melanogaster. *J Biol Chem* **271**: 25041-25048.

Jackson JD, Falciano VT, Gorovsky MA. 1996. A likely histone H2A.F/Z variant in Saccharomyces cerevisiae. *Trends Biochem Sci* **21**: 466-467.

Jacobson RH, Ladurner AG, King DS, Tjian R. 2000. Structure and function of a human TAFII250 double bromodomain module. *Science* **288**: 1422-1425.

James TC, Elgin SC. 1986. Identification of a nonhistone chromosomal protein associated with heterochromatin in Drosophila melanogaster and its gene. *Mol Cell Biol* **6**: 3862-3872.

Jeong EG, Lee SH, Yoo NJ. 2007. Absence of nucleophosmin 1 (NPM1) gene mutations in common solid cancers. *Apmis* **115**: 341-346.

Jeronimo C, Robert F. 2017. The Mediator Complex: At the Nexus of RNA Polymerase II Transcription. *Trends Cell Biol* **27**: 765-783.

Jeronimo C, Watanabe S, Kaplan CD, Peterson CL, Robert F. 2015. The Histone Chaperones FACT and Spt6 Restrict H2A.Z from Intragenic Locations. *Mol Cell* **58**: 1113-1123.

Jian Y, Gao Z, Sun J, Shen Q, Feng F, Jing Y, Yang C. 2009. RNA aptamers interfering with nucleophosmin oligomerization induce apoptosis of cancer cells. *Oncogene* **28**: 4201-4211.

Jiang PS, Chang JH, Yung BY. 2000. Different kinases phosphorylate nucleophosmin/B23 at different sites during G(2) and M phases of the cell cycle. *Cancer Lett* **153**: 151-160.

Jiang S, Qi Y, He R, Huang Y, Liu Z, Ma Y, Guo X, Shao Y, Sun Z, Ruan Q. 2015. Human cytomegalovirus microRNA miR-US25-1-5p inhibits viral replication by targeting multiple cellular genes during infection. *Gene* **570**: 108-114.

Jiao Y, Shi C, Edil BH, de Wilde RF, Klimstra DS, Maitra A, Schulick RD, Tang LH, Wolfgang CL, Choti MA et al. 2011. DAXX/ATRX, MEN1, and mTOR pathway genes are frequently altered in pancreatic neuroendocrine tumors. *Science* **331**: 1199-1203.

Jin C, Kato K, Chimura T, Yamasaki T, Nakade K, Murata T, Li H, Pan J, Zhao M, Sun K et al. 2006. Regulation of histone acetylation and nucleosome assembly by transcription factor JDP2. *Nat Struct Mol Biol* **13**: 331-338.

Jin C, Zang C, Wei G, Cui K, Peng W, Zhao K, Felsenfeld G. 2009. H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. *Nat Genet* **41**: 941-945.

Joerger AC, Fersht AR. 2007a. Structural biology of the tumor suppressor p53 and cancer-associated mutants. *Adv Cancer Res* **97**: 1-23.

Joerger AC, Fersht AR. 2007b. Structure-function-rescue: the diverse nature of common p53 cancer mutants. *Oncogene* **26**: 2226-2242.

Joerger AC, Fersht AR. 2008. Structural biology of the tumor suppressor p53. *Annu Rev Biochem* **77**: 557-582.

Joerger AC, Fersht AR. 2016. The p53 Pathway: Origins, Inactivation in Cancer, and Emerging Therapeutic Approaches. *Annu Rev Biochem* **85**: 375-404.

Johansson H, Simonsson S. 2010. Core transcription factors, Oct4, Sox2 and Nanog, individually form complexes with nucleophosmin (Npm1) to control embryonic stem (ES) cell fate determination. *Aging (Albany NY)* **2**: 815-822.

Johnson ML, Wang R, Sperry AO. 2018. Novel localization of Aurora A kinase in mouse testis suggests multiple roles in spermatogenesis. *Biochem Biophys Res Commun* **503**: 51-55.

Jonkers I, Kwak H, Lis JT. 2014. Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. *Elife* **3**: e02407.

Jullien J, Astrand C, Szenker E, Garrett N, Almouzni G, Gurdon JB. 2012. HIRA dependent H3.3 deposition is required for transcriptional reprogramming following nuclear transfer to Xenopus oocytes. *Epigenetics Chromatin* **5**: 17.

Juo ZS, Chiu TK, Leiberman PM, Baikalov I, Berk AJ, Dickerson RE. 1996. How proteins recognize the TATA box. *J Mol Biol* **261**: 239-254.

Juven-Gershon T, Hsu JY, Theisen JW, Kadonaga JT. 2008. The RNA polymerase II core promoter - the gateway to transcription. *Curr Opin Cell Biol* **20**: 253-259.

Kaimori JY, Maehara K, Hayashi-Takanaka Y, Harada A, Fukuda M, Yamamoto S, Ichimaru N, Umehara T, Yokoyama S, Matsuda R et al. 2016. Histone H4 lysine 20 acetylation is associated with gene repression in human cells. *Sci Rep* **6**: 24318.

Kalra RS, Bapat SA. 2013. Enhanced levels of double-strand DNA break repair proteins protect ovarian cancer cells against genotoxic stress-induced apoptosis. *J Ovarian Res* **6**: 66.

Kamada K, Shu F, Chen H, Malik S, Stelzer G, Roeder RG, Meisterernst M, Burley SK. 2001. Crystal structure of negative cofactor 2 recognizing the TBP-DNA transcription complex. *Cell* **106**: 71-81.

Kamakaka RT, Biggins S. 2005. Histone variants: deviants? *Genes Dev* **19**: 295-310.

Kang JU, Koo SH, Kwon KC, Park JW. 2010. AMY2A: a possible tumor-suppressor gene of 1p21.1 loss in gastric carcinoma. *Int J Oncol* **36**: 1429-1435.

Kang YJ, Olson MO, Busch H. 1974. Phosphorylation of acid-soluble proteins in isolated nucleoli of Novikoff hepatoma ascites cells. Effects of divalent cations. *J Biol Chem* **249**: 5580-5585.

Kanhere A, Bansal M. 2005. Structural properties of promoters: similarities and differences between prokaryotes and eukaryotes. *Nucleic Acids Res* **33**: 3165-3175.

Kaowinn S, Seo EJ, Heo W, Bae JH, Park EJ, Lee S, Kim YJ, Koh SS, Jang IH, Shin DH et al. 2019. Cancer upregulated gene 2 (CUG2), a novel oncogene, promotes stemness-like properties via the NPM1-TGF-beta signaling axis. *Biochem Biophys Res Commun* **514**: 1278-1284.

Kappes F, Waldmann T, Mathew V, Yu J, Zhang L, Khodadoust MS, Chinnaiyan AM, Luger K, Erhardt S, Schneider R et al. 2011. The DEK oncoprotein is a Su(var) that is essential to heterochromatin integrity. *Genes Dev* **25**: 673-678.

Karhemo PR, Rivinoja A, Lundin J, Hyvonen M, Chernenko A, Lammi J, Sihto H, Lundin M, Heikkila P, Joensuu H et al. 2011. An extensive tumor array analysis supports tumor suppressive role for nucleophosmin in breast cancer. *Am J Pathol* **179**: 1004-1014.

Karin M, Liu Z, Zandi E. 1997. AP-1 function and regulation. *Curr Opin Cell Biol* **9**: 240-246.

Karthigeyan D, Bose A, Boopathi R, Rao VJ, Shima H, Bharathy N, Igarashi K, Taneja R, Kundu TK. 2019. Aurora kinase A mediated phosphorylation of mPOU is critical for skeletal muscle differentiation. *bioRxiv*: 589754.

Katabami M, Donninger H, Hommura F, Leaner VD, Kinoshita I, Chick JF, Birrer MJ. 2005. Cyclin A is a c-Jun target gene and is necessary for c-Jun-induced anchorage-independent growth in RAT1a cells. *J Biol Chem* **280**: 16728-16738.

Kato T, Sato N, Hayama S, Yamabuki T, Ito T, Miyamoto M, Kondo S, Nakamura Y, Daigo Y. 2007. Activation of Holliday junction recognizing protein involved in the chromosomal stability and immortality of cancer cells. *Cancer Res* **67**: 8544-8553.

Kawasaki K, Philpott A, Avilion AA, Berrios M, Fisher PA. 1994. Chromatin decondensation in Drosophila embryo extracts. *J Biol Chem* **269**: 10169-10176.

Kaypee S, Sahadevan SA, Patil S, Ghosh P, Roy NS, Roy S, Kundu TK. 2018a. Mutant and Wild-Type Tumor Suppressor p53 Induces p300 Autoacetylation. *iScience* **4**: 260-272.

Kaypee S, Sahadevan SA, Sudarshan D, Halder Sinha S, Patil S, Senapati P, Kodaganur GS, Mohiyuddin A, Dasgupta D, Kundu TK. 2018b. Oligomers of human histone chaperone NPM1 alter p300/KAT3B folding to induce autoacetylation. *Biochim Biophys Acta Gen Subj* **1862**: 1729-1741.

Kazerouninia A, Ngo B, Martinson HG. 2010. Poly(A) signal-dependent degradation of unprocessed nascent transcripts accompanies poly(A) signal-dependent transcriptional pausing in vitro. *Rna* **16**: 197-210.

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. 2002. The human genome browser at UCSC. *Genome Res* **12**: 996-1006.

Kettenberger H, Armache KJ, Cramer P. 2003. Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage. *Cell* **114**: 347-357.

Kharchenko PV, Alekseyenko AA, Schwartz YB, Minoda A, Riddle NC, Ernst J, Sabo PJ, Larschan E, Gorchakov AA, Gu T et al. 2011. Comprehensive analysis of the chromatin landscape in Drosophila melanogaster. *Nature* **471**: 480-485.

Khare SP, Habib F, Sharma R, Gadewal N, Gupta S, Galande S. 2012. HIstome--a relational knowledgebase of human histone proteins and histone modifying enzymes. *Nucleic Acids Res* **40**: D337-342.

Khazaei G, Shamsabadi FT, Yamchi A, Golalipour M, Jhingan GD, Shahbazi M. 2019. Proteomics evaluation of MDA-MB-231 breast cancer cells in response to RNAi-induced silencing of hPTTG. *Life Sci*: 116873.

Kim D, Langmead B, Salzberg SL. 2015. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* **12**: 357-360.

Kim HJ, Jeong SH, Heo JH, Jeong SJ, Kim ST, Youn HD, Han JW, Lee HW, Cho EJ. 2004a. mRNA capping enzyme activity is coupled to an early transcription elongation. *Mol Cell Biol* **24**: 6184-6193.

Kim HJ, Seol JH, Han JW, Youn HD, Cho EJ. 2007. Histone chaperones regulate histone exchange during transcription. *Embo j* **26**: 4467-4474.

Kim J, Parvin JD, Shykind BM, Sharp PA. 1996. A negative cofactor containing Dr1/p19 modulates transcription with TFIIA in a promoter-specific fashion. *J Biol Chem* **271**: 18405-18412.

Kim JL, Nikolov DB, Burley SK. 1993a. Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature* **365**: 520-527.

Kim KH, Yoo BC, Kim WK, Hong JP, Kim K, Song EY, Lee JY, Cho JY, Ku JL. 2014. CD133 and CD133-regulated nucleophosmin linked to 5-fluorouracil susceptibility in human colon cancer cell line SW620. *Electrophoresis* **35**: 522-532.

Kim MP, Lozano G. 2018. Mutant p53 partners in crime. *Cell Death Differ* **25**: 161-168.

Kim MY, Mauro S, Gevry N, Lis JT, Kraus WL. 2004b. NAD+-dependent modulation of chromatin structure and transcription by nucleosome binding properties of PARP-1. *Cell* **119**: 803-814.

Kim TK, Ebright RH, Reinberg D. 2000. Mechanism of ATP-dependent promoter melting by transcription factor IIH. *Science* **288**: 1418-1422.

Kim Y, Geiger JH, Hahn S, Sigler PB. 1993b. Crystal structure of a yeast TBP/TATA-box complex. *Nature* **365**: 512-520.

Kimmins S, Crosio C, Kotaja N, Hirayama J, Monaco L, Hoog C, van Duin M, Gossen JA, Sassone-Corsi P. 2007. Differential functions of the Aurora-B and Aurora-C kinases in mammalian spermatogenesis. *Mol Endocrinol* **21**: 726-739.

Kiran S, Anwar T, Kiran M, Ramakrishna G. 2015. Sirtuin 7 in cell proliferation, stress and disease: Rise of the Seventh Sirtuin! *Cell Signal* **27**: 673-682.

Kireeva N, Lakonishok M, Kireev I, Hirano T, Belmont AS. 2004. Visualization of early chromosome condensation: a hierarchical folding, axial glue model of chromosome structure. *J Cell Biol* **166**: 775-785.

Kleinjan DA, van Heyningen V. 2005. Long-range control of gene expression: emerging mechanisms and disruption in disease. *Am J Hum Genet* **76**: 8-32.

Koga Y, Pelizzola M, Cheng E, Krauthammer M, Sznol M, Ariyan S, Narayan D, Molinaro AM, Halaban R, Weissman SM. 2009. Genome-wide screen of promoter methylation identifies novel markers in melanoma. *Genome Res* **19**: 1462-1470.

Koike A, Nishikawa H, Wu W, Okada Y, Venkitaraman AR, Ohta T. 2010. Recruitment of phosphorylated NPM1 to sites of DNA damage through RNF8-dependent ubiquitin conjugates. *Cancer Res* **70**: 6746-6756.

Kokubo T, Swanson MJ, Nishikawa JI, Hinnebusch AG, Nakatani Y. 1998. The yeast TAF145 inhibitory domain and TFIIA competitively bind to TATA-binding protein. *Mol Cell Biol* **18**: 1003-1012.

Komarnitsky P, Cho EJ, Buratowski S. 2000. Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev* **14**: 2452-2460.

Kondo T, Minamino N, Nagamura-Inoue T, Matsumoto M, Taniguchi T, Tanaka N. 1997. Identification and characterization of nucleophosmin/B23/numatrin which binds the anti-oncogenic transcription factor IRF-1 and manifests oncogenic activity. *Oncogene* **15**: 1275-1281.

Korber P, Barbaric S, Luckenbach T, Schmid A, Schermer UJ, Blaschke D, Horz W. 2006. The histone chaperone Asf1 increases the rate of histone eviction at the yeast PHO5 and PHO8 promoters. *J Biol Chem* **281**: 5539-5545.

Korgaonkar C, Hagen J, Tompkins V, Frazier AA, Allamargot C, Quelle FW, Quelle DE. 2005. Nucleophosmin (B23) targets ARF to nucleoli and inhibits its function. *Mol Cell Biol* **25**: 1258-1271.

Kornberg RD. 1977. Structure of chromatin. *Annu Rev Biochem* **46**: 931-954.

Kornberg RD, Thomas JO. 1974. Chromatin structure; oligomers of the histones. *Science* **184**: 865-868.

Kostrewa D, Zeller ME, Armache KJ, Seizl M, Leike K, Thomm M, Cramer P. 2009. RNA polymerase II-TFIIB structure and mechanism of transcription initiation. *Nature* **462**: 323-330.

Kouzarides T, Ziff E. 1988. The role of the leucine zipper in the fos-jun interaction. *Nature* **336**: 646-651.

Kozlowska M, Tarczewska A, Jakob M, Bystranowska D, Taube M, Kozak M, Czarnocki-Cieciura M, Dziembowski A, Orlowski M, Tkocz K et al. 2017. Nucleoplasmin-like domain of FKBP39 from Drosophila melanogaster forms a tetramer with partly disordered tentacle-like C-terminal segments. *Sci Rep* **7**: 40405.

Krasnov AN, Mazina MY, Nikolenko JV, Vorobyeva NE. 2016. On the way of revealing coactivator complexes cross-talk during transcriptional activation. *Cell Biosci* **6**: 15.

Krause A, Hoffmann I. 2010. Polo-like kinase 2-dependent phosphorylation of NPM/B23 on serine 4 triggers centriole duplication. *PLoS One* **5**: e9849.

Kroeger H, Jelinek J, Estecio MR, He R, Kondo K, Chung W, Zhang L, Shen L, Kantarjian HM, Bueso-Ramos CE et al. 2008. Aberrant CpG island methylation in acute myeloid leukemia is accentuated at relapse. *Blood* **112**: 1366-1373.

Krogan NJ, Dover J, Wood A, Schneider J, Heidt J, Boateng MA, Dean K, Ryan OW, Golshani A, Johnston M et al. 2003. The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. *Mol Cell* **11**: 721-729.

Krohne G, Franke WW. 1980. A major soluble acidic protein located in nuclei of diverse vertebrate species. *Exp Cell Res* **129**: 167-189.

