Mrhl long non-coding RNA regulates Wnt mediated Sox8 gene expression during spermatogenesis through promoter interaction

A thesis submitted for partial fulfillment of degree of
MASTER OF SCIENCE (Biological Sciences)
as part of Integrated PhD program
by

Shubhangini Kataruka

Molecular Biology and Genetics Unit
Jawaharlal Nehru Centre for Advanced Scientific Research
Jakkur, Bangalore-560064
March 2016

DECLARATION

I hereby declare that the work described in this thesis entitled **'*Mrhl* long non-coding RNA regulates Wnt mediated Sox8 gene expression during spermatogenesis through promoter interaction'** is the result of investigations carried out by myself under the guidance of **Prof MRS Rao** at Chromatin Biology Laboratory, Molecular Biology and Genetics Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore-560064, India. This work has not been submitted elsewhere for the award of any other degree.

In keeping with the general practice of reporting scientific observations, due acknowledgements have been made wherever the work described has been based on findings of other investigators. Any omission, which might have occurred by oversight or misjudgment, is regretted.

31st March, 2016 Shubhangini Kataruka

Jawaharlal Nehru Center for Advanced Scientific Research
Jakkur Campus, Jakkur Post
Bengaluru 560 064, INDIA
Office Tel: +91 (80) 2208 2754
Office Fax: +91 (80) 2208 2758
+91 (80) 2208 2468
Email: mrsrao@jncasr.ac.in

Prof MRS Rao,
Honororary Professor and SERB Distinguished Fellow
Molecular Biology and Genetics Unit

CERTIFICATE

This is to certify that the work described in this thesis entitled **'*Mrhl* long non-coding RNA regulates Wnt mediated Sox8 gene expression during spermatogenesis through promoter interaction'** is the result of investigations carried out by Ms. Shubhangini Kataruka at Chromatin Biology laboratory, Molecular Biology and Genetics Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore, under my supervision.

31st March, 2016 Prof MRS Rao

Acknowledgement

First and foremost, I would like to thank my mentor and my thesis supervisor Prof MRS Rao. I am grateful to him for many things- giving me an opportunity to work in the lab, for encouraging me to take risks and trying new techniques, for promoting open intellectual scientific discussion and for teaching me the importance of critical assessment. There are many things that I have learnt from him but the one thing that I wish to imbibe is his passion for science and I hope to prove myself worthy of being his student.

I would like to acknowledge Vijay for his constant guidance and contribution in the Sox8 project. I have had an excellent time in CBL and I would like to thank all the past and present members of CBL for the fun-filled atmosphere- Nikhil, Vijay, Satya, Debosree, Monalisa, Aditya, Roshan, Shalini, Arun, Dhanur, Shrinivas, Raktim, Bhavana, Vandhana and Neha. I have found great friends in many of them and for that I will forever be grateful. I especially want to acknowledge Debosree and Aditya for they also helped me with multiple experiments. I am also thankful to Muniraju for his assistance.

I express my gratitude to Ramesh sir for his training has been immensely helpful during my research. I also want to thank my batchmates- Veena, Priya, Pallabi and Siddharth for we have been in the same boat for three years and it has been a joyous ride with them. I want to express my gratitude to Prof. Maneesha Inamdar for giving me the opportunity to work in her lab during my second rotation and to all the members of VBL for their help especially to Divyesh, Jaspar, Anudeep and Diana. I thank Suma BS for confocal facility and Dr. Prakash for animal facility.

I want to acknowledge my high school teacher- Miss Rosy and my Bachelors professors- Prof Wajeed and Prof Beatrice for it is they who motivated me to pursue science and for making me believe that I could do it.

I consider myself lucky to be surrounded by loyal and crazy friends – Aditi, Anna, Palak, Vinay, Aakash, Kavya, Ale. I am grateful to Nikhil for his support.

Everything that I am and I achieve is because of the efforts of my parents and my brother and I am thankful to them for it is they who taught me to dream and to let my choices not my abilities define me.

Contents

CHAPTER 1
INTRODUCTION

1.1 Spermatogenesis:

The formation of the germ and somatic cell lineages in the testis occur independently. In the pre-gastrulating mouse embryo, germ line allocation occurs in the epiblast-derived extra embryonic mesoderm, to cells residing at the posterior end of the primitive streak. At around 6.0 dpc bone morphogenetic signaling (BMP) designates a region of the epiblast to form Primordial germ cells (PGC) rewiring these cells to suppress somatic cell gene expression and promote pluripotent gene expression such as *Pou5f1 (Oct 4), Nanog, Sox2* [1]. At around 7.5 dpc after re-expression of *Dppa3* (*Stella*) these reprogrammed cells emerge to occupy the proximal epiblast. This cohort of PGCs rapidly proliferate and migrate for 2 days through the posterior primitive streak to eventually colonize the genital ridges, the future gonads [2].

Migratory PGCs require cell-to-cell contact and express numerous adhesion molecules. *E-cadherin* is one of the major players in this process and is aided by *PECAM-1*, which is expressed on the plasma membrane from 9.0 dpc and remains strong in gonadal GCs till 13.0 dpc. PGCs enter the genital ridges between 10.0 and 11.0 dpc. From this time onwards they are no longer locomotory and are sometimes termed gonocytes [3]. Once the germ cells are in the genital ridge they start to express new germ cell specific genes such as *Mvh, Gcna1*. Subsequently *Tnap* and other genes such as *Ssea1* are down regulated [4]. The PGCs adopt a large, rounded morphology upon entry into the genital ridges and after epigenetic reprogramming become susceptible to differentiation cues from the surrounding somatic tissue. Apoptosis or programmed cell death occurs at multiple stages throughout PGC development, both during their migration

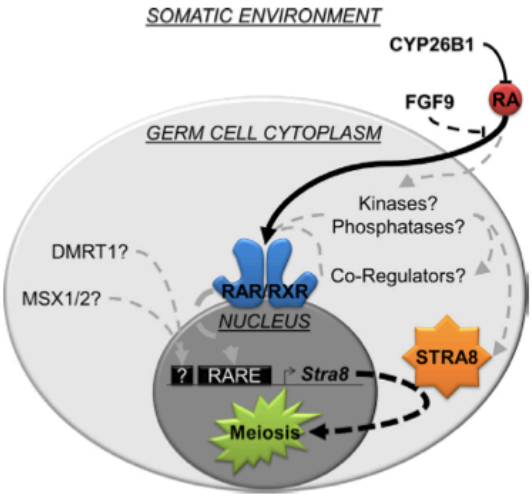


Fig1: Schematic illustration of molecular genetic interactions leading to meiosis in germ cells. *Adapted from: Control of mammalian germ cell entry into meiosis, Chun-Wei Feng et al, 2013. Molecular and Cellular Endocrinology*

and also after colonization of genital ridges. It primarily functions to prevent the occurrence of gonadal tumors and to ensure the integrity of the germ line. In male GCs (Germ cells), an initial wave of apoptosis occurs between 13.0 dpc and 17.0 dpc and a second wave occurs at around the time of birth. Apoptosis of gonadal GCs is a measure to ensure the quality of the genetic information being passed on [5].

At approximately 12.5 dpc when the GCs are already cloistered inside nascent testis cords proliferation ceases, and the GCs undergo mitotic arrest exiting the cell cycle at the G1/G0 transition. Only after birth at 5-10 dpp the prospermatogonia re-enter the cell cycle and resume mitosis with entry into prophase of meiosis 1 occurring at puberty, which is around 7-8 dpp in mice [6]. Some of the prospermatogonia can initially directly form differentiating spermatogonia also referred to as the first wave of spermatogenesis [7]. The first round of spermatogonia develop from a unique *Neurogenin 3* (NGN3) negative pool of prospermatogonia while the NGN3 positive cells remain undifferentiated and comprise the pool of stem cells. The NGN3 negative A1 spermatogonia differentiate into A3 and A4 spermatogonia and finally into B spermatogonia by 5 dpp. Mammalian spermatogenesis is an asynchronous process thus guaranteeing a lifelong production of sperm [8].

The transition to A1 spermatogonia is controlled by Retinoic acid (RA) and signifies the commitment to meiosis [9]. RA is synthesized by the Sertoli cells and acts in a paracrine manner on germ cells for the first round of spermatogenesis. Entry into meiosis occurs cell-autonomously and the timing of this event is also cell autonomous, involving an intrinsic timing mechanism. RA is essential for germ cells of the postnatal testis, not just for successful spermatogenesis but also for entry into meiosis [10]. In mice the first sign of germ cells entry into meiosis is expression of gene stimulated by Retinoic acid 8 (*Stra-8*), which encodes a protein required for pre-meiotic DNA replication and subsequent entry into prophase 1 [11]. Meiotic entry is then marked by up regulation of genes such as synaptonemal complex protein 3 (*Sycp3)*, meiosis-specific sporulation protein (*Spo11*) and meiosis-specific homologous recombinase- dosage suppressor of mck1 homolog, (*Dmc1*). *Sycp3* encodes a component of the synaptonemal complex, a meiosis specific structure

required for synapses of homologous chromosomes during leptotene. Dmc1, which is a meiosis specific recombinase, is required at zygotene for double-stranded break repair. The formation of DSBs comes with the modification of H2A histone family member X (H2AX) to the phosphorylated form, γH2AX, at the site of DNA breaks. Entry into meiosis is also marked by down- regulation of the pluripotency marker-*Oct4* [12]. The timing of entry into meiosis is tightly regulated for germ cells. The precise control is maintained by *Cyp26b1*, which acts as a meiosis- inhibiting factor (MIF) in the developing testis. It does so by bringing about active degradation of RA which is sufficient to block meiosis [13]. Interestingly, *Nanos 2* maintains the suppression of meiosis in male GCs after *Cyp26b1* expression decreases at 14.0 dpc, by preventing the expression of *Stra-8*. Fibroblast growth factor 9 (Fgf-9) is a secreted MPS (meiosis preventing factor). The established MIS (RA) and MPS (FGF9) are extrinsic factors that influence germ cell's decision to undertake meiosis [14]. However, intrinsic factors within the germ cells are required to answer to these cues. A candidate for such factor is the gene deleted in azoospermia-like (*Dazl*) and Double sex-related transcription factor (*DMRT1*) [15]. DMRT1 is a transcription factor expressed by both germ and Sertoli cells. It has been shown that the germ cell-specific loss of DMRT1 in adult spermatogenesis results in the precocious ability of spermatogonia to induce *Stra8* and enter meiosis [16]. The *Dazl* gene is another important intrinsic factor in mammalian competency for meiosis. There is evidence to suggest that Dazl enables germ cells to respond to RA and opposes the activity of DMRT1.

Wnt signaling also plays a non-dispensable role in differentiation of spermatogonia to spermatocytes. Beta-catenin signaling is activated in differentiating male germ cells but not in most spermatogonia. Disruption of Wnt signaling by conditional beta-catenin knockout leads to loss of post mitotic germ cells, indicating that beta–catenin mediated Wnt signaling is vital in the later spermatogenic stages [17].

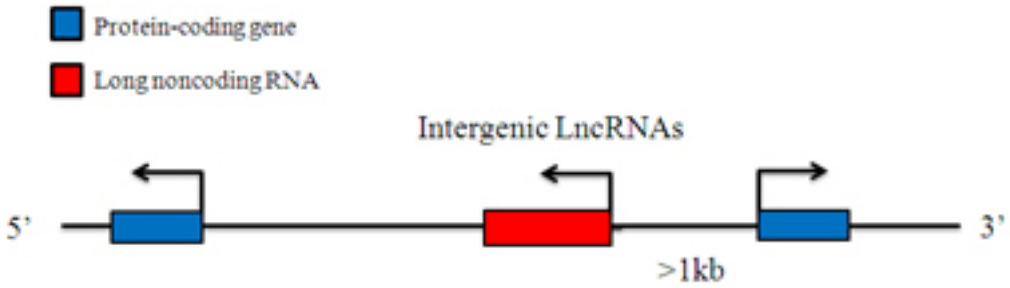
1.2 Lnc RNA:

RNAs have always been considered to be just the intermediates, messengers in the essential pathway-central dogma, which defines the biology of any living being. However, over the past decade many versatile functions of RNA have been discovered in which the major players are the regulatory non-coding RNAs. The mammalian genome is pervasively transcribed but only 2% of the genome codes for translated genes indicating that the majority of the genome codes for non-translated transcripts. It has been suggested that the expansion of the proportion of the noncoding genome contributes to the complexity of the organism as we climb up the evolutionary ladder. The term ncRNA includes a wide variety of RNA species, including, but not restricted to the following: 'classic' ncRNAs that regulate mRNA splicing and translation (i.e., small nuclear RNAs, rRNAs, and tRNAs); short 21– 28 nucleotide regulatory RNAs involved in transcriptional and post-transcriptional gene silencing (i.e., miRNA, small interfering RNAs, Piwi-associated RNAs); and a large number of newly identified and less understood long non-coding RNAs (lncRNAs).

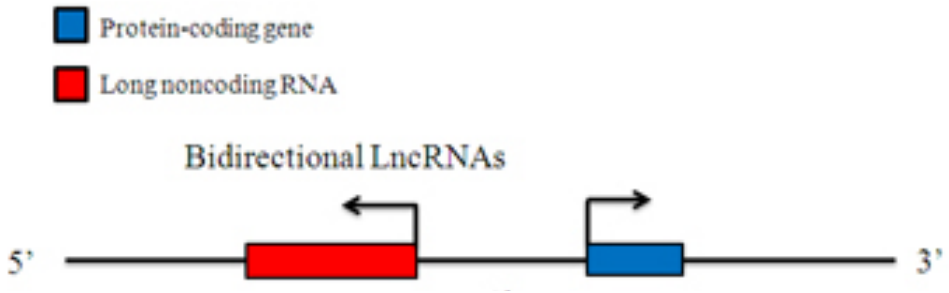
LncRNAs are currently defined as RNA transcripts longer than 200 nucleotides that apparently do not function by encoding proteins [18]. Most of the lncRNAs are transcribed by RNA Polymerase II and are polyadenylated. LncRNAs can be classified as follows-

LncRNA Classification:

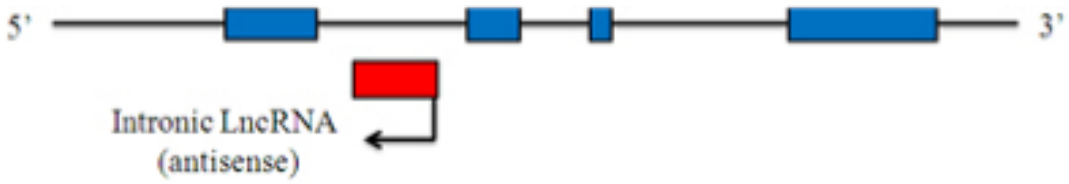
I. Intergenic LncRNAs- Intergenic lncRNAs such as *HOTAIR, MALAT1* are long non-coding RNAs, which locate between annotated protein-coding genes and are at least 1 kb away from the nearest protein-coding genes.



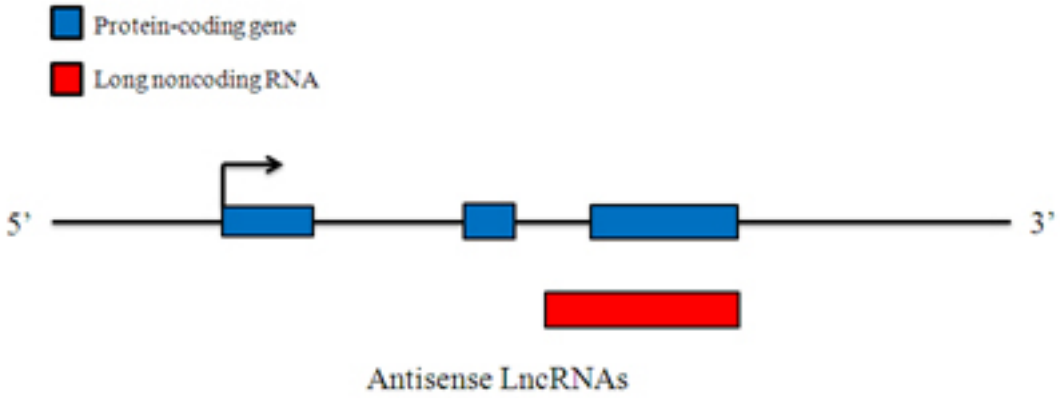
II. Bidirectional LncRNAs- A Bidirectional LncRNA is oriented head to head with a protein-coding gene within 1kb. A Bidirectional LncRNA transcript exhibits a similar expression pattern to its protein-coding counterpart, which suggests that they may be subject to share regulatory pressures. Examples- FMR4, Sox8ot.



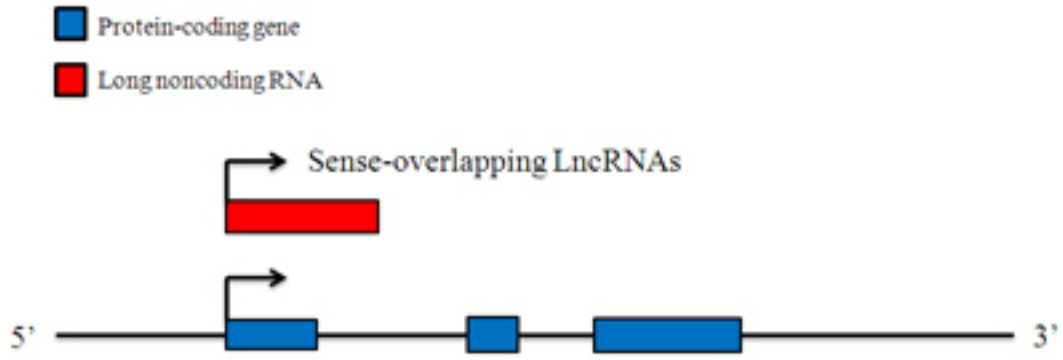
III. Intronic LncRNAs- Intronic LncRNAs such as *mrhl* are RNA molecules that reside within the intron of an annotated coding gene in either sense or antisense orientation.



IV. Antisense LncRNAs- Antisense LncRNAs such as Airn, Kcnq1ot1 are RNA molecules that are transcribed from the antisense strand and overlap in part with well-defined spliced sense or intron less sense RNAs



V. Sense-overlapping LncRNAs-These LncRNAs can be considered transcript variants of protein-coding mRNAs, as they overlap with a known annotated gene on the same genomic strand for example FMR5, Sox2ot



In addition to containing specific sequence information, lncRNAs also possess structural plasticity. They can directly interact with both DNA and RNA through base pairing, and

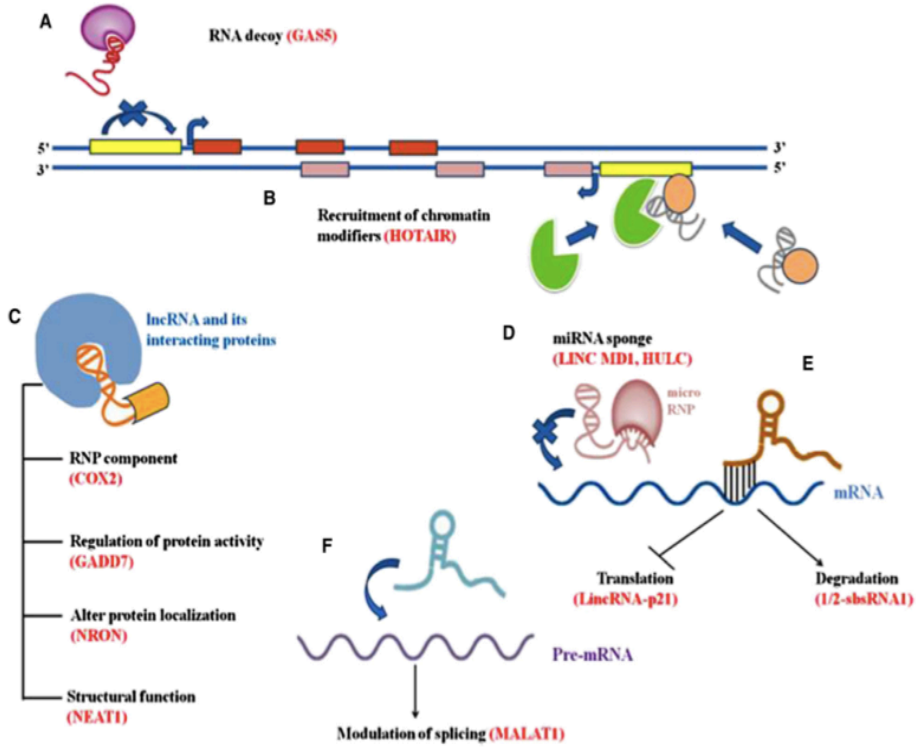


Fig 2: Diverse mechanisms of lncRNA function.
Adapted from: Long noncoding RNAs in development and cancer: potential biomarkers and therapeutic targets, Roshan Fatima et al, 2015, Molecular and Cellular Therapies

can bind to protein partners through specific structural motifs. This unique feature enables lncRNAs to regulate gene expression and function via diverse mechanisms [19]. Biochemical and molecular studies have suggested that lncRNAs can act as an important interface between chromatin modification complexes and the genome — by acting as decoys (GAS5), by recruiting chromatin modifiers (HOTAIR), by modulating splicing (MALAT1) or by altering protein localization (NRON) [20].

Several lncRNAs can directly interact with and recruit chromatin complexes that mediate repressive or activating chromatin modifications, thereby altering the chromatin state and gene expression. Other lncRNAs appear to localize to nuclear territories and the cytoplasm for other chromatin templated processes and post-transcriptional regulation respectively. Regulation of transcription is considered to be interplay of transcription factors (TFs) and chromatin modifying factors acting on enhancer and promoter sequences to facilitate the assembly of the transcription machinery at gene promoters. Transcriptional regulation by lncRNAs could work either in *cis* or in *trans*, and could negatively or positively control protein coding gene expression. LncRNAs work in *cis* when their effects are restricted to the chromosome from which they are transcribed, and work in *trans* when they affect genes on other chromosomes [21].

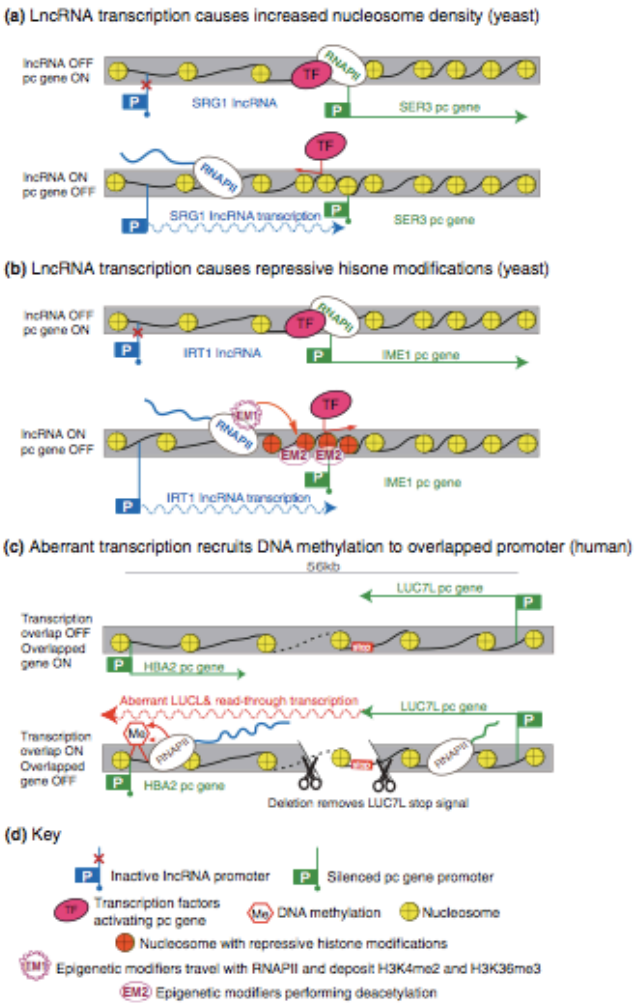


Fig 3: LncRNA transcription interference-mediated silencing by chromatin changes.
Adapted from: Gene regulation by the act of long non-coding RNA transcription*, Kornienko et al, 2013, BMC Biology*

DNA in the nucleus is organized into chromatin with the organizational scaffold consisting

of nucleosomes, which can be densely packed, interfering with protein-DNA interactions, or relaxed, facilitating these interactions. The transcription process, which generates single-stranded DNA as RNAPII progresses along a gene locus, can directly affect nucleosome positioning. Thus, lncRNA transcription could cause transcriptional inhibition by depositing nucleosomes in a manner unfavorable for TF binding on promoters or enhancers [22].

LncRNAs are also known to regulate various important biological processes like genomic imprinting (*Air* and *Kcnq1ot1*), dosage compensation (*Xist* and *roX*), cell differentiation, development (*Fendrr, Bvht, Miat, Hotair,* etc.), pluripotency (*Evx1as* and *Hoxb5/6as*), nuclear architecture (*Neat1*), etc. Since lncRNAs play such a diverse role in development and cellular differentiation, it is not surprising that they have employed multiple molecular mechanisms to exert their biological functions. These mechanisms include translational

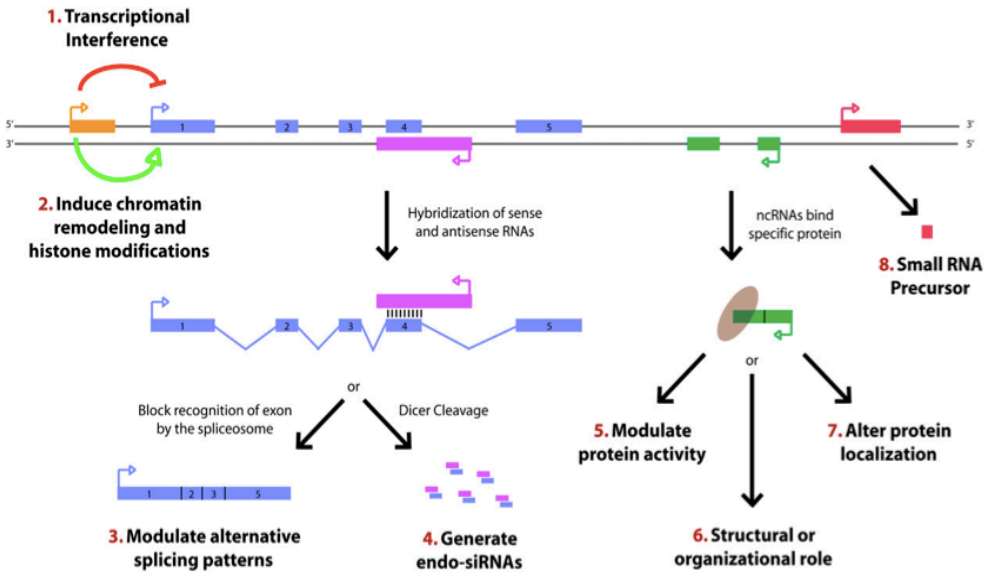


Fig 4: Paradigms of how lncRNA function
Adapted from: Long noncoding RNAs: functional surprises from the RNA world, Wilusz et al, 2009, Genes and Development

inhibition (lincRNA-p21), mRNA degradation (*1/2-sbs* RNAs), RNA decoys (*Gas5*), facilitating recruitment of chromatin modifiers (*Mistral*, *HOTTIP*), regulation of protein activity (*Evi2* and *Lethe*), regulating the availability of miRNAs by sponging mechanism (linc-MD1), etc. In addition to their role in cellular differentiation and development,

expression of lncRNAs are also perturbed in several diseases and more particularly in cancer suggesting that regulation of their expression is an important aspect of cellular homeostasis [18].

LncRNAs are often localized to the nucleus, preferentially in the chromatin fraction. Expression of lncRNAs is more tightly regulated spatiotemporally than protein-coding genes suggesting that they can perform distinct functions in different biological processes. The genomic loci of lncRNAs are more often conserved across species suggesting that syntenic lncRNAs can have conserved functions. They tend to show high structural conservation and low sequential conservation [23].

1.3 Mrhl:

Meiotic recombination hotspot locus RNA or *mrhl* RNA is a 2.4 kb unspliced and polyadenylated long non-coding RNA [24]. It is present on chromosome 8 of mouse within the intron 15 of the phosphorylase kinase beta (PHKB) gene and is transcribed by RNA polymerase II [25]. It has been shown that *mrhl* RNA is expressed in liver, kidney, testis and spleen whereas it is not expressed in brain, heart and lungs [24].

The 2.4-kb primary transcript is nuclear restricted, localizes to the nucleolus, and gets processed to an 80-nucleotide (nt) intermediate RNA by the Drosha machinery. Although Dicer can further process the 80-nt intermediate RNA to 22 nt in an *in vitro* reaction, the mature 22-nt miRNA derived from *mrhl* RNA was not detected *in vivo*, suggesting a nuclear regulatory role for this non- coding RNA. The transcript does not possess a significant open reading frame (ORF) but has a considerable propensity to form a stable secondary structure [25].