Kroumpouzos G, Eberle J, Garbe C, Orfanos CE. 1994. P53 mutation and c-fos overexpression are associated with detection of the antigen VLA-2 in human melanoma cell lines. *Pigment Cell Res* **7**: 348-353.

Kuehner JN, Pearson EL, Moore C. 2011. Unravelling the means to an end: RNA polymerase II transcription termination. *Nat Rev Mol Cell Biol* **12**: 283-294.

Kujirai T, Horikoshi N, Sato K, Maehara K, Machida S, Osakabe A, Kimura H, Ohkawa Y, Kurumizaka H. 2016. Structure and function of human histone H3.Y nucleosome. *Nucleic Acids Res* **44**: 6127-6141.

Kulaeva OI, Gaykalova DA, Studitsky VM. 2007. Transcription through chromatin by RNA polymerase II: histone displacement and exchange. *Mutat Res* **618**: 116-129.

Kunchala P, Kuravi S, Jensen R, McGuirk J, Balusu R. 2018. When the good go bad: Mutant NPM1 in acute myeloid leukemia. *Blood Rev* **32**: 167-183.

Kundu TK, Palhan VB, Wang Z, An W, Cole PA, Roeder RG. 2000. Activator-dependent transcription from chromatin in vitro involving targeted histone acetylation by p300. *Mol Cell* **6**: 551-561.

Kuo LJ, Yang LX. 2008. Gamma-H2AX - a novel biomarker for DNA double-strand breaks. *In Vivo* **22**: 305-309.

Kuo ML, den Besten W, Bertwistle D, Roussel MF, Sherr CJ. 2004. N-terminal polyubiquitination and degradation of the Arf tumor suppressor. *Genes Dev* **18**: 1862-1874.

Kuo ML, den Besten W, Thomas MC, Sherr CJ. 2008. Arf-induced turnover of the nucleolar nucleophosmin-associated SUMO-2/3 protease Senp3. *Cell Cycle* **7**: 3378-3387.

Kurki S, Peltonen K, Laiho M. 2004a. Nucleophosmin, HDM2 and p53: players in UV damage incited nucleolar stress response. *Cell Cycle* **3**: 976-979.

Kurki S, Peltonen K, Latonen L, Kiviharju TM, Ojala PM, Meek D, Laiho M. 2004b. Nucleolar protein NPM interacts with HDM2 and protects tumor suppressor protein p53 from HDM2-mediated degradation. *Cancer Cell* **5**: 465-475.

Kutchy NA, Velho A, Menezes ESB, Jacobsen M, Thibaudeau G, Wills RW, Moura A, Kaya A, Perkins A, Memili E. 2017. Testis specific histone 2B is associated with sperm chromatin dynamics and bull fertility-a pilot study. *Reprod Biol Endocrinol* **15**: 59.

Kwon I, Kato M, Xiang S, Wu L, Theodoropoulos P, Mirzaei H, Han T, Xie S, Corden JL, McKnight SL. 2013. Phosphorylation-regulated binding of RNA polymerase II to fibrous polymers of low-complexity domains. *Cell* **155**: 1049-1060.

Ladurner AG. 2003. Inactivating chromosomes: a macro domain that minimizes transcription. *Mol Cell* **12**: 1-3.

Lagrange T, Kapanidis AN, Tang H, Reinberg D, Ebright RH. 1998. New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. *Genes Dev* **12**: 34-44.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357-359.

Lashgari A, Millau JF, Jacques PE, Gaudreau L. 2017. Global inhibition of transcription causes an increase in histone H2A.Z incorporation within gene bodies. *Nucleic Acids Res* **45**: 12715-12722.

Laskey RA, Honda BM, Mills AD, Finch JT. 1978. Nucleosomes are assembled by an acidic protein which binds histones and transfers them to DNA. *Nature* **275**: 416-420.

Lassar AB, Davis RL, Wright WE, Kadesch T, Murre C, Voronova A, Baltimore D, Weintraub H. 1991. Functional activity of myogenic HLH proteins requires hetero-oligomerization with E12/E47-like proteins in vivo. *Cell* **66**: 305-315.

Lauberth SM, Nakayama T, Wu X, Ferris AL, Tang Z, Hughes SH, Roeder RG. 2013. H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation. *Cell* **152**: 1021-1036.

Lawrence M, Daujat S, Schneider R. 2016. Lateral Thinking: How Histone Modifications Regulate Gene Expression. *Trends Genet* **32**: 42-56.

Le S, Davis C, Konopka JB, Sternglanz R. 1997. Two new S-phase-specific genes from Saccharomyces cerevisiae. *Yeast* **13**: 1029-1042.

Leal MF, Mazzotti TK, Calcagno DQ, Cirilo PD, Martinez MC, Demachki S, Assumpcao PP, Chammas R, Burbano RR, Smith MC. 2014. Deregulated expression of Nucleophosmin 1 in gastric cancer and its clinicopathological implications. *BMC Gastroenterol* **14**: 9.

Lee DF, Su J, Ang YS, Carvajal-Vergara X, Mulero-Navarro S, Pereira CF, Gingold J, Wang HL, Zhao R, Sevilla A et al. 2012a. Regulation of embryonic and induced pluripotency by aurora kinase-p53 signaling. *Cell Stem Cell* **11**: 179-194.

Lee DK, DeJong J, Hashimoto S, Horikoshi M, Roeder RG. 1992. TFIIA induces conformational changes in TFIID via interactions with the basic repeat. *Mol Cell Biol* **12**: 5189-5196.

Lee HH, Kim HS, Kang JY, Lee BI, Ha JY, Yoon HJ, Lim SO, Jung G, Suh SW. 2007a. Crystal structure of human nucleophosmin-core reveals plasticity of the pentamer-pentamer interface. *Proteins* **69**: 672-678.

Lee HZ, Wu CH, Chang SP. 2005a. Release of nucleophosmin from the nucleus: Involvement in aloe-emodin-induced human lung non small carcinoma cell apoptosis. *Int J Cancer* **113**: 971-976.

Lee JS, Shilatifard A. 2007. A site to remember: H3K36 methylation a mark for histone deacetylation. *Mutat Res* **618**: 130-134.

Lee LC, Chen CM, Chen FL, Lin PY, Hsiao YC, Wang PR, Su MT, Hsieh-Li HM, Hwang JC, Wu CH et al. 2009. Altered expression of HSPA5, HSPA8 and PARK7 in spinocerebellar ataxia type 17 identified by 2-dimensional fluorescence difference in gel electrophoresis. *Clin Chim Acta* **400**: 56-62.

Lee N, Kim DK, Kim ES, Park SJ, Kwon JH, Shin J, Park SM, Moon YH, Wang HJ, Gho YS et al. 2014. Comparative interactomes of SIRT6 and SIRT7: Implication of functional links to aging. *Proteomics* **14**: 1610-1622.

Lee SB, Kim CK, Lee KH, Ahn JY. 2012b. S-nitrosylation of B23/nucleophosmin by GAPDH protects cells from the SIAH1-GAPDH death cascade. *J Cell Biol* **199**: 65-76.

Lee SB, Xuan Nguyen TL, Choi JW, Lee KH, Cho SW, Liu Z, Ye K, Bae SS, Ahn JY. 2008. Nuclear Akt interacts with B23/NPM and protects it from proteolytic cleavage, enhancing cell survival. *Proc Natl Acad Sci U S A* **105**: 16584-16589.

Lee SY, Park JH, Kim S, Park EJ, Yun Y, Kwon J. 2005b. A proteomics approach for the identification of nucleophosmin and heterogeneous nuclear ribonucleoprotein C1/C2 as chromatin-binding proteins in response to DNA double-strand breaks. *Biochem J* **388**: 7-15.

Lee W, Tillo D, Bray N, Morse RH, Davis RW, Hughes TR, Nislow C. 2007b. A high-resolution atlas of nucleosome occupancy in yeast. *Nat Genet* **39**: 1235-1244.

Lemon B, Tjian R. 2000. Orchestrated response: a symphony of transcription factors for gene control. *Genes Dev* **14**: 2551-2569.

Leno GH, Mills AD, Philpott A, Laskey RA. 1996. Hyperphosphorylation of nucleoplasmin facilitates Xenopus sperm decondensation at fertilization. *J Biol Chem* **271**: 7253-7256.

Leotoing L, Meunier L, Manin M, Mauduit C, Decaussin M, Verrijdt G, Claessens F, Benahmed M, Veyssiere G, Morel L et al. 2008. Influence of nucleophosmin/B23 on DNA binding and transcriptional activity of the androgen receptor in prostate cancer cell. *Oncogene* **27**: 2858-2867.

Leroy B, Anderson M, Soussi T. 2014. TP53 mutations in human cancer: database reassessment and prospects for the next decade. *Hum Mutat* **35**: 672-688.

Lessard F, Morin F, Ivanchuk S, Langlois F, Stefanovsky V, Rutka J, Moss T. 2010. The ARF tumor suppressor controls ribosome biogenesis by regulating the RNA polymerase I transcription factor TTF-I. *Mol Cell* **38**: 539-550.

Leung A, Jardim FP, Savic N, Monneau YR, Gonzalez-Romero R, Gudavicius G, Eirin-Lopez JM, Bartke T, Mackereth CD, Ausio J et al. 2017. Basic surface features of nuclear FKBPs facilitate chromatin binding. *Sci Rep* **7**: 3795.

Levine AJ, Oren M. 2009. The first 30 years of p53: growing ever more complex. In *Nat Rev Cancer*, Vol 9, pp. 749-758, England.

Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**: 323.

Li D, Lin B, Yusuf N, Burns EM, Yu X, Luo D, Min W. 2017a. Proteomic Analysis and Functional Studies of Baicalin on Proteins Associated with Skin Cancer. *Am J Chin Med* **45**: 599-614.

Li G, Reinberg D. 2011. Chromatin higher-order structures and gene regulation. *Curr Opin Genet Dev* **21**: 175-186.

Li J, Zhang X, Sejas DP, Bagby GC, Pang Q. 2004. Hypoxia-induced nucleophosmin protects cell death through inhibition of p53. *J Biol Chem* **279**: 41275-41279.

Li J, Zhang X, Sejas DP, Pang Q. 2005. Negative regulation of p53 by nucleophosmin antagonizes stress-induced apoptosis in human normal and malignant hematopoietic cells. *Leuk Res* **29**: 1415-1423.

Li P, Ying J, Chang Q, Zhu W, Yang G, Xu T, Yi H, Pan R, Zhang E, Zeng X et al. 2016. Effects of phycoerythrin from Gracilaria lemaneiformis in proliferation and apoptosis of SW480 cells. *Oncol Rep* **36**: 3536-3544.

Li QF, Tang J, Liu QR, Shi SL, Chen XF. 2010. Localization and altered expression of nucleophosmin in the nuclear matrix during the differentiation of human hepatocarcinoma SMMC-7721 cells induced by HMBA. *Cancer Invest* **28**: 1004-1012.

Li S, Zhang X, Zhou Z, Huang Z, Liu L. 2017b. Downregulation of nucleophosmin expression inhibited proliferation and induced apoptosis in salivary gland adenoid cystic carcinoma. *J Oral Pathol Med* **46**: 175-181.

Li Y, Zhang J. 2015. Expression of mutant p53 in oral squamous cell carcinoma is correlated with the effectiveness of intra-arterial chemotherapy. *Oncol Lett* **10**: 2883-2887.

Li YP. 1997. Protein B23 is an important human factor for the nucleolar localization of the human immunodeficiency virus protein Tat. *J Virol* **71**: 4098-4102.

Li Z, Boone D, Hann SR. 2008. Nucleophosmin interacts directly with c-Myc and controls c-Myc-induced hyperproliferation and transformation. *Proc Natl Acad Sci U S A* **105**: 18794-18799.

Li Z, Hann SR. 2009. The Myc-nucleophosmin-ARF network: a complex web unveiled. *Cell Cycle* **8**: 2703-2707.

Liang L, Soyal SM, Dean J. 1997. FIGalpha, a germ cell specific transcription factor involved in the coordinate expression of the zona pellucida genes. *Development* **124**: 4939-4947.

Liao M, Liao W, Xu N, Li B, Liu F, Zhang S, Wang Y, Wang S, Zhu Y, Chen D et al. 2019. LncRNA EPB41L4A-AS1 regulates glycolysis and glutaminolysis by mediating nucleolar translocation of HDAC2. *EBioMedicine* **41**: 200-213.

Licatalosi DD, Geiger G, Minet M, Schroeder S, Cilli K, McNeil JB, Bentley DL. 2002. Functional interaction of yeast pre-mRNA 3' end processing factors with RNA polymerase II. *Mol Cell* **9**: 1101-1111.

Lifantseva N, Koltsova A, Krylova T, Yakovleva T, Poljanskaya G, Gordeeva O. 2011. Expression patterns of cancer-testis antigens in human embryonic stem cells and their cell derivatives indicate lineage tracks. *Stem Cells Int* **2011**: 795239.

Lin CY, Chao A, Wang TH, Lee LY, Yang LY, Tsai CL, Wang HS, Lai CH. 2016. Nucleophosmin/B23 is a negative regulator of estrogen receptor alpha expression via AP2gamma in endometrial cancer cells. *Oncotarget* **7**: 60038-60052.

Lin CY, Lee LY, Wang TH, Hsu CL, Tsai CL, Chao A, Lai CH. 2019a. Palbociclib Promotes Dephosphorylation of NPM/B23 at Threonine 199 and Inhibits Endometrial Cancer Cell Growth. *Cancers (Basel)* **11**.

Lin CY, Liang YC, Yung BY. 2006. Nucleophosmin/B23 regulates transcriptional activation of E2F1 via modulating the promoter binding of NF-kappaB, E2F1 and pRB. *Cell Signal* **18**: 2041-2048.

Lin CY, Tan BC, Liu H, Shih CJ, Chien KY, Lin CL, Yung BY. 2010. Dephosphorylation of nucleophosmin by PP1beta facilitates pRB binding and consequent E2F1-dependent DNA repair. *Mol Biol Cell* **21**: 4409-4417.

Lin J, Kato M, Nagata K, Okuwaki M. 2017. Efficient DNA binding of NF-kappaB requires the chaperone-like function of NPM1. *Nucleic Acids Res* **45**: 3707-3723.

Lin MH, Zheng ZH, Jiang PF, Wu ZJ, Gan DH, Zhang N, Hu JD. 2019b. [Inhibitory Effect of NPM Gene Knockdown on Proliferation of Chronic Myeloid Leukemia Cell Line K562 and Its Mechanism]. *Zhongguo Shi Yan Xue Ye Xue Za Zhi* **27**: 1008-1012.

Lin YC, Choi WS, Gralla JD. 2005. TFIIH XPB mutants suggest a unified bacterial-like mechanism for promoter opening but not escape. *Nat Struct Mol Biol* **12**: 603-607.

Lin YC, Gralla JD. 2005. Stimulation of the XPB ATP-dependent helicase by the beta subunit of TFIIE. *Nucleic Acids Res* **33**: 3072-3081.

Lindstrom MS, Zhang Y. 2008. Ribosomal protein S9 is a novel B23/NPM-binding protein required for normal cell proliferation. *J Biol Chem* **283**: 15568-15576.

Lingenfelter BM, Tripurani SK, Tejomurtula J, Smith GW, Yao J. 2011. Molecular cloning and expression of bovine nucleoplasmin 2 (NPM2): a maternal effect gene regulated by miR-181a. *Reprod Biol Endocrinol* **9**: 40.

Litvin J, King ML. 1988. Expression and segregation of nucleoplasmin during development in Xenopus. *Development* **102**: 9-21.

Liu CD, Chen YL, Min YL, Zhao B, Cheng CP, Kang MS, Chiu SJ, Kieff E, Peng CW. 2012a. The nuclear chaperone nucleophosmin escorts an Epstein-Barr Virus nuclear antigen to establish transcriptional cascades for latent infection in human B cells. *PLoS Pathog* **8**: e1003084.

Liu H, Tan BC, Tseng KH, Chuang CP, Yeh CW, Chen KD, Lee SC, Yung BY. 2007a. Nucleophosmin acts as a novel AP2alpha-binding transcriptional corepressor during cell differentiation. *EMBO Rep* **8**: 394-400.

Liu JC, Baker RE, Chow W, Sun CK, Elsholtz HP. 2005. Epigenetic mechanisms in the dopamine D2 receptor-dependent inhibition of the prolactin gene. *Mol Endocrinol* **19**: 1904-1917.

Liu K, Ling S, Lin WC. 2011. TopBP1 mediates mutant p53 gain of function through NF-Y and p63/p73. *Mol Cell Biol* **31**: 4464-4481.