Mrhl silencing in mouse spermatogonial GC1-spg cells resulted in perturbation of expression of genes involved in cell adhesion, cell signaling and development and differentiation pathway [26]. One such pathway that was affected was the Wnt pathway. It has been conclusively shown the activation of canonical Wnt signaling pathway upon *mrhl* RNA down regulation. It is interesting to note that TCF4, a key transcription factor in canonical Wnt signaling, is also up regulated under *mrhl* RNA gene silencing. A feed-forward loop in the Wnt activation process that in turn activates TCF4 upon *mrhl* RNA down regulation is a probability [26]. Additionally, it has also been observed that down regulation of extracellular negative modulators like Dkk3 and Sfrp1, suggesting a close relationship between Wnt activation and the expression of negative modulators. It is interesting to note that beta-catenin is recruited with TCF4 at the promoter sites of Dkk3, Tsc22d1, Elk3, Sox8, and Sox12 upon mrhl down regulation. It was shown that *mrhl* down regulates Wnt signaling through a protein partner p68. Ddx5/p68 RNA helicase is an interacting partner of *mrhl* RNA in the nucleus and is required for the regulatory function

of *mrhl* RNA. The chromatin occupancy of *mrhl* RNA was mapped and it showed that *mrhl* RNA regulates the expression of several genes in a Ddx5/p68 dependent manner. This exercise resulted in a total of 1370 genomic loci that were associated with *mrhl* RNA. For this purpose only the annotated genes in the 2 data sets namely (a) Gene loci occupied by *mrhl* RNA and (b) Genes perturbed upon *mrhl* RNA down regulation were compared. The genes common to these 2 data sets represent the ones that are bound by *mrhl* RNA and also perturbed upon *mrhl* RNA down regulation. Hence, these genes might represent the most probable regulatory targets of *mrhl* RNA in the Gc1-Spg cells. This exercise resulted in a subset of 37 genes and it was termed as Genes Regulated by Physical Association of *Mrhl* RNA (GRPAM). Among the GRPAM genes, 3 genes belonged to promoter class which includes Sox8, Lrba and H28 [27].

Gene	Ontology Wnt Signaling	Location of ChOP read w.r.t gene	Gene regulation on *mrhl* RNA silencing
Odz4	Wnt signaling, Embryonic Development, differentiation,	Genic	Downregulated
Lamb3	Wnt signaling, differentiation	Genic	Downregulated
Tsc22d1	Wnt signaling, Tgf β signaling,	51.5 kb upstream	Downregulated
*Prickle1***	Wnt signaling, differentiation, development	63.4 kb upstream	Upregulated
Znrf3	Negative regulation of Wnt signaling, stem cell proliferation, tumor supressor	132.3 kb upstream	Upregulated
*Ostm1**	Activator of Wnt signaling, differentiation.	22.2 kb upstream	Downregulated
	Spermatogenesis		
Sox8	Sertoli cell function, germ cell differentiation (Also in Wnt signaling)	Promoter	Upregulated
Rarg	Retinoic acid signaling, spermatogenesis (Also in Wnt signaling)	11.1 kb upstream	Upregulated
Spag16	Spermatogenesis	Genic	Upregulated
Mael	Spermatogenesis, piRNA pathway, differentiation	33 kb downstream	Upregulated
Spam1	Spermatogenesis, sperm maturation	27 kb upstream	Downregulated
Mageb16	testis specific (adult), regulates differentiation in mouse ESCs	122.5 kb upstream	Downregulated
Npepps,***	Spermatogenesis, antigen processing.	14.7 kb downstream	Upregulated
	Cell Adhesion and Transport		
Lrba	Oncogenesis, endosomal transport,	Promoter	Upregulated
Gabrg2	Ion channel transport, post embryonic development	Genic	Upregulated
Kcnq5	Ion transport	Genic	Downregulated
Grik2	Ion transport, apoptosis	440 kb upstream	Downregulated
Kcnh7	Ion transport	276 kb upstream	Downregulated
*Cdh9***	Cell adhesion	180.2 kb upstream	Upregulated
Nrxn1	Cell adhesion, angiogenesis	Genic	Downregulated
	Signaling		
Palm	cAMP signaling, cytoskeleton	Genic	Upregulated
*Rab40b***	Signaling, Protein transport,	Genic	Downregulated
Ksr1	Ras/MAPK signaling, TNF signaling	Genic	Upregulated
Ssx2ip	Regulation of Rac signaling, Cell adhesion,	Genic	Downregulated
Adamts20	Integrin signaling, Proteolysis, apoptosis,	696 kb downstream	Downregulated
Sla2	Calcium signaling, transcription, endocytosis	0.54 kb downstream	Downregulated
Thbs4	PDGF and PI3K/Akt signaling, Focal adhesion,	25.5 kb downstream	Downregulated
*Il1rapl1**	JNK pathway activation, RhoA signaling, differentiation, ion transport	Genic	Downregulated
*Ppargc1a**	Signaling, differentiation	183.8 kb downstream	Downregulated
	Other Functions		
Serpinb8	Regulation of proteolysis	0.7 Mb upstream	Downregulated
Bach2	Cell cycle control, transcriptional repressor, differentiation	Genic	Upregulated
*Zfp455**	Metal ion binding	24.7 kb downstream	Upregulated
*Mrpl32***	Ribosomal, translation	75.75 kb downstream	Downregulated
Hhipl2	Carbohydrate metabolism	581 kb downstream	Upregulated
*Myo18b**	Vasculogenesis, cardiac development	Genic	Downregulated
Stox2	Not Annotated	Genic	Upregulated
H28	Interferon induced	Promoter	Downregulated

Genes that are physically associated with ChOP sequence reads and also showing perturbation of their expression following *mrhl* RNA down regulation in Gc1-Spg cells (GRPAM). These genes were identified by comparing the genes associated with the ChOP sequence reads (present study) and the transcriptome data from GSE19355. The physical position of occupancy of *mrhl* RNA with each of the GRPAM (total 37) and its functions are also given. Asterik (*) indicates the GRPAM loci which are not bound by p68. ** indicates the GRPAM loci where *mrhl* RNA occupancy is not perturbed upon p68 silencing.

Fig 5: List of GRPAM genes and their ontology and function,
Adapted from: Genome wide chromatin occupancy of mrhl RNA and its role in gene regulation in mouse spermatogonial cells, Akhade et al, 2014, RNA Biology

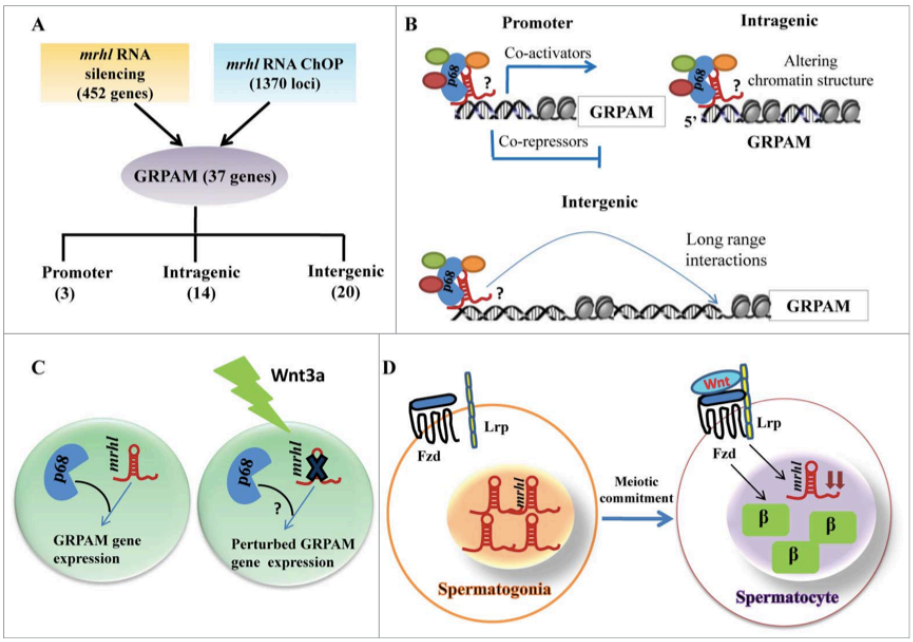


Fig 6: The relationship between chromatin occupancy of mrhl RNA and regulation of gene expression in the context of Wnt signaling in spermatogenesis.
Adapted from: Genome wide chromatin occupancy of mrhl RNA and its role in gene regulation in mouse spermatogonial cells, Akhade et al, 2014, RNA Biology

In a recent study the mechanism by which Wnt signaling regulates *mrhl* RNA has also been explored. Wnt signaling is important in regulating spermatogenesis in mammals. *Mrhl* RNA gets down regulated upon Wnt signaling activation in mouse spermatogonial cells. *Mrhl* RNA has an independent transcription unit and we have identified Ctbp1 as the co- repressor and its occupancy on *mrhl* RNA promoter depends on both B-catenin and TCF4. Upon Wnt signaling activation, Ctbp1 mediates increase in histone repression marks at the *mrhl* RNA promoter. We also demonstrated that Wnt signaling induced *mrhl* RNA down regulation results in an up regulation of various meiotic differentiation marker genes. It is now well documented that Wnt signaling is repressed in spermatogonial cells and is activated in meiotic spermatocytes. Interestingly, expression of *mrhl* RNA is highly down regulated in spermatocytes in comparison to spermatogonial cells. This raises the possibility that *mrhl* RNA plays a significant role in meiotic commitment of spermatogonial cells [28].

1.4 Sox8:

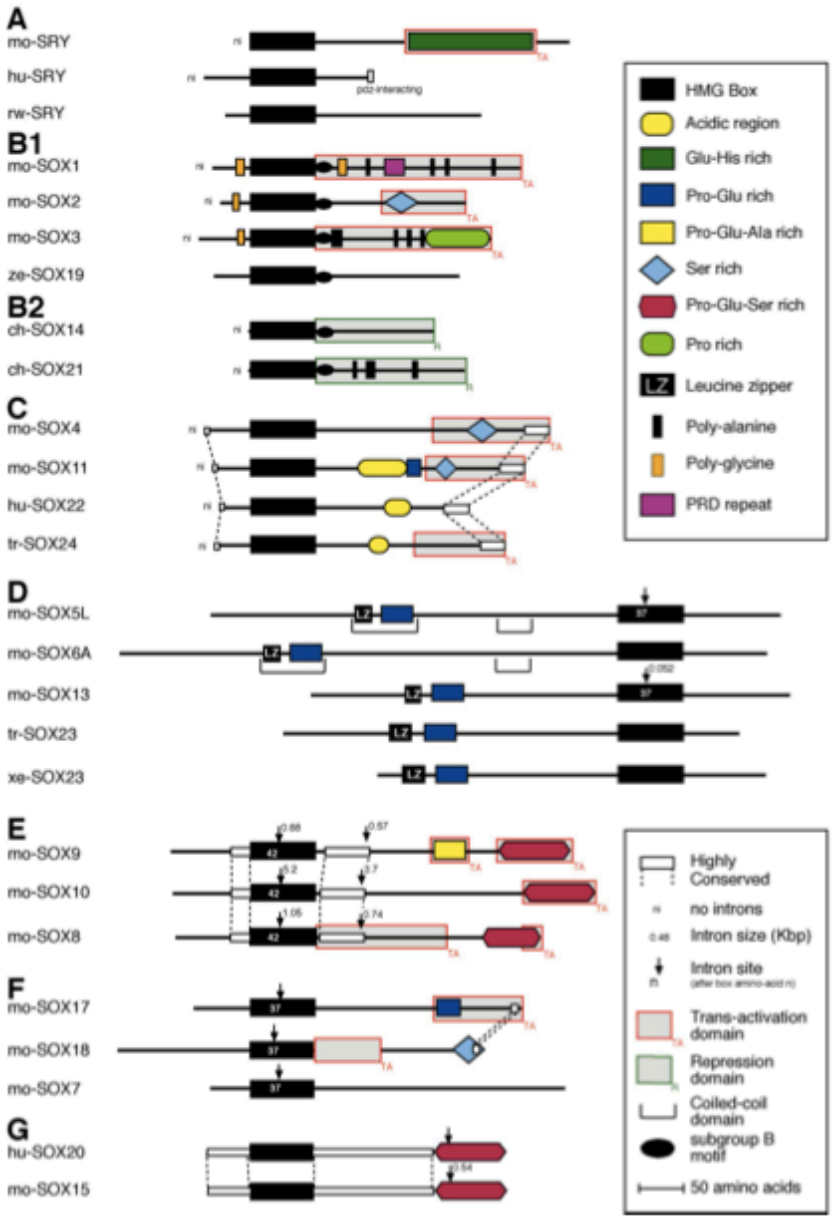


Fig 7: Schematic representation of SOX proteins highlighting conservation within SOX family groups.
Adapted from: Phylogeny of the SOX Family of Developmental Transcription Factors Based on Sequence and Structural Indicators, Bowles et al, 2000, Developmental Biology

Sox proteins belong to the HMG box superfamily of DNA binding proteins and are conserved up the evolutionary tree in a number of vertebrates and invertebrates including various marsupials, reptiles, ascidians and Drosophila melanogaster [29]. They are involved in the regulation of diverse developmental processes such as germ layer formation, organ development and cell type specification [30]. Sox proteins perform their function in a complex interplay with other transcription factors in a manner highly dependent on cell type and promoter context. They exhibit a remarkable crosstalk and functional redundancy among each other. Proteins are grouped into this family if they contain an HMG domain with strong amino acid similarity to the HMG domain of Sry, which is also known as Sry box [31, 32]. It is this box, which gave Sox proteins its name –

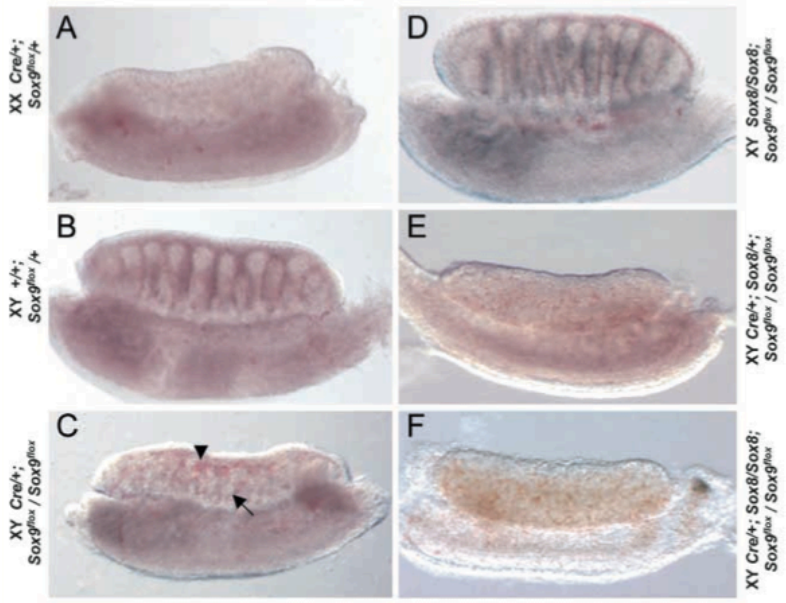


Fig 8: S*ox9* or *Sox8/Sox9* knockout mice show defects in sex cord formation.
Adapted from: Functional analysis of Sox8 and Sox9 during sex determination in the mouse, Chaboissier et al, 2004, Development

Sry related HMG box. Sox proteins bind to specific DNA sequences through this HMG box. Sox proteins are further grouped based on the conservation of their HMG box sequence and the protein structure similarity. Sox E group has three members- Sox8, 9 and 10. All three *SOX E* genes have three coding exons, which are separated by introns at exactly cognate positions. The overall amino acid sequence identity is 54% between SOX9 and SOX10, and 47% for SOX8 versus SOX9 or SOX10. Within the 79 amino acid HMG domain, SOX8 and SOX9 differ by one, SOX8 and SOX10 by five, and SOX9 and SOX10 by four residues. The well-conserved (70–85% identity) 40 residue segment directly N-terminal to the HMG domain comprises the DNA-dependent dimerization domain, specific to the SOX E group. Two additional regions are also well conserved: the very C-terminal 20 residues, with 75–84% identity and a 74–82 residue segment just C-terminal to the second splice site, with 56–71% identity. This segment functions as a strong and a weak transcription activation (TA) domain in SOX8 and SOX10 respectively, but not in SOX9. In contrast to SOX8, a strong TA domain is located at the C terminus in SOX9 and SOX10. The Sox E group has been implicated in a number of human disorders including XY sex reversal, campomelic dysplasia, Waardenberg– Hirschsprung disease type IV (aganglionic megacolon), Yemenite deaf–blind hypopigmentation and chronic intestinal pseudo-obstruction. Consistent with a shared evolutionary origin, SoxE genes frequently display overlapping expression patterns and have been shown to compensate for, or co-operate with, each other in models of gene mutation. For example, in mouse models of Waardenberg–Hirschsprung disease caused by haploinsufficiency of Sox10, ablation of

Sox8 further inhibits enteric neural crest stem cell migration and loss of Sox8 exacerbates the sex-determination defects seen in heterozygous Sox9 mutants. They show abnormal sex cord formation and defects in coelomic vessel formation. These knockout mice also had a reduced number of seminiferous tubules. Interestingly a recent in-vitro study indicated that Sox8 can also activate the Mis promoter and can regulate Amh. When both Sox E genes- Sox8 and Sox9 are ablated at E14.0, only residual levels of Amh transcripts remain, occasionally resulting in partial presence of Mullerian duct derivatives [33].