Liu M, Varela-Ramirez A, Li J, Dai L, Aguilera RJ, Zhang JY. 2015a. Humoral autoimmune response to nucleophosmin in the immunodiagnosis of hepatocellular carcinoma. *Oncol Rep* **33**: 2245-2252.

Liu Q, Liu N, Shangguan Q, Zhang F, Chai W, Tong X, Zhao X, Li Z, Qi D, Ye X. 2019. LncRNA SAMD12-AS1 promotes cell proliferation and inhibits apoptosis by interacting with NPM1. *Sci Rep* **9**: 11593.

Liu WH, Yung BY. 1998. Mortalization of human promyelocytic leukemia HL-60 cells to be more susceptible to sodium butyrate-induced apoptosis and inhibition of telomerase activity by down-regulation of nucleophosmin/B23. *Oncogene* **17**: 3055-3064.

Liu X, Bushnell DA, Wang D, Calero G, Kornberg RD. 2010. Structure of an RNA polymerase II-TFIIB complex and the transcription initiation mechanism. *Science* **327**: 206-209.

Liu X, Kraus WL, Bai X. 2015b. Ready, pause, go: regulation of RNA polymerase II pausing and release by cellular signaling pathways. *Trends Biochem Sci* **40**: 516-525.

Liu X, Liu D, Qian D, Dai J, An Y, Jiang S, Stanley B, Yang J, Wang B, Liu DX. 2012b. Nucleophosmin (NPM1/B23) interacts with activating transcription factor 5 (ATF5) protein and promotes proteasome- and caspase-dependent ATF5 degradation in hepatocellular carcinoma cells. *J Biol Chem* **287**: 19599-19609.

Liu X, Liu Z, Jang SW, Ma Z, Shinmura K, Kang S, Dong S, Chen J, Fukasawa K, Ye K. 2007b. Sumoylation of nucleophosmin/B23 regulates its subcellular localization, mediating cell proliferation and survival. *Proc Natl Acad Sci U S A* **104**: 9679-9684.

Liu Y, Zhang F, Zhang XF, Qi LS, Yang L, Guo H, Zhang N. 2012c. Expression of nucleophosmin/NPM1 correlates with migration and invasiveness of colon cancer cells. *J Biomed Sci* **19**: 53.

Llanos S, Clark PA, Rowe J, Peters G. 2001. Stabilization of p53 by p14ARF without relocation of MDM2 to the nucleolus. *Nat Cell Biol* **3**: 445-452.

Londero AP, Orsaria M, Tell G, Marzinotto S, Capodicasa V, Poletto M, Vascotto C, Sacco C, Mariuzzi L. 2014. Expression and prognostic significance of APE1/Ref-1 and NPM1 proteins in high-grade ovarian serous cancer. *Am J Clin Pathol* **141**: 404-414.

Long M, Sun X, Shi W, Yanru A, Leung STC, Ding D, Cheema MS, MacPherson N, Nelson CJ, Ausio J et al. 2019. A novel histone H4 variant H4G regulates rDNA transcription in breast cancer. *Nucleic Acids Res* **47**: 8399-8409.

Loppin B, Bonnefoy E, Anselme C, Laurencon A, Karr TL, Couble P. 2005. The histone H3.3 chaperone HIRA is essential for chromatin assembly in the male pronucleus. *Nature* **437**: 1386-1390.

Lorain S, Demczuk S, Lamour V, Toth S, Aurias A, Roe BA, Lipinski M. 1996. Structural Organization of the WD repeat protein-encoding gene HIRA in the DiGeorge syndrome critical region of human chromosome 22. *Genome Res* **6**: 43-50.

Loubeau G, Boudra R, Maquaire S, Lours-Calet C, Beaudoin C, Verrelle P, Morel L. 2014. NPM1 silencing reduces tumour growth and MAPK signalling in prostate cancer cells. *PLoS One* **9**: e96293.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550.

Lowndes NF, Toh GW. 2005. DNA repair: the importance of phosphorylating histone H2AX. *Curr Biol* **15**: R99-r102.

Loya TJ, Reines D. 2016. Recent advances in understanding transcription termination by RNA polymerase II. *F1000Res* **5**.

Lu H, Yu D, Hansen AS, Ganguly S, Liu R, Heckert A, Darzacq X, Zhou Q. 2018. Phase-separation mechanism for C-terminal hyperphosphorylation of RNA polymerase II. *Nature* **558**: 318-323.

Lu Z, Zhang C, Zhai Z. 2005. Nucleoplasmin regulates chromatin condensation during apoptosis. *Proc Natl Acad Sci U S A* **102**: 2778-2783.

Luchinat E, Chiarella S, Franceschini M, Di Matteo A, Brunori M, Banci L, Federici L. 2018. Identification of a novel nucleophosmin-interaction motif in the tumor suppressor p14arf. *Febs j* **285**: 832-847.

Luebben WR, Sharma N, Nyborg JK. 2010. Nucleosome eviction and activated transcription require p300 acetylation of histone H3 lysine 14. *Proc Natl Acad Sci U S A* **107**: 19254-19259.

Luger K. 2006. Dynamic nucleosomes. *Chromosome Res* **14**: 5-16.

Luger K, Dechassa ML, Tremethick DJ. 2012. New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? *Nat Rev Mol Cell Biol* **13**: 436-447.

Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**: 251-260.

Luger K, Rechsteiner TJ, Richmond TJ. 1999. Expression and purification of recombinant histones and nucleosome reconstitution. *Methods Mol Biol* **119**: 1-16.

Lukashchuk N, Vousden KH. 2007. Ubiquitination and degradation of mutant p53. *Mol Cell Biol* **27**: 8284-8295.

Luo H, Sun Y, Wei G, Luo J, Yang X, Liu W, Guo M, Chen R. 2015. Functional Characterization of Long Noncoding RNA Lnc_bc060912 in Human Lung Carcinoma Cells. *Biochemistry* **54**: 2895-2902.

Luo Z, Lin C, Shilatifard A. 2012. The super elongation complex (SEC) family in transcriptional control. *Nat Rev Mol Cell Biol* **13**: 543-547.

Luse DS. 2013. Promoter clearance by RNA polymerase II. *Biochim Biophys Acta* **1829**: 63-68.

Ma HL, Jin SF, Ju WT, Fu Y, Tu YY, Wang LZ, Jiang L, Zhang ZY, Zhong LP. 2017. Stathmin is overexpressed and regulated by mutant p53 in oral squamous cell carcinoma. *J Exp Clin Cancer Res* **36**: 109.

MacArthur CA, Shackleford GM. 1997. Npm3: a novel, widely expressed gene encoding a protein related to the molecular chaperones nucleoplasmin and nucleophosmin. *Genomics* **42**: 137-140.

Maggi LB, Jr., Kuchenruether M, Dadey DY, Schwope RM, Grisendi S, Townsend RR, Pandolfi PP, Weber JD. 2008. Nucleophosmin serves as a rate-limiting nuclear export chaperone for the Mammalian ribosome. *Mol Cell Biol* **28**: 7050-7065.

Mai RT, Yeh TS, Kao CF, Sun SK, Huang HH, Wu Lee YH. 2006. Hepatitis C virus core protein recruits nucleolar phosphoprotein B23 and coactivator p300 to relieve the repression effect of transcriptional factor YY1 on B23 gene expression. *Oncogene* **25**: 448-462.

Maiguel DA, Jones L, Chakravarty D, Yang C, Carrier F. 2004. Nucleophosmin sets a threshold for p53 response to UV radiation. *Mol Cell Biol* **24**: 3703-3711.

Maji S, Dahiya P, Waseem M, Dwivedi N, Bhat DS, Dar TH, Thakur JK. 2019. Interaction map of Arabidopsis Mediator complex expounding its topology. *Nucleic Acids Res* **47**: 3904-3920.

Malfatti MC, Gerratana L, Dalla E, Isola M, Damante G, Di Loreto C, Puglisi F, Tell G. 2019. APE1 and NPM1 protect cancer cells from platinum compounds cytotoxicity and their expression pattern has a prognostic value in TNBC. *J Exp Clin Cancer Res* **38**: 309.

Malik S, Roeder RG. 2005. Dynamic regulation of pol II transcription by the mammalian Mediator complex. *Trends Biochem Sci* **30**: 256-263.

Malik S, Roeder RG. 2010. The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. *Nat Rev Genet* **11**: 761-772.

Malik S, Roeder RG. 2013. Biochemistry. Have your PIC! *Science* **342**: 706-707.

Mandal SS, Chu C, Wada T, Handa H, Shatkin AJ, Reinberg D. 2004. Functional interactions of RNA-capping enzyme with factors that positively and negatively regulate promoter escape by RNA polymerase II. *Proc Natl Acad Sci U S A* **101**: 7572-7577.

Mao Z, Pan L, Wang W, Sun J, Shan S, Dong Q, Liang X, Dai L, Ding X, Chen S et al. 2014. Anp32e, a higher eukaryotic histone chaperone directs preferential recognition for H2A.Z. *Cell Res* **24**: 389-399.

Mapendano CK, Lykke-Andersen S, Kjems J, Bertrand E, Jensen TH. 2010. Crosstalk between mRNA 3' end processing and transcription initiation. *Mol Cell* **40**: 410-422.

Margaritis T, Holstege FC. 2008. Poised RNA polymerase II gives pause for thought. *Cell* **133**: 581-584.

Martianov I, Brancorsini S, Catena R, Gansmuller A, Kotaja N, Parvinen M, Sassone-Corsi P, Davidson I. 2005. Polar nuclear localization of H1T2, a histone H1 variant, required for spermatid elongation and DNA condensation during spermiogenesis. *Proc Natl Acad Sci U S A* **102**: 2808-2813.

Martinez E, Kundu TK, Fu J, Roeder RG. 1998. A human SPT3-TAFII31-GCN5-L acetylase complex distinct from transcription factor IID. *J Biol Chem* **273**: 23781-23785.

Martinez E, Zhou Q, L'Etoile ND, Oelgeschlager T, Berk AJ, Roeder RG. 1995. Core promoter-specific function of a mutant transcription factor TFIID defective in TATA-box binding. *Proc Natl Acad Sci U S A* **92**: 11864-11868.

Maruyama K, Todaka D, Mizoi J, Yoshida T, Kidokoro S, Matsukura S, Takasaki H, Sakurai T, Yamamoto YY, Yoshiwara K et al. 2012. Identification of cis-acting promoter elements in cold- and dehydration-induced transcriptional pathways in Arabidopsis, rice, and soybean. *DNA Res* **19**: 37-49.

Marzluff WF. 2005. Metazoan replication-dependent histone mRNAs: a distinct set of RNA polymerase II transcripts. *Curr Opin Cell Biol* **17**: 274-280.

Mason PB, Struhl K. 2003. The FACT complex travels with elongating RNA polymerase II and is important for the fidelity of transcriptional initiation in vivo. *Mol Cell Biol* **23**: 8323-8333.

Matangkasombut O, Buratowski S. 2003. Different sensitivities of bromodomain factors 1 and 2 to histone H4 acetylation. *Mol Cell* **11**: 353-363.

Mathew R, Chatterji D. 2006. The evolving story of the omega subunit of bacterial RNA polymerase. *Trends Microbiol* **14**: 450-455.

Mavrich TN, Jiang C, Ioshikhes IP, Li X, Venters BJ, Zanton SJ, Tomsho LP, Qi J, Glaser RL, Schuster SC et al. 2008. Nucleosome organization in the Drosophila genome. *Nature* **453**: 358-362.

McKittrick E, Gafken PR, Ahmad K, Henikoff S. 2004. Histone H3.3 is enriched in covalent modifications associated with active chromatin. *Proc Natl Acad Sci U S A* **101**: 1525-1530.

Meder VS, Boeglin M, de Murcia G, Schreiber V. 2005. PARP-1 and PARP-2 interact with nucleophosmin/B23 and accumulate in transcriptionally active nucleoli. *J Cell Sci* **118**: 211-222.

Meller VH, Joshi SS, Deshpande N. 2015. Modulation of Chromatin by Noncoding RNA. *Annu Rev Genet* **49**: 673-695.

Mendes A, Fahrenkrog B. 2019. NUP214 in Leukemia: It's More than Transport. *Cells* **8**.

Michaelson JS, Bader D, Kuo F, Kozak C, Leder P. 1999. Loss of Daxx, a promiscuously interacting protein, results in extensive apoptosis in early mouse development. *Genes Dev* **13**: 1918-1923.

Michalik J, Yeoman LC, Busch H. 1981. Nucleolar localization of protein B23 (37/5.1) by immunocytochemical techniques. *Life Sci* **28**: 1371-1379.

Michaud GA, Snyder M. 2002. Proteomic approaches for the global analysis of proteins. *Biotechniques* **33**: 1308-1316.

Minakhin L, Bhagat S, Brunning A, Campbell EA, Darst SA, Ebright RH, Severinov K. 2001. Bacterial RNA polymerase subunit omega and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. *Proc Natl Acad Sci U S A* **98**: 892-897.

Mitrea DM, Cika JA, Stanley CB, Nourse A, Onuchic PL, Banerjee PR, Phillips AH, Park CG, Deniz AA, Kriwacki RW. 2018. Self-interaction of NPM1 modulates multiple mechanisms of liquid-liquid phase separation. *Nat Commun* **9**: 842.

Mitrea DM, Grace CR, Buljan M, Yun MK, Pytel NJ, Satumba J, Nourse A, Park CG, Madan Babu M, White SW et al. 2014. Structural polymorphism in the N-terminal oligomerization domain of NPM1. *Proc Natl Acad Sci U S A* **111**: 4466-4471.

Mitrea DM, Kriwacki RW. 2018. On the relationship status for Arf and NPM1 - it's complicated. *Febs j* **285**: 828-831.

Miwa K, Kojima R, Obita T, Ohkuma Y, Tamura Y, Mizuguchi M. 2016. Crystal Structure of Human General Transcription Factor TFIIE at Atomic Resolution. *J Mol Biol* **428**: 4258-4266.

Mizuguchi G, Shen X, Landry J, Wu WH, Sen S, Wu C. 2004. ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. *Science* **303**: 343-348.

Mladinov M, Sedmak G, Fuller HR, Babic Leko M, Mayer D, Kirincich J, Stajduhar A, Borovecki F, Hof PR, Simic G. 2016. Gene expression profiling of the dorsolateral and medial orbitofrontal cortex in schizophrenia. *Transl Neurosci* **7**: 139-150.

Modak R, Basha J, Bharathy N, Maity K, Mizar P, Bhat AV, Vasudevan M, Rao VK, Kok WK, Natesh N et al. 2013. Probing p300/CBP associated factor (PCAF)-dependent pathways with a small molecule inhibitor. *ACS Chem Biol* **8**: 1311-1323.

Montarras D, Chelly J, Bober E, Arnold H, Ott MO, Gros F, Pinset C. 1991. Developmental patterns in the expression of Myf5, MyoD, myogenin, and MRF4 during myogenesis. *New Biol* **3**: 592-600.

Moreau N, Angelier N, Bonnanfant-Jais ML, Gounon P, Kubisz P. 1986. Association of nucleoplasmin with transcription products as revealed by immunolocalization in the amphibian oocyte. *J Cell Biol* **103**: 683-690.

Morris SW, Kirstein MN, Valentine MB, Dittmer KG, Shapiro DN, Saltman DL, Look AT. 1994. Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin's lymphoma. *Science* **263**: 1281-1284.

Morrison AJ, Shen X. 2005. DNA repair in the context of chromatin. *Cell Cycle* **4**: 568-571.

Moshkin YM, Kan TW, Goodfellow H, Bezstarosti K, Maeda RK, Pilyugin M, Karch F, Bray SJ, Demmers JA, Verrijzer CP. 2009. Histone chaperones ASF1 and NAP1 differentially modulate removal of active histone marks by LID-RPD3 complexes during NOTCH silencing. *Mol Cell* **35**: 782-793.

Motoi N, Suzuki K, Hirota R, Johnson P, Oofusa K, Kikuchi Y, Yoshizato K. 2008. Identification and characterization of nucleoplasmin 3 as a histone-binding protein in embryonic stem cells. *Dev Growth Differ* **50**: 307-320.

Muhlbacher W, Sainsbury S, Hemann M, Hantsche M, Neyer S, Herzog F, Cramer P. 2014. Conserved architecture of the core RNA polymerase II initiation complex. *Nat Commun* **5**: 4310.

Mukhopadhyay A, Sehgal L, Bose A, Gulvady A, Senapati P, Thorat R, Basu S, Bhatt K, Hosing AS, Balyan R et al. 2016. 14-3-3gamma Prevents Centrosome Amplification and Neoplastic Progression. *Sci Rep* **6**: 26580.

Muller F, Tora L. 2014. Chromatin and DNA sequences in defining promoters for transcription initiation. *Biochim Biophys Acta* **1839**: 118-128.