Sox8 is essential for the maintenance of male fertility beyond the first wave of spermatogenesis. The loss of Sox8 leads to progressive degeneration of the seminiferous epithelium through perturbed physical interactions between Sertoli cells and the developing germ cells. It has been shown that SOX8 is a product of Sertoli cells, and while not critically required for testis specification and development, it is critical for the maintenance of adult male fertility. During development and the first wave of spermatogenesis, $Sox8^{-/-}$ mice established histologically normal spermatogenesis and produced motile sperm. In young mice, sperm were capable of producing pups however, by 5 months of age, $Sox8^{-/-}$ mice were sterile, testis weights were significantly reduced by 2-months- of-age as a result of the sloughing of round germ cells, and sperm counts were further reduced as the result of spermiation failure. Of those sperm that did reach the epididymis, many were immotile or incapable of progressive motility. By 5 months of age, virtually no progressively motile (and therefore fertile) sperm were produced and by 9 months there was an extreme deregulation of the cycle of the seminiferous epithelium. Collectively, these data infer that Sox8 is required

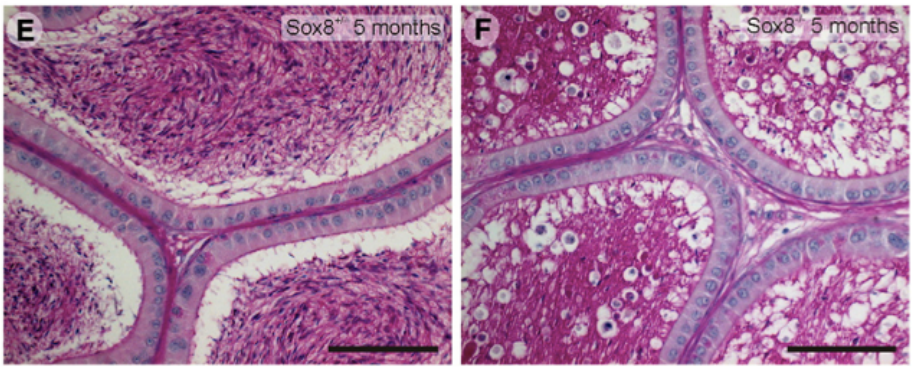


Fig 9: Testis and epididymal histology of Sox8 heterozygotes and Sox8 knockout animals at 5 months of age.
Adapted from: Sox8 is a critical regulator of adult Sertoli cell function and male fertility, O'Bryan et al, 2008, Developmental Biology

for fertility in most male mice, but in a percentage of cases (~20%) young animals can sire pups suggesting that at least some sperm resulting from the first few waves of spermatogenesis can be fertile. There is clear indication that Sox8 is a transcription factor with a role in the maintenance of male fertility and Sox8-initiated transcription in Sertoli cells results in the production of a set of molecules, which are essential for the maintenance of spermatogenesis and normal sperm function [34].

1.5 Aim of the present study-

In our laboratory we have discovered and characterized a novel, 2.4kb, polyadenylated, long noncoding RNA- *mrhl* transcribed from the 15th intron of PHKB gene in mouse. This RNA is negatively regulated by Wnt signaling. On genome wide chromatin mapping 37 loci were identified where the RNA physically interacts with the chromatin and the down regulation of *mrhl* RNA leads to change gene expression of these 37 target genes. One of the targets where *mrhl* RNA binds to the promoter region of the gene is Sox8. Sox8 is an important transcription factor in mammalian spermatogenesis. We were interested in studying the regulation of Sox8 expression by this non-coding RNA under Wnt activation.

The objectives of the project are as follows-

- Expression analysis of Sox8 with Sox9 as control under control and Wnt conditions.
- Dissection of the role of *mrhl* and Wnt in Sox8 regulation
- Analysis of Sox8 promoter and elucidation of the mechanism of Wnt mediated up-regulation of Sox8.
- Role of Sox8 in meiotic commitment.

CHAPTER 2
MATERIALS AND METHODS

2.1 Materials:

Gc1-Spg cell line (ATCC, CRL-2053) and TM4 cell line (ATCC, CRL-1715) were obtained from ATCC. Mouse L-control cell line (ATCC, CRL-2648) and L-Wnt3A cell line (ATCC, CRL-2647) were kind gifts from Dr. Jomon Joseph (NCCS, India).

All chemical reagents were of AR grade and were purchased from Sigma. Protein G dynabeads (10003D) and Lipofectamine 2000 (11668027) were purchased from Invitrogen. DNAseI (MO303), T4 DNA Ligase (M0202S) and Restriction enzymes were purchased from New England Biolabs. Luciferase assay Kit was supplied by Promega, and Sox8 shRNA (SHCLNG-NM_011447) was obtained from Sigma. *Mrhl* siRNA pool was obtained from Dharmacon [26].

The list of antibodies used in the present study is given below with their manufacturer and catalogue number:

Antibody	**Manufacturer**	**Catalogue No.**
Sox8	Pierce	PA1-28072
Sox9	Abcam	Ab3697
Vimentin	Abcam	Ab20346
E-cadherin	BD biosciences	610404
Tubulin	Sigma	T8203
Tcf4	Millipore	05-511
Beta-catenin	BD biosciences	610154
H3K27me3	Abcam	Ab6002
H3K4me3	Abcam	Ab12209
H3	Abcam	Ab46765
c-Myc	Santa-cruz	9E10

List of primers-

Gene	Forward primer	Reverse primer
Beta-actin	AGGTCATCACTATTGGCAACG	TACTCCTGCTTGCTGATCCAC
Sox9	TGCTGAACGAGAGCGAGAAGAGAC	GGACCCTGAGATTGCCCAGAGTG
Sox8 primary	CTGTGTTTTCTGTAGCTTGCTG	CCTGGAGCCCACCTGTGTG
Sox8	GCTGTGGCGCTTGCTGAG	CTGTGTGGTGGTCACTGTG
Mrhl	TGAGGACCATGGCTGGACTCT	AGATGCAGTTTCCAATGTCCAAAT
c-Myc	CTAGTGCTGCATGAGGAGAC	CTCTTGGCAGCTGGATAGTC
Max	GATAACGATGACATCGAGGTG	TTGCTCCAGAAGAGCATTCTG
Mad	CTGTCCACCAAATAGACCAG	TGCATGCTGCCTCGCTCG
Stra-8	CGTGGCAAGTTTCCTGGACAAG	GGCTCTGGTTCCTGGTTTAATGG
Lhx8	GTCTGGAGATAGTTGGCTCGAGT	GGATGGTAGGCTTTGTAAACTAG
c-Kit	CCCGACGCAACTTCCTTA	CGCTTCTGCCTGCTCTTC
Mtl-5	CGTCTGGGAGCTGCTAAA	GGAGGTCCTGAGAACTTGG
Hspa2	CTACGTGGCCTTCACTGACA	GGTCAGGATGGACACATCGA
Beta-actin ChIP	TCCACAAGGGCGGAGGCTAT	GGGTTTTATAGGACGCCACA
Ccnd1	CCATTCTCCCGGTTTAAGAACAG	AGCCTTCGTAGATATGCAAATCG
Cad	CACACAAGACTTCCAGCTCAC	TGTAGTCAATAATACAAATTAGTAC
NM1	GATCCCAGACCTGAAGGTTAG	GTGGTCTCTAATATCTCCATCTC
NM2	GAGATGGAGATATTAGAGACCAC	ACCTGTTTCCCAAGCCCGTG

NM3	CACGGGCTTGGGAAACAGGT	CTTGTCTGGGTGGCATAGAG
NM4	CTCTATGCCACCCAGACAAG	CGAAACTCAAAGGAGGCGATG
NM5	CATCGCCTCCTTTGAGTTTCG	GAGGCGAGAGCCAAGCCTG
NM6	CAGGCTTGGCTCTCGCCTC	CCTAGAGAGAAGGCGGAGAG
NM7	CTCTCCGCCTTCTCTCTAGG	CTGGGTGGTCTGGCGGAG
NM8	CTCCGCCAGACCACCCAG	GTCCTTCCGGGTATGACCTG
NM9	CAGGTCATACCCGGAAGGAC	CTGGCAGTGACTGCTGAGTC
NM10	GACTCAGCAGTCACTGCCAG	GCTAGACAGAGGTGGGAGG
NM11	CCTCCCACCTCTGTCTAGC	CCAGGTGACCCTGGTAACTC
NM12	GAGTTACCAGGGTCACCTGG	GCAACCAAACCCGGGTTCTG
NM13	CAGAACCCGGGTTTGGTTGC	AAGGTTCCTGAGTCCAACATC
NM14	GATGTTGGACTCAGGAACCTT	GCAGCACGGCTTGACTTTCG
NM15	CGAAAGTCAAGCCGTGCTGC	GGTCAGAGGGCTAAGGGTG
NM16	CACCCTTAGCCCTCTGACC	TCCGTACCACGTGGGCCAG
NM17	TGGCCCACGTGGTACGGAG	AATCCAAAGAAGCCTTGCGAC
NM18	TCGCAAGGCTTCTTTGGATTC	TCCGGGTTGGCAGGTCCTC
NM19	AGGACCTGCCAACCCGGAC	CAGTTCTGGGACCCGCAAC
Tcf4 mutation	GTACATCGCCTCCGCTGAGTTTCGCG GGG	CCCCGCGAAACTCAGCGGAGGCGATG TAC

2.2 Cell culture:

Gc1-spg cell line (type B spermatogonia) and TM4 cell line (Sertoli cell) were obtained from the American Type Culture Collection (ATCC). Gc1-spg was cultured in Dulbecco's modified Eagle's medium (Sigma) supplemented with 10% fetal bovine serum (Invitrogen) and 100units/ml penicillin and streptomycin solution (Sigma) whereas TM4 cell line was cultured in Dulbecco's modified Eagle's medium (Sigma) supplemented with 10% Horse serum and 5% fetal bovine serum (Invitrogen) and 100units/ml penicillin streptomycin solution at 37^{o}C in a humidified 5% CO_2 atmosphere.

2.3 Preparation of control and Wnt3a conditioned medium.

L- control or L-Wnt3a cells (1×10^6 cells) were seeded in 90mm culture dishes containing DMEM with 10% FBS. After 48 h, the cells were cultured with fresh medium and incubated for 1 more day. The medium was then collected, centrifuged at 500 × g for 5min, filtered through a 0.2-um syringe filter and subsequently stored at −20°C till further use. For treatment of Gc1-Spg cells, the control as well as Wnt3a conditioned medium were diluted 2:1 with DMEM containing 10% FBS.

2.4 Transfection of mammalian cells with lipofectamine 2000

The cells were transfected with the clone using lipofectamine. 1ug/ml concentration of plasmid was used and 2ul lipofectamine was used for every 1ug of DNA. Serum free media was put in the wells before transfection. Required amount of DNA was added to SF media and required lipofectamine was also added to SF media in separate eppendorfs tubes. Both were incubated for 5 minutes. After that DNA was added drop wise to lipofectamine and incubated for 20 minutes at RT. Then 300ul of this was added to each well. Media change was given 8 hours post transfection (complete media).

2.5 Cloning of 1kb upstream Sox8 and Sox9 promoter and Luciferase assay.

The 1kb sequence upstream to Sox8 and Sox9 gene were cloned in pGL3-Basic vector. Clones were confirmed by DNA sequencing. The TCF4 binding site on the Sox8 gene promoter (−800 bp) was mutated from CTTTGA to CGCTGA using Stratagene QuikChange

Site-Directed Mutagenesis Kit (catalog #200518). For Luciferase assay, Gc1-Spg cells were transfected with 1ug of the pGL3-Basic vector or 1ug of Sox8/Sox9 promoter clone in a six-well plate using Lipofectamine 2000. Transfection with CMV- **β**Gal (0.4ug) plasmid was done as an internal control for transfection efficiency. After 24h, cells were harvested and processed for Luciferase assay as per the manufacturer's proto- col (Promega). Luciferase readings for the pGL3- Basic as well as Sox8/Sox9 promoter plasmid were normalized with the Luminometer readings obtained for CMV- βGal.

2.6 Whole cell/tissue lysate

The cell pellet or tissue was incubated in RIPA buffer (50mM Tris, 150mM NaCl, 1%NP40, 0.5% sodium deoxycholate, 0.1% SDS-sodium dodecyl sulphate, PIC- Protease inhibitor cocktail, PMSF-phenylmethylsulfonyl fluoride) for 30 minutes in ice. Cells were lysed by constant pipetting and vortexing whereas tissues were homogenized using a hand operated homogenizer. The suspension was then centrifuged at 14000g for 10 minutes followed by which the supernatant was collected. Bradford assay was performed to quantitate the amount of protein.