Muller GA, Engeland K. 2010. The central role of CDE/CHR promoter elements in the regulation of cell cycle-dependent gene transcription. *Febs j* **277**: 877-893.

Murakami K, Elmlund H, Kalisman N, Bushnell DA, Adams CM, Azubel M, Elmlund D, Levi-Kalisman Y, Liu X, Gibbons BJ et al. 2013. Architecture of an RNA polymerase II transcription pre-initiation complex. *Science* **342**: 1238724.

Murakami K, Tsai KL, Kalisman N, Bushnell DA, Asturias FJ, Kornberg RD. 2015. Structure of an RNA polymerase II preinitiation complex. *Proc Natl Acad Sci U S A* **112**: 13543-13548.

Murano K, Okuwaki M, Hisaoka M, Nagata K. 2008. Transcription regulation of the rRNA gene by a multifunctional nucleolar protein, B23/nucleophosmin, through its histone chaperone activity. *Mol Cell Biol* **28**: 3114-3126.

Muse GW, Gilchrist DA, Nechaev S, Shah R, Parker JS, Grissom SF, Zeitlinger J, Adelman K. 2007. RNA polymerase is poised for activation across the genome. *Nat Genet* **39**: 1507-1511.

Naar AM, Lemon BD, Tjian R. 2001. Transcriptional coactivator complexes. *Annu Rev Biochem* **70**: 475-501.

Naar AM, Taatjes DJ, Zhai W, Nogales E, Tjian R. 2002. Human CRSP interacts with RNA polymerase II CTD and adopts a specific CTD-bound conformation. *Genes Dev* **16**: 1339-1344.

Nachmani D, Bothmer AH, Grisendi S, Mele A, Bothmer D, Lee JD, Monteleone E, Cheng K, Zhang Y, Bester AC et al. 2019. Germline NPM1 mutations lead to altered rRNA 2'-O-methylation and cause dyskeratosis congenita. *Nat Genet*.

Namboodiri VM, Akey IV, Schmidt-Zachmann MS, Head JF, Akey CW. 2004a. The structure and function of Xenopus NO38-core, a histone chaperone in the nucleolus. *Structure* **12**: 2149-2160.

Namboodiri VM, Dutta S, Akey IV, Head JF, Akey CW. 2003. The crystal structure of Drosophila NLP-core provides insight into pentamer formation and histone binding. *Structure* **11**: 175-186.

Namboodiri VM, Schmidt-Zachmann MS, Head JF, Akey CW. 2004b. Purification, crystallization and preliminary X-ray analysis of the N-terminal domain of NO38, a nucleolar protein from Xenopus laevis. *Acta Crystallogr D Biol Crystallogr* **60**: 2325-2327.

Naoe T, Suzuki T, Kiyoi H, Urano T. 2006. Nucleophosmin: a versatile molecule associated with hematological malignancies. *Cancer Sci* **97**: 963-969.

Nateri AS, Spencer-Dene B, Behrens A. 2005. Interaction of phosphorylated c-Jun with TCF4 regulates intestinal cancer development. *Nature* **437**: 281-285.

Natesan S, Gilman M. 1995. YY1 facilitates the association of serum response factor with the c-fos serum response element. *Mol Cell Biol* **15**: 5975-5982.

Natesan S, Gilman MZ. 1993. DNA bending and orientation-dependent function of YY1 in the c-fos promoter. *Genes Dev* **7**: 2497-2509.

Nechaev S, Adelman K. 2008. Promoter-proximal Pol II: when stalling speeds things up. *Cell Cycle* **7**: 1539-1544.

Nechemia-Arbely Y, Fachinetti D, Miga KH, Sekulic N, Soni GV, Kim DH, Wong AK, Lee AY, Nguyen K, Dekker C et al. 2017. Human centromeric CENP-A chromatin is a homotypic, octameric nucleosome at all cell cycle points. *J Cell Biol* **216**: 607-621.

Negi SS, Olson MO. 2006. Effects of interphase and mitotic phosphorylation on the mobility and location of nucleolar protein B23. *J Cell Sci* **119**: 3676-3685.

Nelson CJ, Santos-Rosa H, Kouzarides T. 2006. Proline isomerization of histone H3 regulates lysine methylation and gene expression. *Cell* **126**: 905-916.

Neo SH, Itahana Y, Alagu J, Kitagawa M, Guo AK, Lee SH, Tang K, Itahana K. 2015. TRIM28 Is an E3 Ligase for ARF-Mediated NPM1/B23 SUMOylation That Represses Centrosome Amplification. *Mol Cell Biol* **35**: 2851-2863.

Neuberg M, Schuermann M, Hunter JB, Muller R. 1989. Two functionally different regions in Fos are required for the sequence-specific DNA interaction of the Fos/Jun protein complex. *Nature* **338**: 589-590.

Nguyen AL, Schindler K. 2017. Specialize and Divide (Twice): Functions of Three Aurora Kinase Homologs in Mammalian Oocyte Meiotic Maturation. *Trends Genet* **33**: 349-363.

Nikolov DB, Chen H, Halay ED, Hoffman A, Roeder RG, Burley SK. 1996. Crystal structure of a human TATA box-binding protein/TATA element complex. *Proc Natl Acad Sci U S A* **93**: 4862-4867.

Nikolov DB, Chen H, Halay ED, Usheva AA, Hisatake K, Lee DK, Roeder RG, Burley SK. 1995. Crystal structure of a TFIIB-TBP-TATA-element ternary complex. *Nature* **377**: 119-128.

Nishida T, Yamada Y. 2008. SMT3IP1, a nucleolar SUMO-specific protease, deconjugates SUMO-2 from nucleolar and cytoplasmic nucleophosmin. *Biochem Biophys Res Commun* **374**: 382-387.

Nishimura Y, Ohkubo T, Furuichi Y, Umekawa H. 2002. Tryptophans 286 and 288 in the C-terminal region of protein B23.1 are important for its nucleolar localization. *Biosci Biotechnol Biochem* **66**: 2239-2242.

Nishio J, Iwasaki H, Nabeshima K, Naito M. 2011. Cytogenetics and molecular genetics of myxoid soft-tissue sarcomas. *Genet Res Int* **2011**: 497148.

Noll M, Kornberg RD. 1977. Action of micrococcal nuclease on chromatin and the location of histone H1. *J Mol Biol* **109**: 393-404.

Nozawa Y, Van Belzen N, Van der Made AC, Dinjens WN, Bosman FT. 1996. Expression of nucleophosmin/B23 in normal and neoplastic colorectal mucosa. *J Pathol* **178**: 48-52.

Ogushi S, Yamagata K, Obuse C, Furuta K, Wakayama T, Matzuk MM, Saitou M. 2017. Reconstitution of the oocyte nucleolus in mice through a single nucleolar protein, NPM2. *J Cell Sci* **130**: 2416-2429.

Ohkuma Y, Roeder RG. 1994. Regulation of TFIIH ATPase and kinase activities by TFIIE during active initiation complex formation. *Nature* **368**: 160-163.

Okada M, Jang SW, Ye K. 2007. Ebp1 association with nucleophosmin/B23 is essential for regulating cell proliferation and suppressing apoptosis. *J Biol Chem* **282**: 36744-36754.

Okuda M, Horn HF, Tarapore P, Tokuyama Y, Smulian AG, Chan PK, Knudsen ES, Hofmann IA, Snyder JD, Bove KE et al. 2000. Nucleophosmin/B23 is a target of CDK2/cyclin E in centrosome duplication. *Cell* **103**: 127-140.

Okuwaki M. 2008. The structure and functions of NPM1/Nucleophsmin/B23, a multifunctional nucleolar acidic protein. *J Biochem* **143**: 441-448.

Okuwaki M, Iwamatsu A, Tsujimoto M, Nagata K. 2001a. Identification of nucleophosmin/B23, an acidic nucleolar protein, as a stimulatory factor for in vitro replication of adenovirus DNA complexed with viral basic core proteins. *J Mol Biol* **311**: 41-55.

Okuwaki M, Matsumoto K, Tsujimoto M, Nagata K. 2001b. Function of nucleophosmin/B23, a nucleolar acidic protein, as a histone chaperone. *FEBS Lett* **506**: 272-276.

Okuwaki M, Sumi A, Hisaoka M, Saotome-Nakamura A, Akashi S, Nishimura Y, Nagata K. 2012. Function of homo- and hetero-oligomers of human nucleoplasmin/nucleophosmin family proteins NPM1, NPM2 and NPM3 during sperm chromatin remodeling. *Nucleic Acids Res* **40**: 4861-4878.

Okuwaki M, Tsujimoto M, Nagata K. 2002. The RNA binding activity of a ribosome biogenesis factor, nucleophosmin/B23, is modulated by phosphorylation with a cell cycle-dependent kinase and by association with its subtype. *Mol Biol Cell* **13**: 2016-2030.

Oleaga C, Welten S, Belloc A, Sole A, Rodriguez L, Mencia N, Selga E, Tapias A, Noe V, Ciudad CJ. 2012. Identification of novel Sp1 targets involved in proliferation and cancer by functional genomics. *Biochem Pharmacol* **84**: 1581-1591.

Olins AL, Olins DE. 1974. Spheroid chromatin units (v bodies). *Science* **183**: 330-332.

Olins AL, Olins DE. 1979. Stereo electron microscopy of the 25-nm chromatin fibers in isolated nuclei. *J Cell Biol* **81**: 260-265.

Olins DE, Olins AL. 2003. Chromatin history: our view from the bridge. *Nat Rev Mol Cell Biol* **4**: 809-814.

Olson MO, Prestayko AW, Jones CF, Busch H. 1974. Phosphorylation of proteins of ribosomes and nucleolar preribosomal particles from Novikoff hepatmoa ascites cells. *J Mol Biol* **90**: 161-168.

Olson MO, Wallace MO, Herrera AH, Marshall-Carlson L, Hunt RC. 1986. Preribosomal ribonucleoprotein particles are a major component of a nucleolar matrix fraction. *Biochemistry* **25**: 484-491.

Onikubo T, Nicklay JJ, Xing L, Warren C, Anson B, Wang WL, Burgos ES, Ruff SE, Shabanowitz J, Cheng RH et al. 2015. Developmentally Regulated Post-translational Modification of Nucleoplasmin Controls Histone Sequestration and Deposition. *Cell Rep*.

Ooi L, Wood IC. 2007. Chromatin crosstalk in development and disease: lessons from REST. *Nat Rev Genet* **8**: 544-554.

Orlandini M, Marconcini L, Ferruzzi R, Oliviero S. 1996. Identification of a c-fos-induced gene that is related to the platelet-derived growth factor/vascular endothelial growth factor family. *Proc Natl Acad Sci U S A* **93**: 11675-11680.

Orphanides G, Lagrange T, Reinberg D. 1996. The general transcription factors of RNA polymerase II. *Genes Dev* **10**: 2657-2683.

Orphanides G, LeRoy G, Chang CH, Luse DS, Reinberg D. 1998. FACT, a factor that facilitates transcript elongation through nucleosomes. *Cell* **92**: 105-116.

Orphanides G, Wu WH, Lane WS, Hampsey M, Reinberg D. 1999. The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSRP1 proteins. *Nature* **400**: 284-288.

Orrick LR, Olson MO, Busch H. 1973. Comparison of nucleolar proteins of normal rat liver and Novikoff hepatoma ascites cells by two-dimensional polyacrylamide gel electrophoresis. *Proc Natl Acad Sci U S A* **70**: 1316-1320.

Osborn JL, Greer SF. 2015. A Role for Histone Chaperones in Regulating RNA Polymerase II. *Advances in Biological Chemistry* **Vol.05No.02**: 10.

Oudet P, Gross-Bellard M, Chambon P. 1975. Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. *Cell* **4**: 281-300.

Ozsolak F, Song JS, Liu XS, Fisher DE. 2007. High-throughput mapping of the chromatin structure of human promoters. *Nat Biotechnol* **25**: 244-248.

Pal M, McKean D, Luse DS. 2001. Promoter clearance by RNA polymerase II is an extended, multistep process strongly affected by sequence. *Mol Cell Biol* **21**: 5815-5825.

Pal M, Ponticelli AS, Luse DS. 2005. The role of the transcription bubble and TFIIB in promoter clearance by RNA polymerase II. *Mol Cell* **19**: 101-110.

Pandit B, Gartel AL. 2015. Mutual Regulation of FOXM1, NPM and ARF Proteins. *J Cancer* **6**: 538-541.

Pang Q, Christianson TA, Koretsky T, Carlson H, David L, Keeble W, Faulkner GR, Speckhart A, Bagby GC. 2003. Nucleophosmin interacts with and inhibits the catalytic function of eukaryotic initiation factor 2 kinase PKR. *J Biol Chem* **278**: 41709-41717.

Pangas SA, Choi Y, Ballow DJ, Zhao Y, Westphal H, Matzuk MM, Rajkovic A. 2006. Oogenesis requires germ cell-specific transcriptional regulators Sohlh1 and Lhx8. *Proc Natl Acad Sci U S A* **103**: 8090-8095.

Parelho V, Hadjur S, Spivakov M, Leleu M, Sauer S, Gregson HC, Jarmuz A, Canzonetta C, Webster Z, Nesterova T et al. 2008. Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* **132**: 422-433.

Park G, Gong Z, Chen J, Kim JE. 2010. Characterization of the DOT1L network: implications of diverse roles for DOT1L. *Protein J* **29**: 213-223.

Park JM, Adam RM, Peters CA, Guthrie PD, Sun Z, Klagsbrun M, Freeman MR. 1999. AP-1 mediates stretch-induced expression of HB-EGF in bladder smooth muscle cells. *Am J Physiol* **277**: C294-301.

Passos-Castilho AM, Marchand C, Archambault D. 2018. B23/nucleophosmin interacts with bovine immunodeficiency virus Rev protein and facilitates viral replication. *Virology* **515**: 158-164.

Patel KR, Vajaria BN, Singh RD, Begum R, Patel PS. 2018. Clinical implications of p53 alterations in oral cancer progression: a review from India. *Exp Oncol* **40**: 10-18.

Pavri R, Zhu B, Li G, Trojer P, Mandal S, Shilatifard A, Reinberg D. 2006. Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II. *Cell* **125**: 703-717.

Peltonen JK, Helppi HM, Paakko P, Turpeenniemi-Hujanen T, Vahakangas KH. 2010. p53 in head and neck cancer: functional consequences and environmental implications of TP53 mutations. *Head Neck Oncol* **2**: 36.

Peng HH, Ko HH, Chi NC, Wang YP, Lee HC, Pan PY, Kuo MY, Cheng SJ. 2019. Upregulated NPM1 is an independent biomarker to predict progression and prognosis of oral squamous cell carcinomas in Taiwan. *Head Neck*.

Peng W, Zhang Y, Zhu R, Mechref Y. 2017. Comparative membrane proteomics analyses of breast cancer cell lines to understand the molecular mechanism of breast cancer brain metastasis. *Electrophoresis* **38**: 2124-2134.

Peng Y, Chen L, Li C, Lu W, Chen J. 2001. Inhibition of MDM2 by hsp90 contributes to mutant p53 stabilization. *J Biol Chem* **276**: 40583-40590.

Perera Y, Farina HG, Gil J, Rodriguez A, Benavent F, Castellanos L, Gomez RE, Acevedo BE, Alonso DF, Perea SE. 2009. Anticancer peptide CIGB-300 binds to nucleophosmin/B23, impairs its CK2-mediated phosphorylation, and leads to apoptosis through its nucleolar disassembly activity. *Mol Cancer Ther* **8**: 1189-1196.

Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. 2015. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol* **33**: 290-295.

Peter M, Nakagawa J, Doree M, Labbe JC, Nigg EA. 1990. Identification of major nucleolar proteins as candidate mitotic substrates of cdc2 kinase. *Cell* **60**: 791-801.

Pfister JA, D'Mello SR. 2016. Regulation of Neuronal Survival by Nucleophosmin 1 (NPM1) Is Dependent on Its Expression Level, Subcellular Localization, and Oligomerization Status. *J Biol Chem* **291**: 20787-20797.

Phatnani HP, Greenleaf AL. 2006. Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev* **20**: 2922-2936.

Phi JH, Sun CH, Lee SH, Lee S, Park I, Choi SA, Park SH, Lee JY, Wang KC, Kim SK et al. 2019. NPM1 as a potential therapeutic target for atypical teratoid/rhabdoid tumors. *BMC Cancer* **19**: 848.

Philpott A, Leno GH, Laskey RA. 1991. Sperm decondensation in Xenopus egg cytoplasm is mediated by nucleoplasmin. *Cell* **65**: 569-578.