2.7 Western Blot

The proteins were resolved by 12% SDS-PAGE and transferred to nitrocellulose membrane using Amersham Biosciences semidry transfer unit. Before transfer, the Whatman filter papers and nitrocellulose membrane were soaked in the transfer buffer (7.5g of Glycine, 1.65g of Tris, 1ml of 10% SDS, 100ml of methanol made upto 500ml with water). The protein bands were transferred at 140mA of current for 70 minutes. After transfer, the membrane was stained with 0.1% Ponceau (0.1g of Ponceau, 5ml of acetic acid made up to 100ml using water) to check the success of transfer of proteins. The membrane was blocked using 5% skimmed milk in 1X PBS for 45 minutes at room temperature followed by one wash with 1X PBS for 5 minutes. The membrane was then incubated with the primary antibody in 1% skimmed milk in 0.05% PBST (1X PBS+ Tween-20) overnight at 4^{o}C. The membrane was washed with 0.05% PBST twice for 5 minutes each. Subsequently the membrane was incubated with the secondary antibody in 1% skimmed milk at a dilution of

1:4000 for 45 minutes at room temperature. This was followed by two washes of 5 minutes each using 0.05% PBST. Detection was carried out using the Amersham ECL-analysis system.

2.8 RNA isolation and qPCR

RNA isolation was done using Trizol (Takara) as per the manufacturers protocol. DNAse (NEB) treatment was carried out for 45 minutes at 37C followed by inactivation at 65C for 10 minutes. The precipitated RNA was quantified using nanodrop and the integrity of RNA was checked on 2% agarose gel. 2ug of RNA was used to set cDNA using RT Revertaid (Thermo scientific) as per the manufacturers protocol.

All qPCR reactions were performed in 15µl reaction volume using 2.5ul of cDNA as template.

2.9 Spreading of testicular cells

In order to obtain the pool of spermatogonia and spermatocytes we selected 7 and 21 days old mice, testes were dissected out, chopped in DMEM medium and filtered through cheese cloth. Cells were fixed in 4% PFA for 15 min at room temperature with gentle agitation. Later, cells were washed once with PBS, permeabilized with 0.5% Triton X-100 in PBS for 10 min with gentle agitation and again washed with PBS at room temperature. Cells were suspended in appropriate volume of 1XPBS, spread over the cover slip and were allowed to dry in humidified chamber for 2–3 h.

2.10 Immunofluorescence staining of spreads

Cells were blocked in 1% BSA for 45 min at 4°C and subsequently incubated at room temperature with primary antibody (diluted with 0.1% BSA) for 30 minutes. Cells were washed three times with PBS for 5 min each, followed by incubation with appropriate Alexa Fluor conjugated secondary antibody (1:400 diluted with 0.1% BSA) for 30 min at room temperature. The cells were washed thrice for 5 min each with PBS and nuclear stained with 4′, 6′-diamidino-2-phenylindole (DAPI) and mounted in 60% glycerol. The images were acquired in an LSM 10 Meta Confocal microscope (Carl Zeiss) and ana- lyzed

by the software provided by Carl Zeiss and intensity quantification was done using ImageJ software.

2.11 Immunofluorescence staining of cell lines

GC1-spg and TM4 cells were seeded on cover slips placed in 24 well plates and after they attained the required confluency the media was removed followed by 1X PBS wash. Then the cells were fixed in 4% PFA for 10 minutes at room temperature. Following this the same procedure as above was carried out.

2.12 Chromatin immunoprecipitation

Cells were harvested and cross-linked by adding 1% formaldehdye for 10 minutes at room temperature with shaking. This was followed by addition of 0.125M of glycine and incubating the mixture for 5 minutes at room temperature. The cross linked pellet was obtained by centrifugation at 2000 rpm at 4^{o}C for 10 minutes. The cell pellet was resuspended in 1ml of SDS lysis buffer (1%SDS, 10mM EDTA, 50mM Tris, PIC, PMSF) and incubated on ice for 15 minutes with resuspension of the pellet every 5 minutes. This was followed by sonication of the lysate using a Biorupter set at high power with a 30 seconds on/off interval for a total of 35 cycles. Insoluble materials were removed by centrifugation at 13000g at 4^{o}C for 10 minutes. The supernatant was transferred to a new tube. 100 µl was saved for use as input material whereas the rest of the lysate was incubated with the antibody for immune-precipitation overnight at 4^{o}C. The immune-complexes were allowed to bind to Protein G Dynabeads by incubating the lysate-containing antibody with 45µl of beads for 3 hours at 4^{o}C. The beads bound by immune-complexes were obtained by magnetic precipitation washed with each of the following buffers:

- Low salt buffer (0.1% SDS, 1% Triton-X 100, 2mM EDTA, 20mM Tris Cl, 150mM NaCl)
- High salt buffer (0.1% SDS, 1% Triton-X 100, 2mM EDTA, 20mM Tris Cl, 500mM NaCl)

In each wash, the beads were incubated with the wash buffer for 5 minutes at 4^{o}C. The immune-precipitated materials were eluted from the beads by adding 400µl of elution

buffer (0.1M $NaHCO_3$, 1% SDS) to each ChIP reaction and the samples were incubated in end-end rotator at room temperature for 45 minutes. The elutes were then processed for DNA precipitation.

2.13 Isolation of DNA by phenol chloroform method

DNA was precipitated from input and elutes by phenol chloroform method. Proteinase k treatment was carried out at 55^{o}C for 4 hours- overnight. The samples were centrifuged at 12000g for 10 minutes followed by which the supernatant was collected. Phenol-chloroform-Isoamyl alcohol (25:24:1) mix was added to each sample and vortexed vigorously. Supernatant was collected after centrifugation at 10000g for 5 minutes. Equal volume of phenol-chloroform was added and thoroughly mixed. The samples were centrifuged at 10000g for 5 minutes to obtain the aqueous phase, which was collected into a new tube. For DNA precipitation, 1/10th the volume of sodium acetate (3M pH 5.2) was added along with twice the volume of ethanol and incubated kept at -20^{o}C for 1 hour. The precipitated DNA is pelleted by centrifugation at 12000g for 10 minutes followed by 70% ethanol wash. The pellet was dried and dissolved in 25µl of nuclease free water.

2.14 Microccocal Nuclease digestion

Cells were harvested and pelleted at 2000g for 3 minutes at 4^{o}C. It was resuspended in lysis buffer (60mM KCl, 15mM NaCl, 15mM Tris, 0.34M sucrose, 2mM EDTA, 0.5mM EGTA, 1mM DTT, 0.03% Triton, 1% Glycerol, PIC) and incubated in ice for 10 minutes with regular vortexing. The suspension was then centrifuged at 14000g for 10 minutes at 4^{o}C. The pellet was then resuspended in wash buffer (60mM KCl, 15mM NaCl, 15mM Tris, 0.34M Sucrose, 1mM DTT, PMSF, PIC) and centrifuged at 14000g for 10 minutes at 4^{o}C. The pellet was then suspended in MNase buffer (10mM Tris, 10mM KCl, 2mM $CaCl_2$) and micrococcal nuclease enzyme (1.5 units) was added followed by incubation at 37^{o}C for 40 minutes. The reaction was then stopped by addition of 5mM EGTA and incubation on ice for 5 minutes. The nucleosomes were then isolated by centrifugation at 2000g, 10 minutes at 4^{o}C. The pellet was then suspended in LSDB 250 buffer (20% Glycerol, 50mM HEPES, 3mM MgCl2, 250mM KCl, PIC) and was incubated overnight

with the appropriate antibody. The beads were blocked with 0.1% BSA at 4^{o}C for 30 minutes followed by which they were added to the nucleosomes and incubated for further 4 hours at 4^{o}C. The beads were then washed with LSDB 250 buffer 3 times, 5 minutes each at 4^{o}C and eluted like described for ChIP. The elute was subjected to DNA extraction as described above.

CHAPTER 3
RESULTS

3.1 Sox8 is expressed in germ cells

As there are no reports of Sox8 expression in mammalian male germ cells, we first checked for its expression for which testis cellular spreads were made from 7 day old and 21 day old mice and the following markers were used- E-cadherin- spermatogonia, SCP3- spermatocytes, Vimentin- Sertoli cells. The reason for choosing these two particular ages is that at P7 spermatogonia are the major population among germ cells in addition to the somatic cells (Sertoli and Leydig cells) and Wnt signaling is inactive at this stage whereas in P21 spermatocytes are the major population and Wnt signaling is active. Sox9 was used as a control as Sox9 belongs to the same group E of Sox proteins as Sox8 but doesn't interact with mrhl.

It was observed that both Sox8 and Sox9 are expressed in spermatogonial cells as well as in spermatocytes as shown in Fig 1A. We also checked for the expression in Sertoli cells. As shown in Fig 2A it was seen that both Sox8 and Sox9 are expressed in Sertoli cells at both the stages. The spermatogonial cell line used for most of the experiments is GC1-spg and the Sertoli cell line used is TM4. Sox8 expression was validated in both these cell lines as well as shown in Fig 1C.

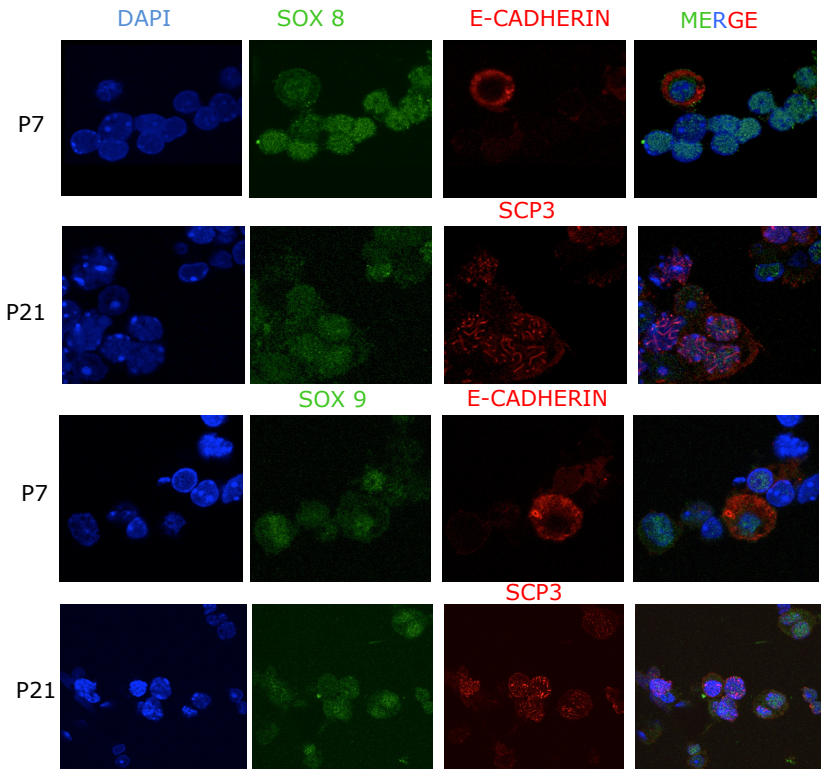


Fig 1A: IF for Sox8 and Sox9 in P7 and P21 mice testes with spermatogonium and spermatocyte marker

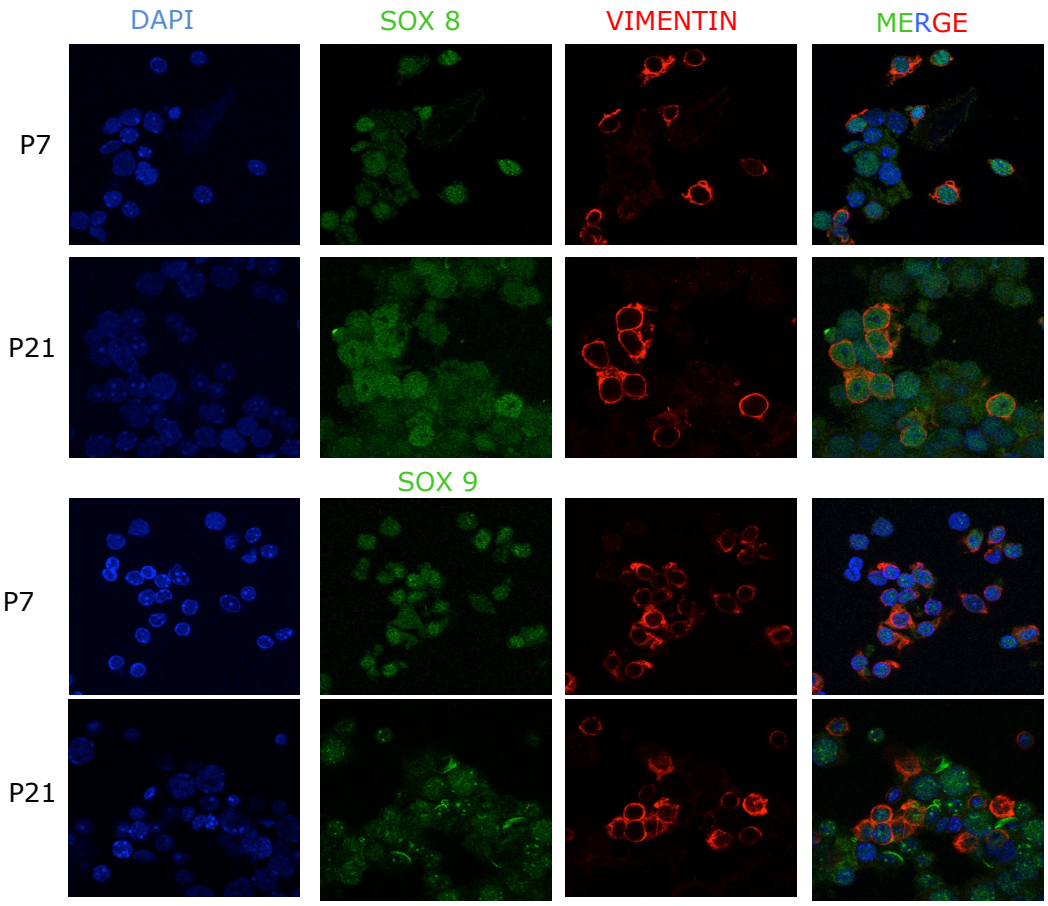


Fig 1B: IF for Sox8 and Sox9 in P7 and P21 mice testes with sertoli cell marker- Vimentin