Pianta A, Puppin C, Franzoni A, Fabbro D, Di Loreto C, Bulotta S, Deganuto M, Paron I, Tell G, Puxeddu E et al. 2010. Nucleophosmin is overexpressed in thyroid tumors. *Biochem Biophys Res Commun* **397**: 499-504.

Pianta A, Puppin C, Passon N, Franzoni A, Romanello M, Tell G, Di Loreto C, Bulotta S, Russo D, Damante G. 2011. Nucleophosmin delocalization in thyroid tumour cells. *Endocr Pathol* **22**: 18-23.

Pikaard CS, Haag JR, Ream T, Wierzbicki AT. 2008. Roles of RNA polymerase IV in gene silencing. *Trends Plant Sci* **13**: 390-397.

Planells-Palop V, Hazazi A, Feichtinger J, Jezkova J, Thallinger G, Alsiwiehri NO, Almutairi M, Parry L, Wakeman JA, McFarlane RJ. 2017. Human germ/stem cell-specific gene TEX19 influences cancer cell proliferation and cancer prognosis. *Mol Cancer* **16**: 84.

Platonova O, Akey IV, Head JF, Akey CW. 2011. Crystal structure and function of human nucleoplasmin (npm2): a histone chaperone in oocytes and embryos. *Biochemistry* **50**: 8078-8089.

Poletto M, Lirussi L, Wilson DM, 3rd, Tell G. 2014. Nucleophosmin modulates stability, activity, and nucleolar accumulation of base excision repair proteins. *Mol Biol Cell* **25**: 1641-1652.

Poletto M, Malfatti MC, Dorjsuren D, Scognamiglio PL, Marasco D, Vascotto C, Jadhav A, Maloney DJ, Wilson DM, 3rd, Simeonov A et al. 2016. Inhibitors of the apurinic/apyrimidinic endonuclease 1 (APE1)/nucleophosmin (NPM1) interaction that display anti-tumor properties. *Mol Carcinog* **55**: 688-704.

Polo SE, Theocharis SE, Grandin L, Gambotti L, Antoni G, Savignoni A, Asselain B, Patsouris E, Almouzni G. 2010. Clinical significance and prognostic value of chromatin assembly factor-1 overexpression in human solid tumours. *Histopathology* **57**: 716-724.

Pozzo F, Bittolo T, Vendramini E, Bomben R, Bulian P, Rossi FM, Zucchetto A, Tissino E, Degan M, D'Arena G et al. 2017. NOTCH1-mutated chronic lymphocytic leukemia cells are characterized by a MYC-related overexpression of nucleophosmin 1 and ribosome-associated components. *Leukemia* **31**: 2407-2415.

Pradeepa MM, Nikhil G, Hari Kishore A, Bharath GN, Kundu TK, Rao MR. 2009. Acetylation of transition protein 2 (TP2) by KAT3B (p300) alters its DNA condensation property and interaction with putative histone chaperone NPM3. *J Biol Chem* **284**: 29956-29967.

Price DH. 2018. Transient pausing by RNA polymerase II. *Proc Natl Acad Sci U S A* **115**: 4810-4812.

Proudfoot N. 2004. New perspectives on connecting messenger RNA 3' end formation to transcription. *Curr Opin Cell Biol* **16**: 272-278.

Proudfoot NJ, Furger A, Dye MJ. 2002. Integrating mRNA processing with transcription. *Cell* **108**: 501-512.

Qi W, Shakalya K, Stejskal A, Goldman A, Beeck S, Cooke L, Mahadevan D. 2008. NSC348884, a nucleophosmin inhibitor disrupts oligomer formation and induces apoptosis in human cancer cells. *Oncogene* **27**: 4210-4220.

Qin FX, Shao HY, Chen XC, Tan S, Zhang HJ, Miao ZY, Wang L, Hui C, Zhang L. 2011. Knockdown of NPM1 by RNA interference inhibits cells proliferation and induces apoptosis in leukemic cell line. *Int J Med Sci* **8**: 287-294.

Qin Y, Choi Y, Zhao H, Simpson JL, Chen ZJ, Rajkovic A. 2007. NOBOX homeobox mutation causes premature ovarian failure. *Am J Hum Genet* **81**: 576-581.

Qu S, Li XY, Liang ZG, Li L, Huang ST, Li JQ, Li DR, Zhu XD. 2016. Protein expression of nucleophosmin, annexin A3 and nm23-H1 correlates with human nasopharyngeal carcinoma radioresistance in vivo. *Oncol Lett* **12**: 615-620.

Quartuccio SM, Schindler K. 2015. Functions of Aurora kinase C in meiosis and cancer. *Front Cell Dev Biol* **3**: 50.

Quevedo M, Meert L, Dekker MR, Dekkers DHW, Brandsma JH, van den Berg DLC, Ozgur Z, van IWFJ, Demmers J, Fornerod M et al. 2019. Mediator complex interaction partners organize the transcriptional network that defines neural stem cells. *Nat Commun* **10**: 2669.

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841-842.

Quivy JP, Grandi P, Almouzni G. 2001. Dimerization of the largest subunit of chromatin assembly factor 1: importance in vitro and during Xenopus early development. *Embo j* **20**: 2015-2027.

Radman-Livaja M, Rando OJ. 2010. Nucleosome positioning: how is it established, and why does it matter? *Dev Biol* **339**: 258-266.

Ragos V, N SM, Tsiambas E, Baliou E, S NM, Tsoukalas N, E EP, P PF. 2018. p53 mutations in oral cavity carcinoma. *J buon* **23**: 1569-1572.

Raisner RM, Hartley PD, Meneghini MD, Bao MZ, Liu CL, Schreiber SL, Rando OJ, Madhani HD. 2005. Histone variant H2A.Z marks the 5' ends of both active and inactive genes in euchromatin. *Cell* **123**: 233-248.

Raivich G, Bohatschek M, Da Costa C, Iwata O, Galiano M, Hristova M, Nateri AS, Makwana M, Riera-Sans L, Wolfer DP et al. 2004. The AP-1 transcription factor c-Jun is required for efficient axonal regeneration. *Neuron* **43**: 57-67.

Rajkovic A, Pangas SA, Ballow D, Suzumori N, Matzuk MM. 2004. NOBOX deficiency disrupts early folliculogenesis and oocyte-specific gene expression. *Science* **305**: 1157-1159.

Ramirez F, Dundar F, Diehl S, Gruning BA, Manke T. 2014. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* **42**: W187-191.

Ramos I, Prado A, Finn RM, Muga A, Ausio J. 2005. Nucleoplasmin-mediated unfolding of chromatin involves the displacement of linker-associated chromatin proteins. *Biochemistry* **44**: 8274-8281.

Ramsamooj P, Notario V, Dritschilo A. 1995. Modification of nucleolar protein B23 after exposure to ionizing radiation. *Radiat Res* **143**: 158-164.

Rau R, Brown P. 2009. Nucleophosmin (NPM1) mutations in adult and childhood acute myeloid leukaemia: towards definition of a new leukaemia entity. *Hematol Oncol* **27**: 171-181.

Reboutier D, Troadec MB, Cremet JY, Fukasawa K, Prigent C. 2012. Nucleophosmin/B23 activates Aurora A at the centrosome through phosphorylation of serine 89. *J Cell Biol* **197**: 19-26.

Reddy BA, Jeronimo C, Robert F. 2017. Recent Perspectives on the Roles of Histone Chaperones in Transcription Regulation. *Current Molecular Biology Reports* **3**: 1-10.

Ren Z, Aerts JL, Pen JJ, Heirman C, Breckpot K, De Greve J. 2015. Phosphorylated STAT3 physically interacts with NPM and transcriptionally enhances its expression in cancer. *Oncogene* **34**: 1650-1657.

Ren Z, Aerts JL, Vandenplas H, Wang JA, Gorbenko O, Chen JP, Giron P, Heirman C, Goyvaerts C, Zacksenhaus E et al. 2016. Phosphorylated STAT5 regulates p53 expression via BRCA1/BARD1-NPM1 and MDM2. *Cell Death Dis* **7**: e2560.

Richard P, Manley JL. 2009. Transcription termination by nuclear RNA polymerases. *Genes Dev* **23**: 1247-1269.

Richards EJ, Elgin SC. 2002. Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. *Cell* **108**: 489-500.

Roberts C, Sutherland HF, Farmer H, Kimber W, Halford S, Carey A, Brickman JM, Wynshaw-Boris A, Scambler PJ. 2002. Targeted mutagenesis of the Hira gene results in gastrulation defects and patterning abnormalities of mesoendodermal derivatives prior to early embryonic lethality. *Mol Cell Biol* **22**: 2318-2328.

Robinson DCL, Dilworth FJ. 2018. Epigenetic Regulation of Adult Myogenesis. *Curr Top Dev Biol* **126**: 235-284.

Robinson PJ, Fairall L, Huynh VA, Rhodes D. 2006. EM measurements define the dimensions of the "30-nm" chromatin fiber: evidence for a compact, interdigitated structure. *Proc Natl Acad Sci U S A* **103**: 6506-6511.

Robinson PJ, Trnka MJ, Bushnell DA, Davis RE, Mattei PJ, Burlingame AL, Kornberg RD. 2016. Structure of a Complete Mediator-RNA Polymerase II Pre-Initiation Complex. *Cell* **166**: 1411-1422.e1416.

Roeder RG. 1996. The role of general initiation factors in transcription by RNA polymerase II. *Trends Biochem Sci* **21**: 327-335.

Roeder RG. 1998. Role of general and gene-specific cofactors in the regulation of eukaryotic transcription. *Cold Spring Harb Symp Quant Biol* **63**: 201-218.

Roeder RG. 2003. Lasker Basic Medical Research Award. The eukaryotic transcriptional machinery: complexities and mechanisms unforeseen. *Nat Med* **9**: 1239-1244.

Roeder RG. 2005. Transcriptional regulation and the role of diverse coactivators in animal cells. *FEBS Lett* **579**: 909-915.

Roeder RG, Rutter WJ. 1969. Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. *Nature* **224**: 234-237.

Roeder RG, Rutter WJ. 1970. Specific nucleolar and nucleoplasmic RNA polymerases. *Proc Natl Acad Sci U S A* **65**: 675-682.

Rosinska S, Filipek A. 2018. Interaction of CacyBP/SIP with NPM1 and its influence on NPM1 localization and function in oxidative stress. *J Cell Physiol* **233**: 8826-8838.

Rosonina E, Kaneko S, Manley JL. 2006. Terminating the transcript: breaking up is hard to do. *Genes Dev* **20**: 1050-1056.

Roulland Y, Ouararhni K, Naidenov M, Ramos L, Shuaib M, Syed SH, Lone IN, Boopathi R, Fontaine E, Papai G et al. 2016. The Flexible Ends of CENP-A Nucleosome Are Required for Mitotic Fidelity. *Mol Cell* **63**: 674-685.

Routh A, Sandin S, Rhodes D. 2008. Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure. *Proc Natl Acad Sci U S A* **105**: 8872-8877.

Rowley MJ, Corces VG. 2018. Organizational principles of 3D genome architecture. *Nat Rev Genet* **19**: 789-800.

Roy AL, Singer DS. 2015. Core promoters in transcription: old problem, new insights. *Trends Biochem Sci* **40**: 165-171.

Rubbi CP, Milner J. 2003. Disruption of the nucleolus mediates stabilization of p53 in response to DNA damage and other stresses. *Embo j* **22**: 6068-6077.

Ruggero D, Pandolfi PP. 2003. Does the ribosome translate cancer? *Nat Rev Cancer* **3**: 179-192.

Sachdev R, Mandal AK, Singh I, Agarwal AK. 2008. Progressive rise of c fos expression from premalignant to malignant lesions of oral cavity. *Med Oral Patol Oral Cir Bucal* **13**: E683-686.

Sainsbury S, Niesser J, Cramer P. 2013. Structure and function of the initially transcribing RNA polymerase II-TFIIB complex. *Nature* **493**: 437-440.

Sakhinia E, Glennie C, Hoyland JA, Menasce LP, Brady G, Miller C, Radford JA, Byers RJ. 2007. Clinical quantitation of diagnostic and predictive gene expression levels in follicular and diffuse large B-cell lymphoma by RT-PCR gene expression profiling. *Blood* **109**: 3922-3928.

Samad MA, Okuwaki M, Haruki H, Nagata K. 2007. Physical and functional interaction between a nucleolar protein nucleophosmin/B23 and adenovirus basic core proteins. *FEBS Lett* **581**: 3283-3288.

Sampath J, Sun D, Kidd VJ, Grenet J, Gandhi A, Shapiro LH, Wang Q, Zambetti GP, Schuetz JD. 2001. Mutant p53 cooperates with ETS and selectively up-regulates human MDR1 not MRP1. *J Biol Chem* **276**: 39359-39367.

Sanchez F, Adriaenssens T, Romero S, Smitz J. 2009. Quantification of oocyte-specific transcripts in follicle-enclosed oocytes during antral development and maturation in vitro. *Mol Hum Reprod* **15**: 539-550.

Sansoni V, Casas-Delucchi CS, Rajan M, Schmidt A, Bonisch C, Thomae AW, Staege MS, Hake SB, Cardoso MC, Imhof A. 2014. The histone variant H2A.Bbd is enriched at sites of DNA synthesis. *Nucleic Acids Res* **42**: 6405-6420.

Santenard A, Ziegler-Birling C, Koch M, Tora L, Bannister AJ, Torres-Padilla ME. 2010. Heterochromatin formation in the mouse embryo requires critical residues of the histone variant H3.3. *Nat Cell Biol* **12**: 853-862.

Santos-Rosa H, Schneider R, Bannister AJ, Sherriff J, Bernstein BE, Emre NC, Schreiber SL, Mellor J, Kouzarides T. 2002. Active genes are tri-methylated at K4 of histone H3. *Nature* **419**: 407-411.

Sanulli S, Trnka MJ, Dharmarajan V, Tibble RW, Pascal BD, Burlingame AL, Griffin PR, Gross JD, Narlikar GJ. 2019. HP1 reshapes nucleosome core to promote heterochromatin phase separation. *Nature*.

Saredi G, Huang H, Hammond CM, Alabert C, Bekker-Jensen S, Forne I, Reveron-Gomez N, Foster BM, Mlejnkova L, Bartke T et al. 2016. H4K20me0 marks post-replicative chromatin and recruits the TONSL-MMS22L DNA repair complex. *Nature* **534**: 714-718.

Sartorelli V, Puri PL. 2018. Shaping Gene Expression by Landscaping Chromatin Architecture: Lessons from a Master. *Mol Cell* **71**: 375-388.

Satkunanathan S, Thorpe R, Zhao Y. 2017. The function of DNA binding protein nucleophosmin in AAV replication. *Virology* **510**: 46-54.

Sato K, Hayami R, Wu W, Nishikawa T, Nishikawa H, Okuda Y, Ogata H, Fukuda M, Ohta T. 2004. Nucleophosmin/B23 is a candidate substrate for the BRCA1-BARD1 ubiquitin ligase. *J Biol Chem* **279**: 30919-30922.

Savkur RS, Olson MO. 1998. Preferential cleavage in pre-ribosomal RNA byprotein B23 endoribonuclease. *Nucleic Acids Res* **26**: 4508-4515.

Sawatsubashi S, Murata T, Lim J, Fujiki R, Ito S, Suzuki E, Tanabe M, Zhao Y, Kimura S, Fujiyama S et al. 2010. A histone chaperone, DEK, transcriptionally coactivates a nuclear receptor. *Genes Dev* **24**: 159-170.

Schlissel G, Rine J. 2019. The nucleosome core particle remembers its position through DNA replication and RNA transcription. *Proc Natl Acad Sci U S A* **116**: 20605-20611.

Schmidt-Zachmann MS, Hugle-Dorr B, Franke WW. 1987. A constitutive nucleolar protein identified as a member of the nucleoplasmin family. *Embo j* **6**: 1881-1890.

Schones DE, Cui K, Cuddapah S, Roh TY, Barski A, Wang Z, Wei G, Zhao K. 2008. Dynamic regulation of nucleosome positioning in the human genome. *Cell* **132**: 887-898.

Schubeler D, MacAlpine DM, Scalzo D, Wirbelauer C, Kooperberg C, van Leeuwen F, Gottschling DE, O'Neill LP, Turner BM, Delrow J et al. 2004. The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev* **18**: 1263-1271.

Schudrowitz N, Takagi S, Wessel GM, Yajima M. 2017. Germline factor DDX4 functions in blood-derived cancer cell phenotypes. *Cancer Sci* **108**: 1612-1619.

Schwabish MA, Struhl K. 2006. Asf1 mediates histone eviction and deposition during elongation by RNA polymerase II. *Mol Cell* **22**: 415-422.