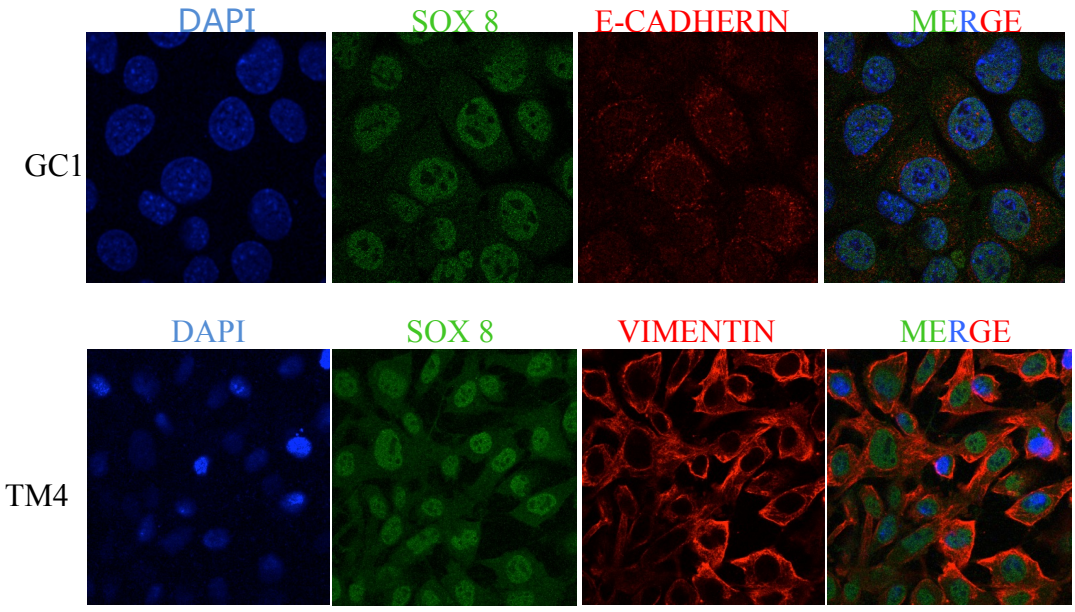


Fig 1C: IF for Sox8 in GC1-spg and TM4 cell line

3.2 Sox8 expression is up regulated in germ cells and Sertoli cells on Wnt activation

It is known that Wnt signaling can down regulate mrhl RNA and mrhl RNA binds to the promoter of Sox8 and plays a role in the expression of Sox8. Therefore we went ahead to check the expression of Sox8 on Wnt activation. The expression of Sox8 and Sox9 under the cue of Wnt was checked by qPCR. The prespliced primary transcript of Sox8 was also scored for by using intron-exon junction primers to establish that it is a transcriptional activation. The result was also supported by luciferase assay in GC1-spg and TM4 as shown in Fig 2A and 2B. The results were also validated by western blot analysis of Sox8 protein. It was observed that Sox9 expression is unaffected by Wnt activation in both germ cells and Sertoli cells (Fig 2C) whereas Sox8 was up regulated in both the cell lines as shown in Fig 2D. In-vivo, the results were confirmed by western blot of P7 and P21 testes (Fig 2E).

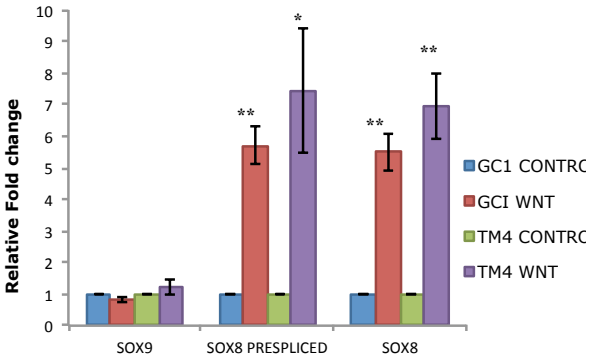


Fig 2A: qPCR analysis of Sox8 and Sox9 in control and Wnt treated GC1-spg and TM4

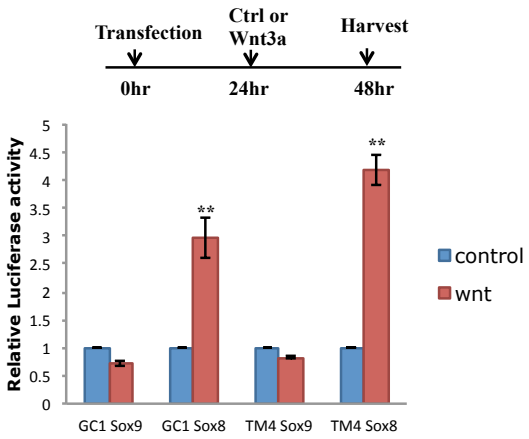


Fig 2B:Luciferase assay of Sox8 and Sox9 in control and Wnt treated GC1-spg

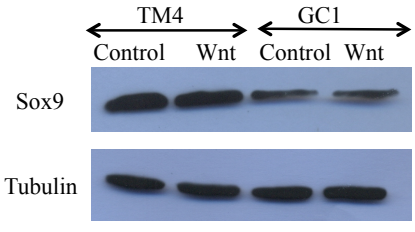


Fig 2C:Western blot of Sox9 in control and Wnt treated GC1-spg and TM4

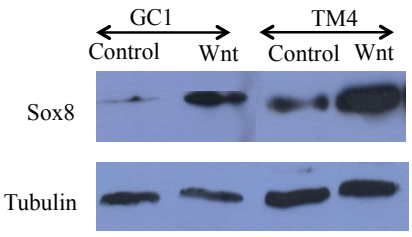


Fig 2D: Western blot of Sox8 in control and Wnt treated GC1-spg and TM4

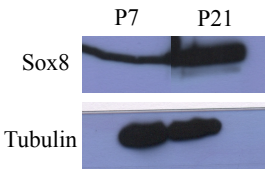


Fig 2E: Western blot of Sox8 in P7 and P21 mouse testes

3.3 Sox8 promoter has a Tcf-4 binding site in addition to the ChOP site.

Sox8 has an interesting genomic architecture and is transcribed from a bidirectional promoter. The ChOP site is located at around 141 bp upstream of the Transcriptional start site (TSS) in the promoter. Bioinformatically a Tcf-4 binding site was also mapped at around 800 bp upstream of TSS in the promoter (Fig 3A). Now that a Tcf-4 site is also present it is a possibility that Tcf-4 and b-catenin can directly regulate the expression of Sox8 by occupying this site. The possible modes of regulation are shown in Fig 3B and are tested in the subsequent sections.

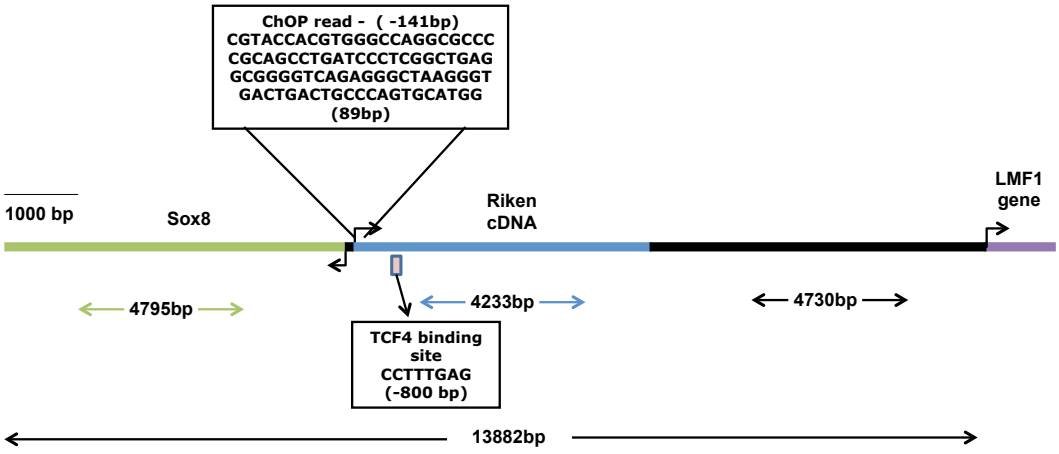


Fig 3A: Sox8 genomic architecture

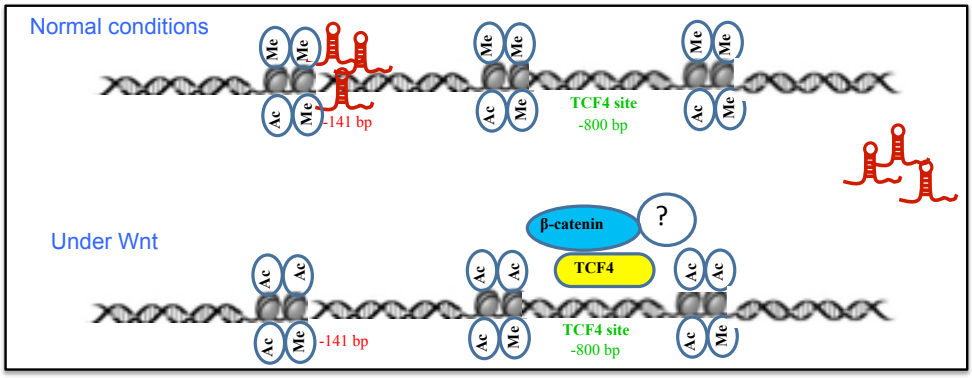


Fig 3B: Model showing the modes of regulation of Sox8

3.4. Dissection of the role of *mrhl* and Wnt in Sox8 regulation

3.4.1 Tcf-4 and beta-catenin occupy the Tcf-4 site on Sox8 promoter

The occupancy of Tcf-4 and beta-catenin on Sox8 promoter was checked by ChIP-qPCR. It was found that Tcf-4 constitutively occupies the site (Fig 4A) whereas beta-catenin is recruited only upon Wnt activation (Fig 4B) in GC1-spg cells. The results were also validated in P7 and P21 mouse testes. Given that the site is occupied by Beta-catenin on Wnt activation, next the functionality of the site was checked. Site directed mutagenesis was carried out in the Sox8 promoter followed by luciferase assay. It was seen that even though there was a reduction in the Sox8 promoter function on mutation (Fig 4C) it was not as drastic as would be expected if Tcf-4 binding site was the sole regulator of Sox8 expression.

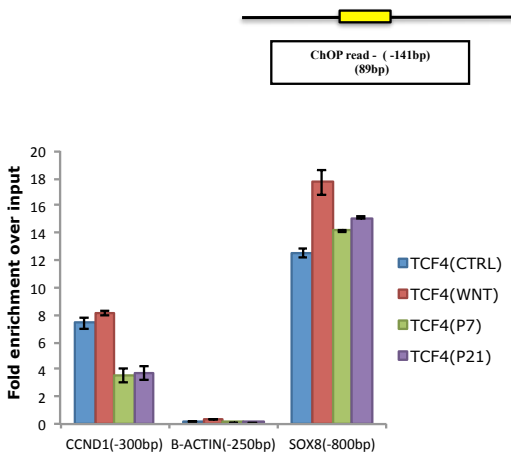


Fig 4A: ChIP-qPCR for Tcf-4

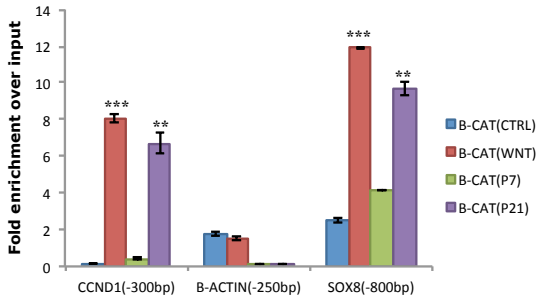


Fig 4B: ChIP-qPCR for Beta-catenin

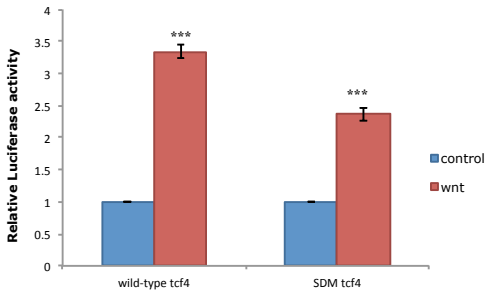


Fig 4C: Luciferase assay of Sox8
in wild-type and mutated promoter

3.4.2 Mrhl RNA independently regulates Sox8

Next the role of mrhl RNA in regulation of Sox8 was investigated. In control conditions mrhl RNA is expressed whereas under Wnt activation mrhl RNA is down regulated. To test the role of Mrhl in Sox8 regulation mrhl overexpression and mrhl silencing was done under control and Wnt conditions. It was observed that under control conditions when mrhl was silenced Sox8 expression was up regulated (Fig 4D) whereas under Wnt conditions when mrhl RNA was overexpressed in trans the up regulation of Sox8 was abrogated (Fig 4D). The results were also corroborated by Luciferase assay (Fig 4E).

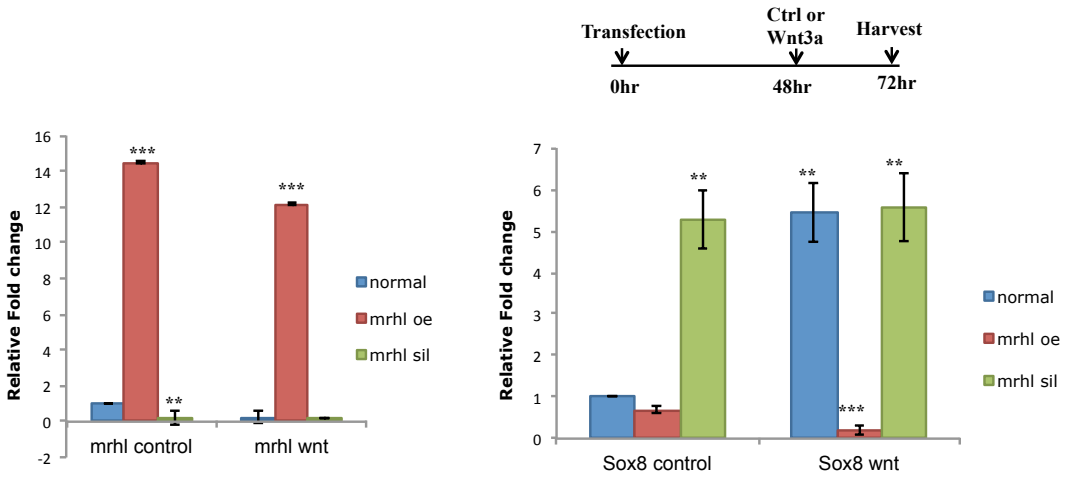


Fig 4D: qPCR analysis of Sox8 under *mrhl* overexpression and *mrhl* silenced conditions in GC1-spg

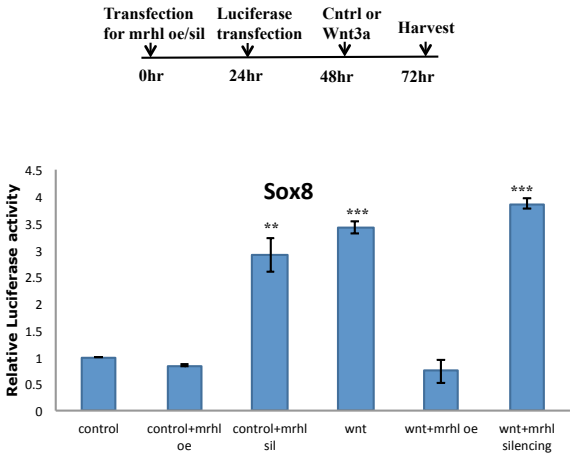


Fig 4E: Luciferase assay of Sox8 under *mrhl* overexpression and *mrhl* silenced conditions in GC1-spg

3.5 Chromatin dynamics change to facilitate the transcriptional activation of Sox8 on Wnt activation

3.5.1 Change in histone activation and repression marks on Sox8 promoter upon Wnt activation

We have established so far that Sox8 is up regulated on Wnt activation and the removal of mrhl RNA from the Sox8 locus is important for this transcriptional activation. The Tcf-4 binding site is occupied by beta-catenin on Wnt activation but is minimally active. After this we were interested in exploring the chromatin changes taking place in the promoter region to facilitate this transcriptional activation. We looked at the histone activation and repression marks at both the ChOP site and the Tcf4 binding site. There was an increase in activation mark- H3K4me3 and decrease in repression mark- H3K27me3 at both the sites on Wnt activation in GC1-spg as shown in Fig 5A and 5B. The results were also corroborated in-vivo in P7 and P21 mouse testes.