Schwartzentruber J, Korshunov A, Liu XY, Jones DT, Pfaff E, Jacob K, Sturm D, Fontebasso AM, Quang DA, Tonjes M et al. 2012. Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. *Nature* **482**: 226-231.

Scott DD, Oeffinger M. 2016. Nucleolin and nucleophosmin: nucleolar proteins with multiple functions in DNA repair. *Biochem Cell Biol* **94**: 419-432.

Sekhar KR, Benamar M, Venkateswaran A, Sasi S, Penthala NR, Crooks PA, Hann SR, Geng L, Balusu R, Abbas T et al. 2014. Targeting nucleophosmin 1 represents a rational strategy for radiation sensitization. *Int J Radiat Oncol Biol Phys* **89**: 1106-1114.

Senapati P, Dey S, Sudarshan D, Bhattacharya A, G S, Das S, Sudevan S, Maliekal TT, Kundu TK. 2019. Histone chaperone Nucleophosmin regulates transcription of key genes involved in oral tumorigenesis. *bioRxiv*: 852095.

Senapati P, Dey S, Sudarshan D, Das S, Kumar M, Kaypee S, Mohiyuddin A, Kodaganur GS, Kundu TK. 2018. Oncogene c-fos and mutant R175H p53 regulate expression of Nucleophosmin implicating cancer manifestation. *Febs j* **285**: 3503-3524.

Senapati P, Sudarshan D, Gadad SS, Shandilya J, Swaminathan V, Kundu TK. 2015. Methods to study histone chaperone function in nucleosome assembly and chromatin transcription. *Methods Mol Biol* **1288**: 375-394.

Shackleford GM, Ganguly A, MacArthur CA. 2001. Cloning, expression and nuclear localization of human NPM3, a member of the nucleophosmin/nucleoplasmin family of nuclear chaperones. *BMC Genomics* **2**: 8.

Shaiu WL, Hsieh TS. 1998. Targeting to transcriptionally active loci by the hydrophilic N-terminal domain of Drosophila DNA topoisomerase I. *Mol Cell Biol* **18**: 4358-4367.

Shandilya J, Senapati P, Dhanasekaran K, Bangalore SS, Kumar M, Kishore AH, Bhat A, Kodaganur GS, Kundu TK. 2014a. Phosphorylation of multifunctional nucleolar protein nucleophosmin (NPM1) by aurora kinase B is critical for mitotic progression. *FEBS Lett* **588**: 2198-2205.

Shandilya J, Senapati P, Hans F, Menoni H, Bouvet P, Dimitrov S, Angelov D, Kundu TK. 2014b. Centromeric histone variant CENP-A represses acetylation-dependent chromatin transcription that is relieved by histone chaperone NPM1. *J Biochem* **156**: 221-227.

Shandilya J, Swaminathan V, Gadad SS, Choudhari R, Kodaganur GS, Kundu TK. 2009. Acetylated NPM1 localizes in the nucleoplasm and regulates transcriptional activation of genes implicated in oral cancer manifestation. *Mol Cell Biol* **29**: 5115-5127.

Shang C, Wang W, Liao Y, Chen Y, Liu T, Du Q, Huang J, Liang Y, Liu J, Zhao Y et al. 2018. LNMICC Promotes Nodal Metastasis of Cervical Cancer by Reprogramming Fatty Acid Metabolism. *Cancer Res* **78**: 877-890.

Sharpless NE. 2005. INK4a/ARF: a multifunctional tumor suppressor locus. *Mutat Res* **576**: 22-38.

Shields LB, Gercel-Taylor C, Yashar CM, Wan TC, Katsanis WA, Spinnato JA, Taylor DD. 1997. Induction of immune responses to ovarian tumor antigens by multiparity. *J Soc Gynecol Investig* **4**: 298-304.

Shinagawa T, Huynh LM, Takagi T, Tsukamoto D, Tomaru C, Kwak HG, Dohmae N, Noguchi J, Ishii S. 2015. Disruption of Th2a and Th2b genes causes defects in spermatogenesis. *Development* **142**: 1287-1292.

Sidi S, Bouchier-Hayes L. 2017. Direct pro-apoptotic role for NPM1 as a regulator of PIDDosome formation. *Mol Cell Oncol* **4**: e1348325.

Sierecki E. 2018. The Mediator complex and the role of protein-protein interactions in the gene regulation machinery. *Semin Cell Dev Biol*.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J et al. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**: 539.

Simpson AJ, Caballero OL, Jungbluth A, Chen YT, Old LJ. 2005. Cancer/testis antigens, gametogenesis and cancer. *Nat Rev Cancer* **5**: 615-625.

Simpson RT. 1978. Structure of the chromatosome, a chromatin particle containing 160 base pairs of DNA and all the histones. *Biochemistry* **17**: 5524-5531.

Sims RJ, 3rd, Belotserkovskaya R, Reinberg D. 2004. Elongation by RNA polymerase II: the short and long of it. *Genes Dev* **18**: 2437-2468.

Singh S, Kumar M, Kumar S, Sen S, Upadhyay P, Bhattacharjee S, M N, Tomar VS, Roy S, Dutt A et al. 2019. The cancer-associated, gain-of-function TP53 variant P152Lp53 activates multiple signaling pathways implicated in tumorigenesis. *J Biol Chem* **294**: 14081-14095.

Skaar TC, Prasad SC, Sharareh S, Lippman ME, Brunner N, Clarke R. 1998. Two-dimensional gel electrophoresis analyses identify nucleophosmin as an estrogen regulated protein associated with acquired estrogen-independence in human breast cancer cells. *J Steroid Biochem Mol Biol* **67**: 391-402.

So CH, Michal AM, Mashayekhi R, Benovic JL. 2012. G protein-coupled receptor kinase 5 phosphorylates nucleophosmin and regulates cell sensitivity to polo-like kinase 1 inhibition. *J Biol Chem* **287**: 17088-17099.

Soboleva TA, Nekrasov M, Pahwa A, Williams R, Huttley GA, Tremethick DJ. 2011. A unique H2A histone variant occupies the transcriptional start site of active genes. *Nat Struct Mol Biol* **19**: 25-30.

Soboleva TA, Parker BJ, Nekrasov M, Hart-Smith G, Tay YJ, Tng WQ, Wilkins M, Ryan D, Tremethick DJ. 2017. A new link between transcriptional initiation and pre-mRNA splicing: The RNA binding histone variant H2A.B. *PLoS Genet* **13**: e1006633.

Soekarman D, von Lindern M, Daenen S, de Jong B, Fonatsch C, Heinze B, Bartram C, Hagemeijer A, Grosveld G. 1992. The translocation (6;9) (p23;q34) shows consistent rearrangement of two genes and defines a myeloproliferative disorder with specific clinical features. *Blood* **79**: 2990-2997.

Song F, Chen P, Sun D, Wang M, Dong L, Liang D, Xu RM, Zhu P, Li G. 2014. Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. *Science* **344**: 376-380.

Song G, Wang L. 2009. Nuclear receptor SHP activates miR-206 expression via a cascade dual inhibitory mechanism. *PLoS One* **4**: e6880.

Soshnev AA, Josefowicz SZ, Allis CD. 2016. Greater Than the Sum of Parts: Complexity of the Dynamic Epigenome. *Mol Cell* **62**: 681-694.

Soussi T, Wiman KG. 2015. TP53: an oncogene in disguise. *Cell Death Differ* **22**: 1239-1249.

Soutourina J. 2018. Transcription regulation by the Mediator complex. *Nat Rev Mol Cell Biol* **19**: 262-274.

Soyal SM, Amleh A, Dean J. 2000. FIGalpha, a germ cell-specific transcription factor required for ovarian follicle formation. *Development* **127**: 4645-4654.

Spector DL, Ochs RL, Busch H. 1984. Silver staining, immunofluorescence, and immunoelectron microscopic localization of nucleolar phosphoproteins B23 and C23. *Chromosoma* **90**: 139-148.

Sportoletti P, Grisendi S, Majid SM, Cheng K, Clohessy JG, Viale A, Teruya-Feldstein J, Pandolfi PP. 2008. Npm1 is a haploinsufficient suppressor of myeloid and lymphoid malignancies in the mouse. *Blood* **111**: 3859-3862.

Springhetti EM, Istomina NE, Whisstock JC, Nikitina T, Woodcock CL, Grigoryev SA. 2003. Role of the M-loop and reactive center loop domains in the folding and bridging of nucleosome arrays by MENT. *J Biol Chem* **278**: 43384-43393.

Stambolsky P, Tabach Y, Fontemaggi G, Weisz L, Maor-Aloni R, Siegfried Z, Shiff I, Kogan I, Shay M, Kalo E et al. 2010. Modulation of the vitamin D3 response by cancer-associated mutant p53. *Cancer Cell* **17**: 273-285.

Stark C, Breitkreutz BJ, Reguly T, Boucher L, Breitkreutz A, Tyers M. 2006. BioGRID: a general repository for interaction datasets. *Nucleic Acids Res* **34**: D535-539.

Stein A. 1979. DNA folding by histones: the kinetics of chromatin core particle reassembly and the interaction of nucleosomes with histones. *J Mol Biol* **130**: 103-134.

Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, Stein TI, Nudel R, Lieder I, Mazor Y et al. 2016. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Curr Protoc Bioinformatics* **54**: 1.30.31-31.30.33.

Steurer B, Janssens RC, Geverts B, Geijer ME, Wienholz F, Theil AF, Chang J, Dealy S, Pothof J, van Cappellen WA et al. 2018. Live-cell analysis of endogenous GFP-RPB1 uncovers rapid turnover of initiating and promoter-paused RNA Polymerase II. *Proc Natl Acad Sci U S A* **115**: E4368-e4376.

Storck S, Shukla M, Dimitrov S, Bouvet P. 2007. Functions of the histone chaperone nucleolin in diseases. *Subcell Biochem* **41**: 125-144.

Strahl BD, Allis CD. 2000. The language of covalent histone modifications. *Nature* **403**: 41-45.

Straube K, Blackwell JS, Jr., Pemberton LF. 2010. Nap1 and Chz1 have separate Htz1 nuclear import and assembly functions. *Traffic* **11**: 185-197.

Struhl K. 1999. Fundamentally different logic of gene regulation in eukaryotes and prokaryotes. *Cell* **98**: 1-4.

Struhl K, Segal E. 2013. Determinants of nucleosome positioning. *Nat Struct Mol Biol* **20**: 267-273.

Subong EN, Shue MJ, Epstein JI, Briggman JV, Chan PK, Partin AW. 1999. Monoclonal antibody to prostate cancer nuclear matrix protein (PRO:4-216) recognizes nucleophosmin/B23. *Prostate* **39**: 298-304.

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES et al. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**: 15545-15550.

Sui G, Affar el B, Shi Y, Brignone C, Wall NR, Yin P, Donohoe M, Luke MP, Calvo D, Grossman SR. 2004. Yin Yang 1 is a negative regulator of p53. *Cell* **117**: 859-872.

Suzuki H, Ahn HW, Chu T, Bowden W, Gassei K, Orwig K, Rajkovic A. 2012. SOHLH1 and SOHLH2 coordinate spermatogonial differentiation. *Dev Biol* **361**: 301-312.

Swaminathan V, Kishore AH, Febitha KK, Kundu TK. 2005. Human histone chaperone nucleophosmin enhances acetylation-dependent chromatin transcription. *Mol Cell Biol* **25**: 7534-7545.

Sylvester AM, Chen D, Krasinski K, Andrés V. 1998. Role of c-fos and E2F in the induction of cyclin A transcription and vascular smooth muscle cell proliferation. *J Clin Invest* **101**: 940-948.

Sylvestre EL, Pennetier S, Bureau M, Robert C, Sirard MA. 2010. Investigating the potential of genes preferentially expressed in oocyte to induce chromatin remodeling in somatic cells. *Cell Reprogram* **12**: 519-528.

Szebeni A, Herrera JE, Olson MO. 1995. Interaction of nucleolar protein B23 with peptides related to nuclear localization signals. *Biochemistry* **34**: 8037-8042.

Szebeni A, Hingorani K, Negi S, Olson MO. 2003. Role of protein kinase CK2 phosphorylation in the molecular chaperone activity of nucleolar protein b23. *J Biol Chem* **278**: 9107-9115.

Szebeni A, Mehrotra B, Baumann A, Adam SA, Wingfield PT, Olson MO. 1997. Nucleolar protein B23 stimulates nuclear import of the HIV-1 Rev protein and NLS-conjugated albumin. *Biochemistry* **36**: 3941-3949.

Szebeni A, Olson MO. 1999. Nucleolar protein B23 has molecular chaperone activities. *Protein Sci* **8**: 905-912.

Szenker E, Lacoste N, Almouzni G. 2012. A developmental requirement for HIRA-dependent H3.3 deposition revealed at gastrulation in Xenopus. *Cell Rep* **1**: 730-740.

Szenker E, Ray-Gallet D, Almouzni G. 2011. The double face of the histone variant H3.3. *Cell Res* **21**: 421-434.

Tachiwana H, Kurumizaka H. 2011. Structure of the CENP-A nucleosome and its implications for centromeric chromatin architecture. *Genes Genet Syst* **86**: 357-364.

Tago K, Chiocca S, Sherr CJ. 2005. Sumoylation induced by the Arf tumor suppressor: a p53-independent function. *Proc Natl Acad Sci U S A* **102**: 7689-7694.

Takemura M, Ohta N, Furuichi Y, Takahashi T, Yoshida S, Olson MO, Umekawa H. 1994. Stimulation of calf thymus DNA polymerase alpha activity by nucleolar protein B23. *Biochem Biophys Res Commun* **199**: 46-51.

Takemura M, Sato K, Nishio M, Akiyama T, Umekawa H, Yoshida S. 1999. Nucleolar protein B23.1 binds to retinoblastoma protein and synergistically stimulates DNA polymerase alpha activity. *J Biochem* **125**: 904-909.

Talbert PB, Henikoff S. 2010. Histone variants--ancient wrap artists of the epigenome. *Nat Rev Mol Cell Biol* **11**: 264-275.

Talbert PB, Meers MP, Henikoff S. 2019. Old cogs, new tricks: the evolution of gene expression in a chromatin context. *Nat Rev Genet* **20**: 283-297.

Tamada H, Van Thuan N, Reed P, Nelson D, Katoku-Kikyo N, Wudel J, Wakayama T, Kikyo N. 2006. Chromatin decondensation and nuclear reprogramming by nucleoplasmin. *Mol Cell Biol* **26**: 1259-1271.

Tan S, Hunziker Y, Sargent DF, Richmond TJ. 1996. Crystal structure of a yeast TFIIA/TBP/DNA complex. *Nature* **381**: 127-151.

Tanabe H, Muller S, Neusser M, von Hase J, Calcagno E, Cremer M, Solovei I, Cremer C, Cremer T. 2002. Evolutionary conservation of chromosome territory arrangements in cell nuclei from higher primates. *Proc Natl Acad Sci U S A* **99**: 4424-4429.

Tanaka H, Matsuoka Y, Onishi M, Kitamura K, Miyagawa Y, Nishimura H, Tsujimura A, Okuyama A, Nishimune Y. 2006. Expression profiles and single-nucleotide polymorphism analysis of human HANP1/H1T2 encoding a histone H1-like protein. *Int J Androl* **29**: 353-359.

Tanaka M, Hennebold JD, Macfarlane J, Adashi EY. 2001. A mammalian oocyte-specific linker histone gene H1oo: homology with the genes for the oocyte-specific cleavage stage histone (cs-H1) of sea urchin and the B4/H1M histone of the frog. *Development* **128**: 655-664.

Tanaka M, Sasaki H, Kino I, Sugimura T, Terada M. 1992. Genes preferentially expressed in embryo stomach are predominantly expressed in gastric cancer. *Cancer Res* **52**: 3372-3377.

Tanaka N, Zhao M, Tang L, Patel AA, Xi Q, Van HT, Takahashi H, Osman AA, Zhang J, Wang J et al. 2018. Gain-of-function mutant p53 promotes the oncogenic potential of head and neck squamous cell carcinoma cells by targeting the transcription factors FOXO3a and FOXM1. *Oncogene* **37**: 1279-1292.

Tang J, Chen H, Wong CC, Liu D, Li T, Wang X, Ji J, Sung JJ, Fang JY, Yu J. 2018. DEAD-box helicase 27 promotes colorectal cancer growth and metastasis and predicts poor survival in CRC patients. *Oncogene* **37**: 3006-3021.

Tanikawa C, Ueda K, Nakagawa H, Yoshida N, Nakamura Y, Matsuda K. 2009. Regulation of protein Citrullination through p53/PADI4 network in DNA damage response. *Cancer Res* **69**: 8761-8769.