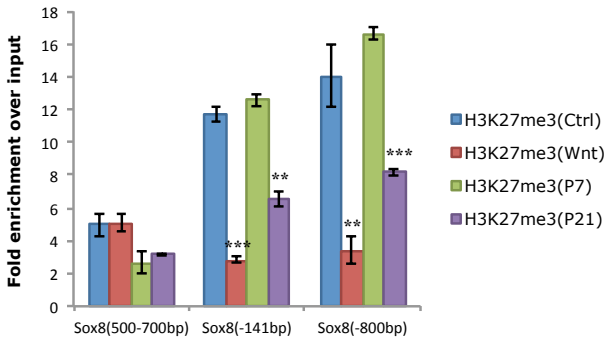


Fig 5A: ChIP-qPCR for H3K27me3 histone modification in control and Wnt treated GC1-spg and in P7 and P21 testes

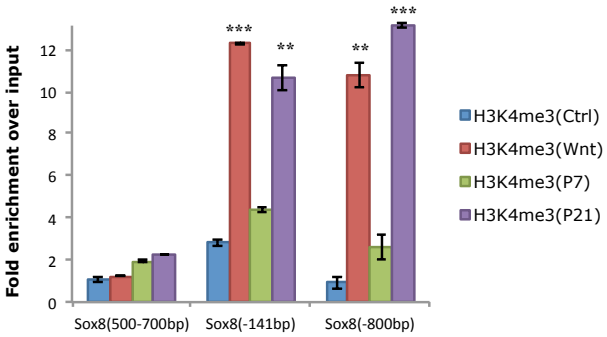


Fig 5B: ChIP-qPCR for H3K4me3 histone modification in control and Wnt treated GC1-spg and in P7 and P21 testes

3.5.2 Nucleosome mobility at the ChOP site and the Tcf-4 binding site on Wnt activation

Nucleosome depletion is a common phenomenon, which occurs primarily at transcriptionally active promoters to facilitate the binding of the transcriptional machinery and various co-activators. Sox8 also shows a transcriptional activation on Wnt treatment. Therefore the nucleosome position was mapped at every 50bp distance on the Sox8 promoter in control and Wnt treated GC1-spg cells by H3 ChIP-qPCR. Nucleosome depletion was observed at both the ChOP site and the Tcf-4 site as shown in Fig 5C.

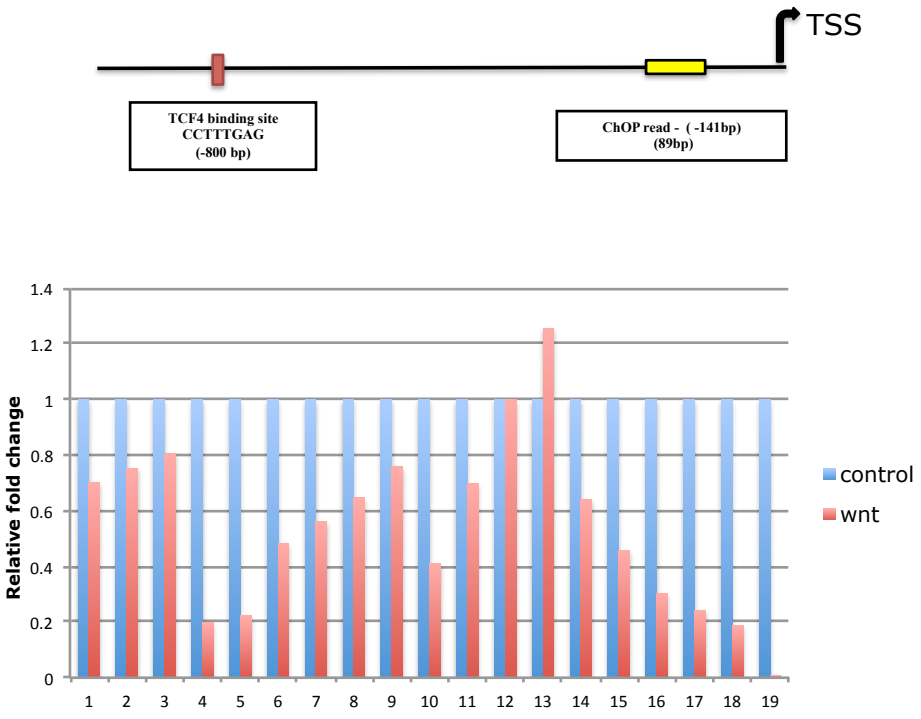


Fig 5C: Nucleosome mapping in control and Wnt treated GC1-spg

- Average of two independent experiments

3.6. Transcription factors involved in regulation of Sox8

3.6.1 Myc-Max binding site is present on Sox8 promoter

The 1kb upstream promoter of Sox8 was bioinformatically analyzed for putative transcription factor binding sites. Five transcription factor binding sites were identified out of which one was Myc-Max as shown in Figure 6A. Interestingly, it was observed that the binding site of all transcription factors including Myc-Max overlapped with the ChOP site indicating the importance of this locus in Sox8 gene regulation.

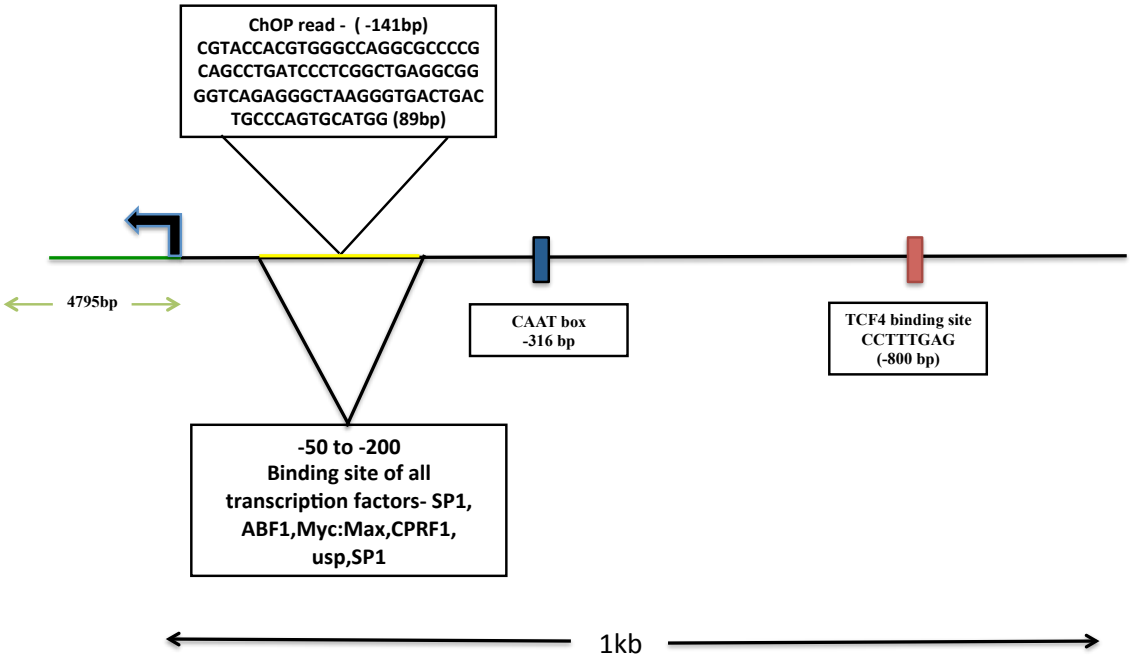


Fig 6A: Identification of putative transcription factor binding sites on Sox8 promoter

3.6.2 c-Myc occupies Sox8 promoter on Wnt activation

There have been reports earlier showing up regulation of c-Myc on Wnt activation. We checked the same in our model system and it was observed that c-Myc is up regulated on Wnt activation in GC1-spg cell line whereas there was no significant effect on the expression of Max or Mad as shown in Fig 6B. Next we checked for the occupancy of c-Myc on Sox8 promoter on Wnt activation. As shown in Fig 6C c-Myc exhibits a strong occupancy of the Sox8 promoter on Wnt activation establishing it to be a strong candidate for a co-activator in Wnt mediated up regulation of Sox8.

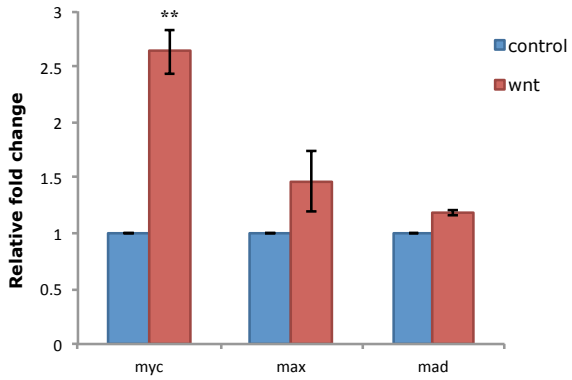


Fig 6B: Expression of Myc Max Mad on Wnt treatment in GC1- spg

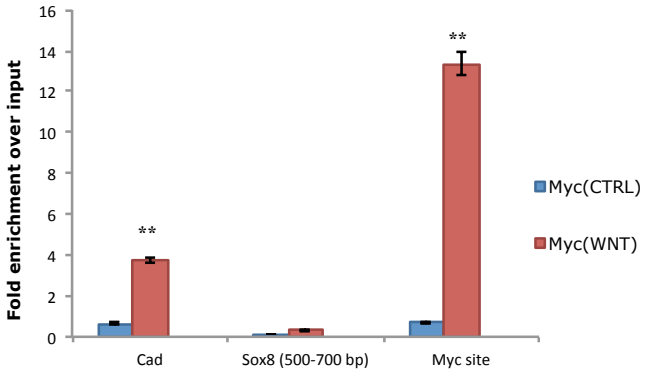


Fig 6C: ChIP-qPCR for c-Myc in control and Wnt treated GC1-spg

3.7 Sox8 binding site is present in promoters of pre-meiotic and meiotic marker genes

Sox8 expression was known to be limited to Sertoli cells, But now that we see that Sox8 is not only expressed in germ cells but also temporally regulated by Wnt signaling pathway we were interested in seeing the role of Sox8 in germ cells. It has been shown earlier that Wnt mediated down regulation of mrhl RNA is essential for various pre-meiotic and meiotic marker genes up regulation in spermatogonial cells thus paving the path for the differentiation of spermatogonia to spermatocytes. Now that it is established that mrhl RNA can also regulate Sox8 expression it is a possibility that Wnt mediated mrhl RNA down regulation leads to up regulation of Sox8 which in turn lead to activation of these meiotic commitment marker genes. When the Sox8 binding site was mapped on the promoter of these genes it was observed all of these genes with the exception of Lhx-8 possess Sox8 binding site as shown in Fig 7A. Further the functionality of Sox8 in regulation of these genes was checked. Sox8 was silenced under Wnt conditions and it was observed that when the Wnt mediated up regulation of Sox8 is perturbed it abrogates the up regulation of these pre-meiotic and meiotic marker genes as shown in Fig 7B.

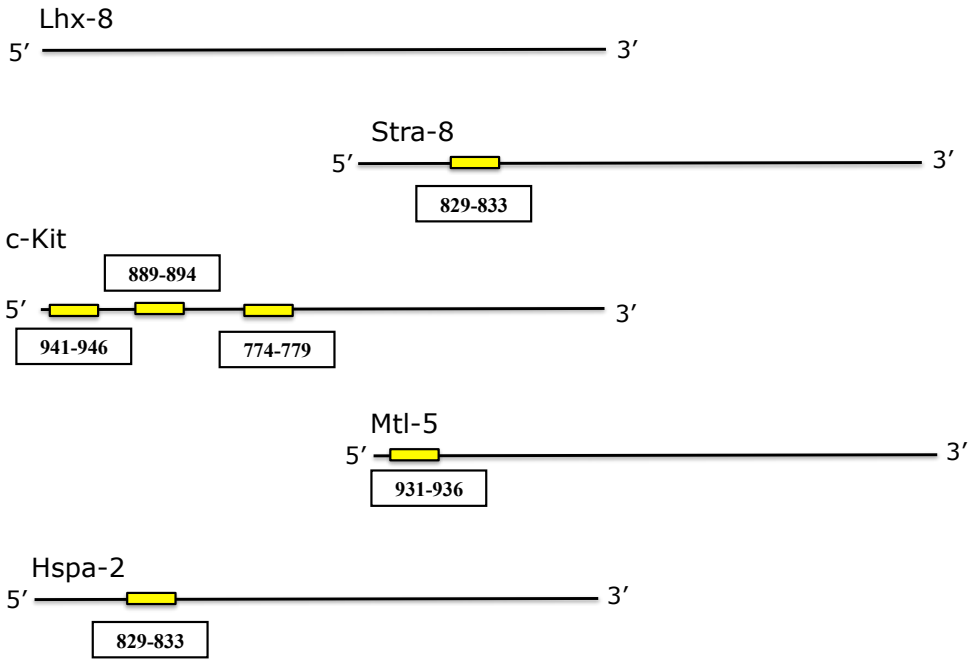


Fig 7A: Sox8 binding sites on the promoters of pre-meiotic and meiotic marker genes

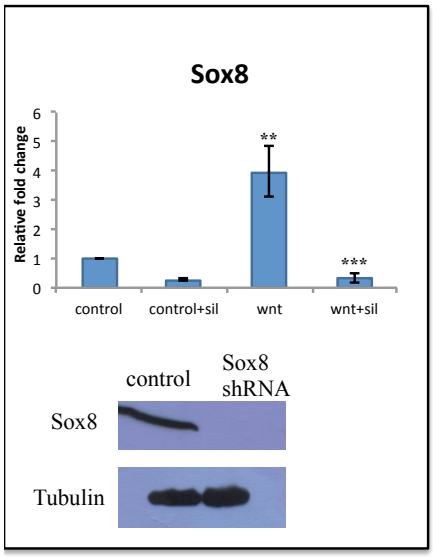
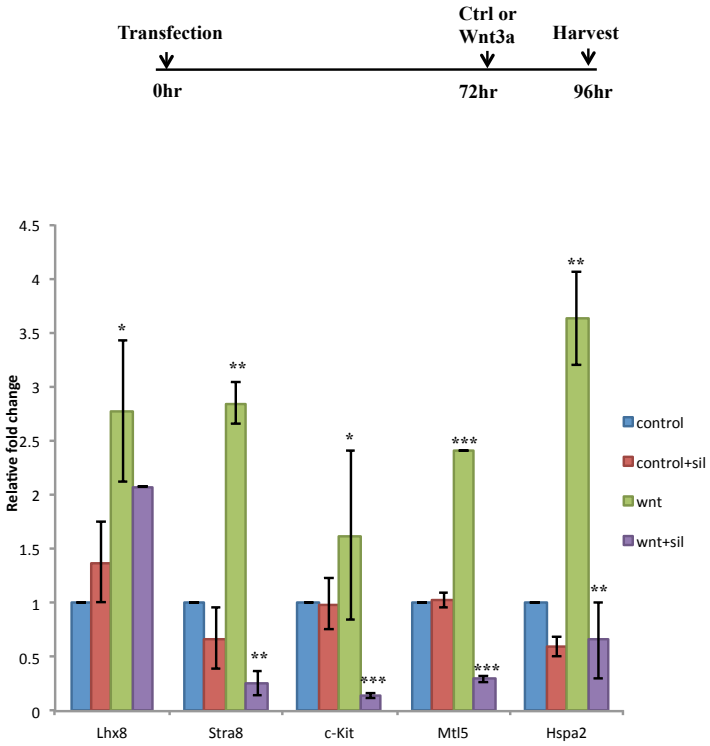


Fig 7B: qPCR analysis of meiotic markers under Sox8 silencing in control and Wnt treated GC1-spg



CHAPTER 4
DISCUSSION

Earlier report from our laboratory has shown that *mrhl* RNA physically associates at 37 loci [27] out of which one is the Sox8 genomic locus where the RNA binds to it's promoter region 141 base pairs upstream of the transcription start site. In the present study, we have explored the regulation of Sox8 by *mrhl* RNA and have identified possible players in Wnt mediated regulation of Sox8.