Tarapore P, Shinmura K, Suzuki H, Tokuyama Y, Kim SH, Mayeda A, Fukasawa K. 2006. Thr199 phosphorylation targets nucleophosmin to nuclear speckles and represses pre-mRNA processing. *FEBS Lett* **580**: 399-409.

Taylor A, Allende CC, Weinmann R, Allende JE. 1987. The phosphorylation of nucleoplasmin by casein kinase-2 is resistant to heparin inhibition. *FEBS Lett* **226**: 109-114.

Teufel DP, Freund SM, Bycroft M, Fersht AR. 2007. Four domains of p300 each bind tightly to a sequence spanning both transactivation subdomains of p53. *Proc Natl Acad Sci U S A* **104**: 7009-7014.

Thakar A, Gupta P, Ishibashi T, Finn R, Silva-Moreno B, Uchiyama S, Fukui K, Tomschik M, Ausio J, Zlatanova J. 2009. H2A.Z and H3.3 histone variants affect nucleosome structure: biochemical and biophysical studies. *Biochemistry* **48**: 10852-10857.

Thinakaran G, Ojala J, Bag J. 1993. Expression of c-jun/AP-1 during myogenic differentiation in mouse C2C12 myoblasts. *FEBS Lett* **319**: 271-276.

Thoma F, Koller T, Klug A. 1979. Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. *J Cell Biol* **83**: 403-427.

Thomas JO, Stott K. 2012. H1 and HMGB1: modulators of chromatin structure. *Biochem Soc Trans* **40**: 341-346.

Thomas JO, Travers AA. 2001. HMG1 and 2, and related 'architectural' DNA-binding proteins. *Trends Biochem Sci* **26**: 167-174.

Thomas MC, Chiang CM. 2006. The general transcription machinery and general cofactors. *Crit Rev Biochem Mol Biol* **41**: 105-178.

Thompson PR, Wang D, Wang L, Fulco M, Pediconi N, Zhang D, An W, Ge Q, Roeder RG, Wong J et al. 2004. Regulation of the p300 HAT domain via a novel activation loop. *Nat Struct Mol Biol* **11**: 308-315.

Tirode F, Busso D, Coin F, Egly JM. 1999. Reconstitution of the transcription factor TFIIH: assignment of functions for the three enzymatic subunits, XPB, XPD, and cdk7. *Mol Cell* **3**: 87-95.

Tokusumi Y, Ma Y, Song X, Jacobson RH, Takada S. 2007. The new core promoter element XCPE1 (X Core Promoter Element 1) directs activator-, mediator-, and TATA-binding protein-dependent but TFIID-independent RNA polymerase II transcription from TATA-less promoters. *Mol Cell Biol* **27**: 1844-1858.

Tokuyama Y, Horn HF, Kawamura K, Tarapore P, Fukasawa K. 2001. Specific phosphorylation of nucleophosmin on Thr(199) by cyclin-dependent kinase 2-cyclin E and its role in centrosome duplication. *J Biol Chem* **276**: 21529-21537.

Tolstorukov MY, Goldman JA, Gilbert C, Ogryzko V, Kingston RE, Park PJ. 2012. Histone variant H2A.Bbd is associated with active transcription and mRNA processing in human cells. *Mol Cell* **47**: 596-607.

Trouche D, Grigoriev M, Lenormand JL, Robin P, Leibovitch SA, Sassone-Corsi P, Harel-Bellan A. 1993. Repression of c-fos promoter by MyoD on muscle cell differentiation. *Nature* **363**: 79-82.

Trouche D, Masutani H, Groisman R, Robin P, Lenormand JL, Harel-Bellan A. 1995. Myogenin binds to and represses c-fos promoter. *FEBS Lett* **361**: 140-144.

Tsai FT, Sigler PB. 2000. Structural basis of preinitiation complex assembly on human pol II promoters. *Embo j* **19**: 25-36.

Tsubota T, Berndsen CE, Erkmann JA, Smith CL, Yang L, Freitas MA, Denu JM, Kaufman PD. 2007. Histone H3-K56 acetylation is catalyzed by histone chaperone-dependent complexes. *Mol Cell* **25**: 703-712.

Tsui KH, Cheng AJ, Chang P, Pan TL, Yung BY. 2004. Association of nucleophosmin/B23 mRNA expression with clinical outcome in patients with bladder carcinoma. *Urology* **64**: 839-844.

Tsunemoto K, Anzai M, Matsuoka T, Tokoro M, Shin SW, Amano T, Mitani T, Kato H, Hosoi Y, Saeki K et al. 2008. Cis-acting elements (E-box and NBE) in the promoter region of three maternal genes (Histone H1oo, Nucleoplasmin 2, and Zygote Arrest 1) are required for oocyte-specific gene expression in the mouse. *Mol Reprod Dev* **75**: 1104-1108.

Turatti E, da Costa Neves A, de Magalhaes MH, de Sousa SO. 2005. Assessment of c-Jun, c-Fos and cyclin D1 in premalignant and malignant oral lesions. *J Oral Sci* **47**: 71-76.

Ueda J, Harada A, Urahama T, Machida S, Maehara K, Hada M, Makino Y, Nogami J, Horikoshi N, Osakabe A et al. 2017. Testis-Specific Histone Variant H3t Gene Is Essential for Entry into Spermatogenesis. *Cell Rep* **18**: 593-600.

Ugai H, Dobbins GC, Wang M, Le LP, Matthews DA, Curiel DT. 2012. Adenoviral protein V promotes a process of viral assembly through nucleophosmin 1. *Virology* **432**: 283-295.

Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson A, Kampf C, Sjostedt E, Asplund A et al. 2015. Proteomics. Tissue-based map of the human proteome. *Science* **347**: 1260419.

Ulanet DB, Flavahan NA, Casciola-Rosen L, Rosen A. 2004. Selective cleavage of nucleolar autoantigen B23 by granzyme B in differentiated vascular smooth muscle cells: insights into the association of specific autoantibodies with distinct disease phenotypes. *Arthritis Rheum* **50**: 233-241.

Umekawa H, Sato K, Takemura M, Watanabe Y, Usui S, Takahashi T, Yoshida S, Olson MO, Furuichi Y. 2001. The carboxyl terminal sequence of nucleolar protein B23.1 is important in its DNA polymerase alpha-stimulatory activity. *J Biochem* **130**: 199-205.

Vancurova I, Paine TM, Lou W, Paine PL. 1995. Nucleoplasmin associates with and is phosphorylated by casein kinase II. *J Cell Sci* **108 ( Pt 2)**: 779-787.

Vanden Bush TJ, Bishop GA. 2011. CDK-mediated regulation of cell functions via c-Jun phosphorylation and AP-1 activation. *PLoS One* **6**: e19468.

Vardabasso C, Manganaro L, Lusic M, Marcello A, Giacca M. 2008. The histone chaperone protein Nucleosome Assembly Protein-1 (hNAP-1) binds HIV-1 Tat and promotes viral transcription. *Retrovirology* **5**: 8.

Vascotto C, Fantini D, Romanello M, Cesaratto L, Deganuto M, Leonardi A, Radicella JP, Kelley MR, D'Ambrosio C, Scaloni A et al. 2009. APE1/Ref-1 interacts with NPM1 within nucleoli and plays a role in the rRNA quality control process. *Mol Cell Biol* **29**: 1834-1854.

Vascotto C, Lirussi L, Poletto M, Tiribelli M, Damiani D, Fabbro D, Damante G, Demple B, Colombo E, Tell G. 2014. Functional regulation of the apurinic/apyrimidinic endonuclease 1 by nucleophosmin: impact on tumor biology. *Oncogene* **33**: 2876-2887.

Vassilev LT, Vu BT, Graves B, Carvajal D, Podlaski F, Filipovic Z, Kong N, Kammlott U, Lukacs C, Klein C et al. 2004. In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. *Science* **303**: 844-848.

Vavougios GD, Zarogiannis SG, Krogfelt KA, Gourgoulianis K, Mitsikostas DD, Hadjigeorgiou G. 2018. Novel candidate genes of the PARK7 interactome as mediators of apoptosis and acetylation in multiple sclerosis: An in silico analysis. *Mult Scler Relat Disord* **19**: 8-14.

Veija T, Koljonen V, Bohling T, Kero M, Knuutila S, Sarhadi VK. 2017. Aberrant expression of ALK and EZH2 in Merkel cell carcinoma. *BMC Cancer* **17**: 236.

Venkatesh S, Smolle M, Li H, Gogol MM, Saint M, Kumar S, Natarajan K, Workman JL. 2012. Set2 methylation of histone H3 lysine 36 suppresses histone exchange on transcribed genes. *Nature* **489**: 452-455.

Venkatesh S, Workman JL. 2015. Histone exchange, chromatin structure and the regulation of transcription. *Nat Rev Mol Cell Biol* **16**: 178-189.

Venters BJ, Pugh BF. 2013. Genomic organization of human transcription initiation complexes. *Nature* **502**: 53-58.

Verger A, Monte D, Villeret V. 2019. Twenty years of Mediator complex structural studies. *Biochem Soc Trans* **47**: 399-410.

Vermeulen M, Mulder KW, Denissov S, Pijnappel WW, van Schaik FM, Varier RA, Baltissen MP, Stunnenberg HG, Mann M, Timmers HT. 2007. Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* **131**: 58-69.

Vettese-Dadey M, Grant PA, Hebbes TR, Crane- Robinson C, Allis CD, Workman JL. 1996. Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nucleosomal DNA in vitro. *Embo j* **15**: 2508-2518.

Vishwamitra D, Curry CV, Shi P, Alkan S, Amin HM. 2015. SUMOylation Confers Posttranslational Stability on NPM-ALK Oncogenic Protein. *Neoplasia* **17**: 742-754.

Vitale AM, Calvert ME, Mallavarapu M, Yurttas P, Perlin J, Herr J, Coonrod S. 2007. Proteomic profiling of murine oocyte maturation. *Mol Reprod Dev* **74**: 608-616.

Vogiatzi F, Brandt DT, Schneikert J, Fuchs J, Grikscheit K, Wanzel M, Pavlakis E, Charles JP, Timofeev O, Nist A et al. 2016. Mutant p53 promotes tumor progression and metastasis by the endoplasmic reticulum UDPase ENTPD5. *Proc Natl Acad Sci U S A* **113**: E8433-e8442.

von Lindern M, Fornerod M, Soekarman N, van Baal S, Jaegle M, Hagemeijer A, Bootsma D, Grosveld G. 1992. Translocation t(6;9) in acute non-lymphocytic leukaemia results in the formation of a DEK-CAN fusion gene. *Baillieres Clin Haematol* **5**: 857-879.

Wakita K, Ohyanagi H, Yamamoto K, Tokuhisa T, Saitoh Y. 1992. Overexpression of c-Ki-ras and c-fos in human pancreatic carcinomas. *Int J Pancreatol* **11**: 43-47.

Waldmann T, Eckerich C, Baack M, Gruss C. 2002. The ubiquitous chromatin protein DEK alters the structure of DNA by introducing positive supercoils. *J Biol Chem* **277**: 24988-24994.

Waldmann T, Scholten I, Kappes F, Hu HG, Knippers R. 2004. The DEK protein--an abundant and ubiquitous constituent of mammalian chromatin. *Gene* **343**: 1-9.

Walter PP, Owen-Hughes TA, Cote J, Workman JL. 1995. Stimulation of transcription factor binding and histone displacement by nucleosome assembly protein 1 and nucleoplasmin requires disruption of the histone octamer. *Mol Cell Biol* **15**: 6178-6187.

Wang AH, Zare H, Mousavi K, Wang C, Moravec CE, Sirotkin HI, Ge K, Gutierrez-Cruz G, Sartorelli V. 2013. The histone chaperone Spt6 coordinates histone H3K27 demethylation and myogenesis. *Embo j* **32**: 1075-1086.

Wang D, Umekawa H, Olson MO. 1993. Expression and subcellular locations of two forms of nucleolar protein B23 in rat tissues and cells. *Cell Mol Biol Res* **39**: 33-42.

Wang HT, Tong X, Zhang ZX, Sun YY, Yan W, Xu ZM, Fu WN. 2019. MYCT1 represses apoptosis of laryngeal cancerous cells through the MAX/miR-181a/NPM1 pathway. *Febs j* **286**: 3892-3908.

Wang T, Liu M, Zheng SJ, Bian DD, Zhang JY, Yao J, Zheng QF, Shi AM, Li WH, Li L et al. 2017. Tumor-associated autoantibodies are useful biomarkers in immunodiagnosis of alpha-fetoprotein-negative hepatocellular carcinoma. *World J Gastroenterol* **23**: 3496-3504.

Wang W, Budhu A, Forgues M, Wang XW. 2005. Temporal and spatial control of nucleophosmin by the Ran-Crm1 complex in centrosome duplication. *Nat Cell Biol* **7**: 823-830.

Wang X, Spangler L, Dvir A. 2003. Promoter escape by RNA polymerase II. Downstream promoter DNA is required during multiple steps of early transcription. *J Biol Chem* **278**: 10250-10256.

Wang ZT, Chen ZJ, Jiang GM, Wu YM, Liu T, Yi YM, Zeng J, Du J, Wang HS. 2016. Histone deacetylase inhibitors suppress mutant p53 transcription via HDAC8/YY1 signals in triple negative breast cancer cells. *Cell Signal* **28**: 506-515.

Wanzel M, Russ AC, Kleine-Kohlbrecher D, Colombo E, Pelicci PG, Eilers M. 2008. A ribosomal protein L23-nucleophosmin circuit coordinates Mizl function with cell growth. *Nat Cell Biol* **10**: 1051-1061.

Warneboldt J, Haller F, Horstmann O, Danner BC, Fuzesi L, Doenecke D, Happel N. 2008. Histone H1x is highly expressed in human neuroendocrine cells and tumours. *BMC Cancer* **8**: 388.

Warren C, Matsui T, Karp JM, Onikubo T, Cahill S, Brenowitz M, Cowburn D, Girvin M, Shechter D. 2017. Dynamic intramolecular regulation of the histone chaperone nucleoplasmin controls histone binding and release. *Nat Commun* **8**: 2215.

Wasserman WW, Sandelin A. 2004. Applied bioinformatics for the identification of regulatory elements. *Nat Rev Genet* **5**: 276-287.

Weber CM, Henikoff S. 2014. Histone variants: dynamic punctuation in transcription. *Genes Dev* **28**: 672-682.

Weber JD, Jeffers JR, Rehg JE, Randle DH, Lozano G, Roussel MF, Sherr CJ, Zambetti GP. 2000. p53-independent functions of the p19(ARF) tumor suppressor. *Genes Dev* **14**: 2358-2365.

Weber JD, Taylor LJ, Roussel MF, Sherr CJ, Bar-Sagi D. 1999. Nucleolar Arf sequesters Mdm2 and activates p53. *Nat Cell Biol* **1**: 20-26.

Weinhold N, Moreaux J, Raab MS, Hose D, Hielscher T, Benner A, Meissner T, Ehrbrecht E, Brough M, Jauch A et al. 2010. NPM1 is overexpressed in hyperdiploid multiple myeloma due to a gain of chromosome 5 but is not delocalized to the cytoplasm. *Genes Chromosomes Cancer* **49**: 333-341.

Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, Schirghuber E, Tsutsumi S, Nagae G, Ishihara K, Mishiro T et al. 2008. Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* **451**: 796-801.

Westover KD, Bushnell DA, Kornberg RD. 2004. Structural basis of transcription: separation of RNA from DNA by RNA polymerase II. *Science* **303**: 1014-1016.

White RJ. 2008. RNA polymerases I and III, non-coding RNAs and cancer. *Trends Genet* **24**: 622-629.

Widom J, Klug A. 1985. Structure of the 300A chromatin filament: X-ray diffraction from oriented samples. *Cell* **43**: 207-213.

Wiedemann SM, Mildner SN, Bonisch C, Israel L, Maiser A, Matheisl S, Straub T, Merkl R, Leonhardt H, Kremmer E et al. 2010. Identification and characterization of two novel primate-specific histone H3 variants, H3.X and H3.Y. *J Cell Biol* **190**: 777-791.

Wiesmann N, Gieringer R, Grus F, Brieger J. 2019. Phosphoproteome Profiling Reveals Multifunctional Protein NPM1 as part of the Irradiation Response of Tumor Cells. *Transl Oncol* **12**: 308-319.

Wilczek C, Chitta R, Woo E, Shabanowitz J, Chait BT, Hunt DF, Shechter D. 2011. Protein arginine methyltransferase Prmt5-Mep50 methylates histones H2A and H4 and the histone chaperone nucleoplasmin in Xenopus laevis eggs. *J Biol Chem* **286**: 42221-42231.

Wiles ET, Selker EU. 2017. H3K27 methylation: a promiscuous repressive chromatin mark. *Curr Opin Genet Dev* **43**: 31-37.

Willems E, Dedobbeleer M, Digregorio M, Lombard A, Lumapat PN, Rogister B. 2018. The functional diversity of Aurora kinases: a comprehensive review. *Cell Div* **13**: 7.

Williams SK, Truong D, Tyler JK. 2008. Acetylation in the globular core of histone H3 on lysine-56 promotes chromatin disassembly during transcriptional activation. *Proc Natl Acad Sci U S A* **105**: 9000-9005.

Williams SP, Athey BD, Muglia LJ, Schappe RS, Gough AH, Langmore JP. 1986. Chromatin fibers are left-handed double helices with diameter and mass per unit length that depend on linker length. *Biophys J* **49**: 233-248.

Wind M, Reines D. 2000. Transcription elongation factor SII. *Bioessays* **22**: 327-336.

Winkler C, Steingrube DS, Altermann W, Schlaf G, Max D, Kewitz S, Emmer A, Kornhuber M, Banning-Eichenseer U, Staege MS. 2012. Hodgkin's lymphoma RNA-transfected dendritic cells induce cancer/testis antigen-specific immune responses. *Cancer Immunol Immunother* **61**: 1769-1779.

Winkler DD, Luger K. 2011. The histone chaperone FACT: structural insights and mechanisms for nucleosome reorganization. *J Biol Chem* **286**: 18369-18374.

Wisdom R, Johnson RS, Moore C. 1999. c-Jun regulates cell cycle progression and apoptosis by distinct mechanisms. *Embo j* **18**: 188-197.

Wise-Draper TM, Mintz-Cole RA, Morris TA, Simpson DS, Wikenheiser-Brokamp KA, Currier MA, Cripe TP, Grosveld GC, Wells SI. 2009. Overexpression of the cellular DEK protein promotes epithelial transformation in vitro and in vivo. *Cancer Res* **69**: 1792-1799.

Wong JC, Hasan MR, Rahman M, Yu AC, Chan SK, Schaeffer DF, Kennecke HF, Lim HJ, Owen D, Tai IT. 2013. Nucleophosmin 1, upregulated in adenomas and cancers of the colon, inhibits p53-mediated cellular senescence. *Int J Cancer* **133**: 1567-1577.

Wood K, Tellier M, Murphy S. 2018. DOT1L and H3K79 Methylation in Transcription and Genomic Stability. *Biomolecules* **8**.

Wu M, Wei W, Chen J, Cong R, Shi T, Bouvet P, Li J, Wong J, Du JX. 2017. Acidic domains differentially read histone H3 lysine 4 methylation status and are widely present in chromatin-associated proteins. *Sci China Life Sci* **60**: 138-151.

Wu MH, Chang JH, Chou CC, Yung BY. 2002a. Involvement of nucleophosmin/B23 in the response of HeLa cells to UV irradiation. *Int J Cancer* **97**: 297-305.

Wu MH, Chang JH, Yung BY. 2002b. Resistance to UV-induced cell-killing in nucleophosmin/B23 over-expressed NIH 3T3 fibroblasts: enhancement of DNA repair and up-regulation of PCNA in association with nucleophosmin/B23 over-expression. *Carcinogenesis* **23**: 93-100.

Wu MH, Yung BY. 2002. UV stimulation of nucleophosmin/B23 expression is an immediate-early gene response induced by damaged DNA. *J Biol Chem* **277**: 48234-48240.

Wu N, Li CJ, Gui JF. 2009. Molecular characterization and functional commonality of nucleophosmin/nucleoplasmin in two cyprinid fish. *Biochem Genet* **47**: 749-762.

Xi R, Lee S, Xia Y, Kim TM, Park PJ. 2016. Copy number analysis of whole-genome data using BIC-seq2 and its application to detection of cancer susceptibility variants. *Nucleic Acids Res* **44**: 6274-6286.

Xia X, Liu S, Xiao Z, Zhu F, Song NY, Zhou M, Liu B, Shen J, Nagashima K, Veenstra TD et al. 2013. An IKKalpha-nucleophosmin axis utilizes inflammatory signaling to promote genome integrity. *Cell Rep* **5**: 1243-1255.

Xie H, Hu H, Chang M, Huang D, Gu X, Xiong X, Xiong R, Hu L, Li G. 2016. Identification of chaperones in a MPP(+)-induced and ATRA/TPA-differentiated SH-SY5Y cell PD model. *Am J Transl Res* **8**: 5659-5671.

Xin H, Takahata S, Blanksma M, McCullough L, Stillman DJ, Formosa T. 2009. yFACT induces global accessibility of nucleosomal DNA without H2A-H2B displacement. *Mol Cell* **35**: 365-376.

Xu R, Yu S, Zhu D, Huang X, Xu Y, Lao Y, Tian Y, Zhang J, Tang Z, Zhang Z et al. 2019. hCINAP regulates the DNA-damage response and mediates the resistance of acute myelocytic leukemia cells to therapy. *Nat Commun* **10**: 3812.

Xue YM, Kowalska AK, Grabowska K, Przybyt K, Cichewicz MA, Del Rosario BC, Pemberton LF. 2013. Histone chaperones Nap1 and Vps75 regulate histone acetylation during transcription elongation. *Mol Cell Biol* **33**: 1645-1656.

Yamaguchi Y, Shibata H, Handa H. 2013. Transcription elongation factors DSIF and NELF: promoter-proximal pausing and beyond. *Biochim Biophys Acta* **1829**: 98-104.

Yan W, Ma L, Burns KH, Matzuk MM. 2003. HILS1 is a spermatid-specific linker histone H1-like protein implicated in chromatin remodeling during mammalian spermiogenesis. *Proc Natl Acad Sci U S A* **100**: 10546-10551.

Yang C, Maiguel DA, Carrier F. 2002. Identification of nucleolin and nucleophosmin as genotoxic stress-responsive RNA-binding proteins. *Nucleic Acids Res* **30**: 2251-2260.

Yang K, Wang M, Zhao Y, Sun X, Yang Y, Li X, Zhou A, Chu H, Zhou H, Xu J et al. 2016a. A redox mechanism underlying nucleolar stress sensing by nucleophosmin. *Nat Commun* **7**: 13599.

Yang MR, Zhang Y, Wu XX, Chen W. 2016b. Critical genes of hepatocellular carcinoma revealed by network and module analysis of RNA-seq data. *Eur Rev Med Pharmacol Sci* **20**: 4248-4256.

Yang YX, Hu HD, Zhang DZ, Ren H. 2007. Identification of proteins responsible for the development of adriamycin resistance in human gastric cancer cells using comparative proteomics analysis. *J Biochem Mol Biol* **40**: 853-860.

Yeh CW, Huang SS, Lee RP, Yung BY. 2006. Ras-dependent recruitment of c-Myc for transcriptional activation of nucleophosmin/B23 in highly malignant U1 bladder cancer cells. *Mol Pharmacol* **70**: 1443-1453.

Yin JW, Wang G. 2014. The Mediator complex: a master coordinator of transcription and cell lineage development. *Development* **141**: 977-987.

Yokota T, Kouno J, Adachi K, Takahashi H, Teramoto A, Matsumoto K, Sugisaki Y, Onda M, Tsunoda T. 2006. Identification of histological markers for malignant glioma by genome-wide expression analysis: dynein, alpha-PIX and sorcin. *Acta Neuropathol* **111**: 29-38.

Yonaha M, Proudfoot NJ. 2000. Transcriptional termination and coupled polyadenylation in vitro. *Embo j* **19**: 3770-3777.

Yu Y, Maggi LB, Jr., Brady SN, Apicelli AJ, Dai MS, Lu H, Weber JD. 2006. Nucleophosmin is essential for ribosomal protein L5 nuclear export. *Mol Cell Biol* **26**: 3798-3809.

Yuan GC, Liu YJ, Dion MF, Slack MD, Wu LF, Altschuler SJ, Rando OJ. 2005. Genome-scale identification of nucleosome positions in S. cerevisiae. *Science* **309**: 626-630.

Yudkovsky N, Ranish JA, Hahn S. 2000. A transcription reinitiation intermediate that is stabilized by activator. *Nature* **408**: 225-229.

Yue F Cheng Y Breschi A Vierstra J Wu W Ryba T Sandstrom R Ma Z Davis C Pope BD et al. 2014. A comparative encyclopedia of DNA elements in the mouse genome. *Nature* **515**: 355-364.

Yuen MF, Wu PC, Lai VC, Lau JY, Lai CL. 2001. Expression of c-Myc, c-Fos, and c-jun in hepatocellular carcinoma. *Cancer* **91**: 106-112.

Yun JP, Chew EC, Liew CT, Chan JY, Jin ML, Ding MX, Fai YH, Li HK, Liang XM, Wu QL. 2003. Nucleophosmin/B23 is a proliferate shuttle protein associated with nuclear matrix. *J Cell Biochem* **90**: 1140-1148.

Yun JP, Miao J, Chen GG, Tian QH, Zhang CQ, Xiang J, Fu J, Lai PB. 2007. Increased expression of nucleophosmin/B23 in hepatocellular carcinoma and correlation with clinicopathological parameters. *Br J Cancer* **96**: 477-484.

Yung BY. 2004. c-Myc-mediated expression of nucleophosmin/B23 decreases during retinoic acid-induced differentiation of human leukemia HL-60 cells. *FEBS Lett* **578**: 211-216.

Yung BY, Busch H, Chan PK. 1985. Translocation of nucleolar phosphoprotein B23 (37 kDa/pI 5.1) induced by selective inhibitors of ribosome synthesis. *Biochim Biophys Acta* **826**: 167-173.

Yusufzai TM, Tagami H, Nakatani Y, Felsenfeld G. 2004. CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species. *Mol Cell* **13**: 291-298.

Zatsepina OV, Rousselet A, Chan PK, Olson MO, Jordan EG, Bornens M. 1999. The nucleolar phosphoprotein B23 redistributes in part to the spindle poles during mitosis. *J Cell Sci* **112 ( Pt 4)**: 455-466.

Zatsepina OV, Todorov IT, Philipova RN, Krachmarov CP, Trendelenburg MF, Jordan EG. 1997. Cell cycle-dependent translocations of a major nucleolar phosphoprotein, B23, and some characteristics of its variants. *Eur J Cell Biol* **73**: 58-70.

Zeitlinger J, Stark A, Kellis M, Hong JW, Nechaev S, Adelman K, Levine M, Young RA. 2007. RNA polymerase stalling at developmental control genes in the Drosophila melanogaster embryo. *Nat Genet* **39**: 1512-1516.

Zeller KI, Haggerty TJ, Barrett JF, Guo Q, Wonsey DR, Dang CV. 2001. Characterization of nucleophosmin (B23) as a Myc target by scanning chromatin immunoprecipitation. *J Biol Chem* **276**: 48285-48291.

Zeng, Xiao Y, Zhu J, Peng C, Liang W, Lin H. 2019. Knockdown of nucleophosmin 1 suppresses proliferation of triple-negative breast cancer cells through activating CDH1/Skp2/p27kip1 pathway. *Cancer Manag Res* **11**: 143-156.

Zhang G, Campbell EA, Minakhin L, Richter C, Severinov K, Darst SA. 1999. Crystal structure of Thermus aquaticus core RNA polymerase at 3.3 A resolution. *Cell* **98**: 811-824.

Zhang H, Shi X, Paddon H, Hampong M, Dai W, Pelech S. 2004. B23/nucleophosmin serine 4 phosphorylation mediates mitotic functions of polo-like kinase 1. *J Biol Chem* **279**: 35726-35734.

Zhang MJ, Ding YL, Xu CW, Yang Y, Lian WX, Zhan YQ, Li W, Xu WX, Yu M, Ge CH et al. 2012. Erythroid differentiation-associated gene interacts with NPM1 (nucleophosmin/B23) and increases its protein stability, resisting cell apoptosis. *Febs j* **279**: 2848-2862.

Zhang Q, Wang HY, Liu X, Wasik MA. 2007. STAT5A is epigenetically silenced by the tyrosine kinase NPM1-ALK and acts as a tumor suppressor by reciprocally inhibiting NPM1-ALK expression. *Nat Med* **13**: 1341-1348.

Zhang WJ, Wu XN, Shi TT, Xu HT, Yi J, Shen HF, Huang MF, Shu XY, Wang FF, Peng BL et al. 2016. Regulation of Transcription Factor Yin Yang 1 by SET7/9-mediated Lysine Methylation. *Sci Rep* **6**: 21718.

Zhang X, Bernatavichute YV, Cokus S, Pellegrini M, Jacobsen SE. 2009. Genome-wide analysis of mono-, di- and trimethylation of histone H3 lysine 4 in Arabidopsis thaliana. *Genome Biol* **10**: R62.

Zhang Y, Ku WL, Liu S, Cui K, Jin W, Tang Q, Lu W, Ni B, Zhao K. 2017. Genome-wide identification of histone H2A and histone variant H2A.Z-interacting proteins by bPPI-seq. *Cell Res* **27**: 1258-1274.

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W et al. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**: R137.

Zhao H, Chen ZJ, Qin Y, Shi Y, Wang S, Choi Y, Simpson JL, Rajkovic A. 2008. Transcription factor FIGLA is mutated in patients with premature ovarian failure. *Am J Hum Genet* **82**: 1342-1348.

Zhao X, Herr W. 2002. A regulated two-step mechanism of TBP binding to DNA: a solvent-exposed surface of TBP inhibits TATA box recognition. *Cell* **108**: 615-627.

Zhao Y, Garcia BA. 2015. Comprehensive Catalog of Currently Documented Histone Modifications. *Cold Spring Harb Perspect Biol* **7**: a025064.

Zheng S, Crickard JB, Srikanth A, Reese JC. 2014. A highly conserved region within H2B is important for FACT to act on nucleosomes. *Mol Cell Biol* **34**: 303-314.

Zhou BR, Jiang J, Feng H, Ghirlando R, Xiao TS, Bai Y. 2015. Structural Mechanisms of Nucleosome Recognition by Linker Histones. *Mol Cell* **59**: 628-638.

Zhou F, Muller-Tidow C. 2019. NPM1 functions in epitranscriptomics. *Nat Genet*.

Zhou G, Liu Z, Myers JN. 2016. TP53 Mutations in Head and Neck Squamous Cell Carcinoma and Their Impact on Disease Progression and Treatment Response. *J Cell Biochem* **117**: 2682-2692.

Zhou Q, Gedrich RW, Engel DA. 1995. Transcriptional repression of the c-fos gene by YY1 is mediated by a direct interaction with ATF/CREB. *J Virol* **69**: 4323-4330.

Zhou VW, Goren A, Bernstein BE. 2011. Charting histone modifications and the functional organization of mammalian genomes. *Nat Rev Genet* **12**: 7-18.

Zhou Y, Shen J, Xia L, Wang Y. 2014. Estrogen mediated expression of nucleophosmin 1 in human endometrial carcinoma clinical stages through estrogen receptor-alpha signaling. *Cancer Cell Int* **14**: 540.

Zhou YX, Wang C, Mao LW, Wang YL, Xia LQ, Zhao W, Shen J, Chen J. 2018. Long noncoding RNA HOTAIR mediates the estrogen-induced metastasis of endometrial cancer cells via the miR-646/NPM1 axis. *Am J Physiol Cell Physiol* **314**: C690-c701.

Zhu W, Olson SY, Garban H. 2011. Transcription regulator Yin-yang 1: from silence to cancer. *Crit Rev Oncog* **16**: 227-238.

Zhu Y, Shi M, Chen H, Gu J, Zhang J, Shen B, Deng X, Xie J, Zhan X, Peng C. 2015. NPM1 activates metabolic changes by inhibiting FBP1 while promoting the tumorigenicity of pancreatic cancer cells. *Oncotarget* **6**: 21443-21451.

Zirwes RF, Schmidt-Zachmann MS, Franke WW. 1997. Identification of a small, very acidic constitutive nucleolar protein (NO29) as a member of the nucleoplasmin family. *Proc Natl Acad Sci U S A* **94**: 11387-11392.

Ziv O, Zeisel A, Mirlas-Neisberg N, Swain U, Nevo R, Ben-Chetrit N, Martelli MP, Rossi R, Schiesser S, Canman CE et al. 2014. Identification of novel DNA-damage tolerance genes reveals regulation of translesion DNA synthesis by nucleophosmin. *Nat Commun* **5**: 5437.

Zlatanova J, Doenecke D. 1994. Histone H1 zero: a major player in cell differentiation? *Faseb j* **8**: 1260-1268.

Zou Y, Wu J, Giannone RJ, Boucher L, Du H, Huang Y, Johnson DK, Liu Y, Wang Y. 2008. Nucleophosmin/B23 negatively regulates GCN5-dependent histone acetylation and transactivation. *J Biol Chem* **283**: 5728-5737.