There have been few scattered reports on the expression pattern of Sox8 in various tissues, and expression has been reported only in Sertoli cells. In our study we find that Sox8 is expressed in Type B spermatogonial cell line – GC1-spg and is also expressed in germ cells at P7 and P21 testes. In 7-day-old testes spermatogonia is the major germ cell population and using E-cadherin as a marker for spermatogonial stem cells we see the expression of Sox8 and also in P21 testes wherein spermatocytes are the major germ cell population. We have used Scp3 as a spermatocyte marker to confirm the expression of Sox8 in germ cells at this stage. Expression of Sox8 is also observed in Sertoli cells at both these stages. Vimentin was used as a marker for Sertoli cells. The spermatogonial cell line GC1-spg and the Sertoli cell line TM4 also showed expression of Sox8.

It is known that *mrhl* RNA is negatively regulated by Wnt signaling [28] and that down regulation of *mrhl* RNA can regulate Sox8 gene expression. Therefore, we were interested in studying the regulation of Sox8 under Wnt signaling activation. We see that Sox8 is significantly up regulated on Wnt activation in germ cells as well as in Sertoli cells. The over expression of Sox8 was observed at the RNA level and also at the protein level. In testes as well Sox8 was up regulated in 21-day-old mice testes when Wnt signaling is active as compared to the 7-day-old testes when it is inactive.

Bioinformatically it was found that Sox8 promoter also has a Tcf-4 binding site at 800 base pairs upstream of the TSS in the promoter. It was found that this site is constitutively occupied by Tcf-4 and by beta-catenin upon Wnt activation in GC1-spg cell line and in testes. Even though the site is occupied by beta-catenin its role in regulation of Sox8 is negligible as compared to the occupancy of *mrhl* RNA on the promoter. It was found that removal of *mrhl* RNA from the Sox8 promoter even in control conditions can lead to up regulation of the gene whereas in Wnt conditions if *mrhl* is ectopically over expressed in

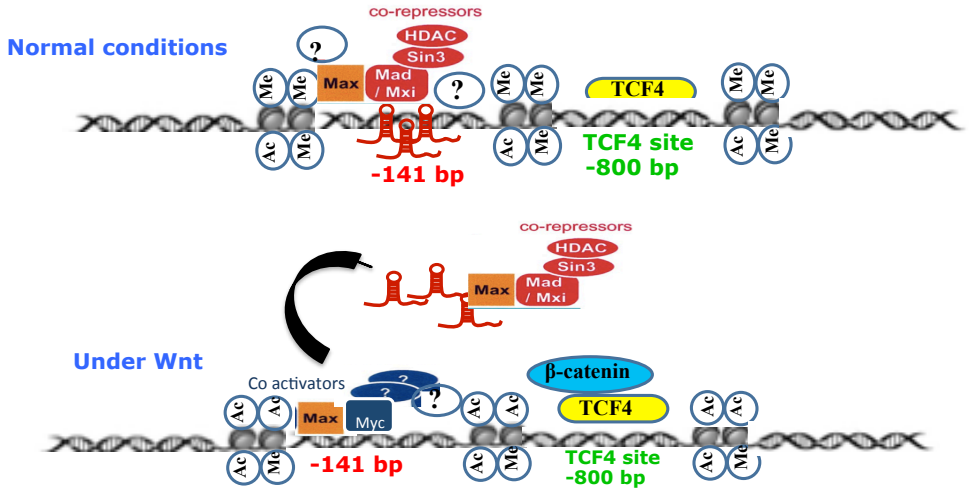
trans it can abrogate the up regulation of Sox8. It can therefore be concluded that *mrhl* RNA down regulation is the major factor in Sox8 gene up regulation.

We have shown that the Wnt mediated up regulation of Sox8 is a transcription regulation event by scoring for the pre-spliced primary transcript of Sox8. It is well known that for transcriptional activation multiple chromatin level changes are induced for efficient recruitment of the transcriptional machinery. Therefore we dissected the changes in chromatin dynamics at the Sox8 promoter locus under Wnt activation. It was observed that there is an increase in histone activation mark (H3K4me3) and decrease in histone repression mark (H3K27me3) at both the ChOP site and Tcf-4 binding site upon Wnt activation. The histone marks also lead to nucleosome depletion at both these sites thereby facilitating the transcriptional activation. Multiple transcription factors have binding sites on Sox8 promoter and interestingly all of them are clustered at the ChOP site showing the importance of this locus in Sox8 promoter regulation. Myc-Max binding site is present in Sox8 promoter and overlaps with the ChOP site. We observed c-Myc occupancy on Sox8 promoter on Wnt activation proving that it might have a role in regulation of Sox8 gene expression.

Our results conclusively show that Sox8 is not only expressed in germ cells but is also temporally regulated by Wnt signaling activation. It has been reported that Wnt signaling activation paves the path for differentiation of spermatogonia to spermatocytes by up regulation of various pre-meiotic and meiotic marker genes. Earlier result for our laboratory has shown that Wnt mediated down regulation of *mrhl* RNA is essential for up regulation of these genes [28]. Now that we know that *mrhl* RNA can in turn regulate Sox8 we were interested in seeing the role of Sox8 in regulation of these marker genes and in turn in meiotic progression. We found that silencing of Sox8 abrogates the up regulation of these marker genes showing that Sox8 can be an important transcription factor in spermatogonial differentiation.

Our study so far conclusively proves the expression of Sox8 in germ cells, its regulation by *mrhl* RNA and its role in germ cells as a transcription factor. There are still many questions we wish to address. The Myc-Max-Mad network can work in complex ways to

regulate a gene. We still have to explore the possibility of Mad-Max acting as co-repressors in control conditions as shown in figure below.



It would also be interesting to explore if the whole 2.4kb RNA or just a part of it is required for its function at the Sox8 locus. The binding of the RNA can also be brought about by various possibilities, to name a few; it can be either through a RNA-DNA triplex structure or through bridging proteins. Answering these questions would widen our knowledge on the complex and diverse ways by which lncRNAs can bring about specific spatio-temporal regulation of gene expression.

CHAPTER 5
REFERENCES

1. Griswold MD (2015) Spermatogenesis: The commitment to meiosis. Physiology Review 96:1-17
2. Anderson R, Copeland TK, SchoÈler H, Heasman J and Wylie C (2000) The onset of germ cell migration in the mouse embryo. Mechanisms of development 91: 61-68
3. Bendel-Stenzel MR, Gomperts M, Anderson R, Heasman J and Wylie C (2000) The role of cadherins during primordial germ cell migration and early gonad formation in the mouse. Mechanisms of development 91: 143-152
4. Ewen KA and Koopman P (2010) Mouse germ cell development: From specification to sex determination. Molecular and Cellular Endocrinology 323:76-93.
5. Gobel U, Schneider DT, Calaminus G, Haas R.J., Schmidt P. and Harms D (2000) Germ-cell tumors in childhood and adolescence. Annals of oncology 11: 263-271
6. McLaren, A. (1984) Meiosis and differentiation of mouse germ cells. In: Controlling Events in Meiosis: 38th Symposium of the Society for Experimental Biology. 7–23
7. Nakagawa T, Nabeshima Y-i and Yoshida S (2007) Functional Identification of the Actual and Potential Stem Cell Compartments in Mouse Spermatogenesis. Developmental Cell 12:195-206.
8. Drumond AL, Meistrich ML and Chiarini-Garcia H (2011) Spermatogonial morphology and kinetics during testis development in mice: a high-resolution light microscopy approach. Reproduction 142:145-155.
9. Lin Y,* Gill ME,* Koubova J,Page DC (2008) Germ Cell–Intrinsic and –Extrinsic Factors Govern Meiotic Initiation in Mouse Embryos. Science 322: 1685-1687
10. Raverdeau M, Gely-Pernot A, Féret B, Dennefeld C, Benoit G, Davidson I, Chambon P, Mark M, and Ghyselinck NB (2012) Retinoic acid induces Sertoli cell paracrine signals for spermatogonia differentiation but cell autonomously drives spermatocyte meiosis. Proceedings of National Academy of Sciences 109: 16582–16587

11. Baltus AE, Menke DB, Hu Y-C, Goodheart ML, Carpenter AE, de Rooij DG and Page DC (2006) In germ cells of mouse embryonic ovaries, the decision to enter meiosis precedes premeiotic DNA replication. Nature Genetics 38:1430-1434.
12. Feng C-W, Bowles J and Koopman P (2014) Control of mammalian germ cell entry into meiosis. Molecular and Cellular Endocrinology 382:488-497.
13. Bowles J, Knight D, Smith C, Wilhelm D, Richman J, Mamiya S, Yashiro K, Chawengsaksophak K, Wilson MJ, Rossant J, Hamada H, Koopman P (2006) Retinoid Signaling Determines Germ Cell Fate in Mice. Science 312: 596-600
14. Bowles J, Feng C-W, Spiller C, Davidson T-L, Jackson A and Koopman P (2010) FGF9 Suppresses Meiosis and Promotes Male Germ Cell Fate in Mice. Developmental Cell 19:440-449.
15. Saunders PTK, Turner JMA, Ruggiu M, Taggart M, Burgoyne PS, Elliott D and Cooke HJ (2003) Absence of mDazl produces a final block on germ cell development at meiosis. Reproduction 126:589-597
16. Matson CK, Murphy MW, Griswold MD, Yoshida S, Bardwell VJ and Zarkower D (2010) The Mammalian Doublesex Homolog DMRT1 Is a Transcriptional Gatekeeper that Controls the Mitosis versus Meiosis Decision in Male Germ Cells. Developmental Cell 19:612-624.
17. Kerr GE, Young JC, Horvay K, Abud HE and Loveland KL (2013) Regulated Wnt/Beta-Catenin Signaling Sustains Adult Spermatogenesis in Mice. Biology of Reproduction 90:3-3.
18. Rinn JL and Chang HY (2012) Genome Regulation by Long Noncoding RNAs. Annual Review of Biochemistry 81:145-166.
19. Wilusz JE, Sunwoo H and Spector DL (2009) Long noncoding RNAs: functional surprises from the RNA world. Genes & Development 23:1494-1504.
20. Wang Kevin C and Chang Howard Y (2011) Molecular Mechanisms of Long Noncoding RNAs. Molecular Cell 43:904-914.
21. Chen L-L and Carmichael GG (2010) Decoding the function of nuclear long non-coding RNAs. Current Opinion in Cell Biology 22:357-364.
22. Kornienko AE, Guenzl PM, Barlow DP and Pauler FM (2013) Gene regulation by

the act of long non-coding RNA transcription. BMC Biology 11:59

23. Bergmann JH and Spector DL (2014) Long non-coding RNAs: modulators of nuclear structure and function. Current Opinion in Cell Biology 26:10-18.
24. Nishant KT, Ravishankar H and Rao MRS (2004) Characterization of a Mouse Recombination Hot Spot Locus Encoding a Novel Non-Protein-Coding RNA. Molecular and Cellular Biology 24:5620-5634.
25. Ganesan G and Rao SMR (2008) A novel noncoding RNA processed by Drosha is restricted to nucleus in mouse. Rna 14:1399-1410.
26. Arun G, Akhade VS, Donakonda S and Rao MRS (2012) mrhl RNA, a Long Noncoding RNA, Negatively Regulates Wnt Signaling through Its Protein Partner Ddx5/p68 in Mouse Spermatogonial Cells. Molecular and Cellular Biology 32:3140-3152.
27. Akhade VS, Arun G, Donakonda S and Satyanarayana Rao MR (2015) Genome wide chromatin occupancy of mrhl RNA and its role in gene regulation in mouse spermatogonial cells. RNA Biology 11:1262-1279.
28. Akhade VS, Dighe SN, Kataruka S and Rao Manchanahalli RS (2016) Mechanism of Wnt signaling induced down regulation ofmrhllong non-coding RNA in mouse spermatogonial cells. Nucleic Acids Research 44:387-401.
29. Bowles J, Schepers G and Koopman P (2000) Phylogeny of the SOX Family of Developmental Transcription Factors Based on Sequence and Structural Indicators. Developmental Biology 227:239-255.
30. Kamachi Y and Kondoh H (2013) Sox proteins: regulators of cell fate specification and differentiation. Development 140:4129-4144
31. Wegner M (1999) From head to toes: the multiple facets of Sox proteins. Nucleic Acids Research 27:1409-1420.
32. Barrionuevo F and Scherer G (2010) SOX E genes: SOX9 and SOX8 in mammalian testis development. The International Journal of Biochemistry & Cell Biology 42:433-436.
33. Chaboissier MC (2004) Functional analysis of Sox8 and Sox9 during sex determination in the mouse. Development 131:1891-1901.

34. O'Bryan MK, Takada S, Kennedy CL, Scott G, Harada S-i, Ray MK, Dai Q, Wilhelm D, de Kretser DM, Eddy EM, Koopman P and Mishina Y (2008) Sox8 is a critical regulator of adult Sertoli cell function and male fertility. Developmental Biology 316:359-370.