Towards understanding molecular and cellular attributes of potential genes for hot water epilepsy

A thesis submitted for the degree of

Doctor of Philosophy

By

Sambhavi Puri

Molecular Biology and Genetics Unit
Jawaharlal Nehru Centre for Advanced Scientific Research
(A Deemed University)
Jakkur, Bengaluru 560064, India
September 2021

Dedicated to

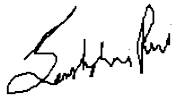
Nanu

And my family

For all the love and support

Declaration

I hereby declare that this thesis, entitled "**Towards understanding molecular and cellular attributes of potential genes for hot water epilepsy**", is an authentic record of research work carried out by me under the guidance of **Prof. Anuranjan Anand** at the Molecular Biology and Genetics Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore, India and that this work has not been submitted elsewhere for the award of any other degree. In keeping the norm of reporting scientific observations, due acknowledgements have been made wherever the work described here has been based on the findings of other investigators. Any omission, which might have occurred by oversight or misjudgment, is regretted.



Sambhavi Puri

Place: Bengaluru

Date : 27th September 2021

Certificate

This is to certify that the work described in this thesis entitled " **Towards understanding molecular and cellular attributes of potential genes for hot water epilepsy**" is the result of the investigations carried out by **Ms. Sambhavi Puri** in the Molecular Biology and Genetics Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore, India, under my guidance, and that the results presented in this thesis have not previously formed the basis for the award of any other diploma, degree or fellowship.

Prof. Anuranjan Anand

Place: Bengaluru

Date: 27th September 2021

Acknowledgements

I take this as an opportunity to thank all the incredible people who have stood along with me and supported me in this journey.

Firstly, I express my gratitude to my advisor and mentor, Prof. Anuranjan Anand. He gave me the opportunity to work in his lab and trusted me with the work on two uncharacterized proteins. He has given an exceptional amount of freedom to plan and shape the work. He has always been encouraging towards pursuing new ideas and directions. I am thankful to him for an organized lab environment and ensuring that the best of facilities are provided to the student. He always tells me to be not so worked up about everything and try to be "cool," so I try to implement that in life. The guidance provided by him throughout my Ph.D. has been a constant source of inspiration.

I thank our clinical collaborators from NIMHANS, Bangalore: Dr. P. Satishchandra and Dr. Sanjib Sinha, for their help in ascertaining the epilepsy families. I am grateful to all the families and subjects who participated in this study. I thank our collaborator, Dr. Ravi Mudashetty, and his student Bharti Nawalpuri, inStem, Bangalore, who helped us with primary neuronal culture work. I also thank them for all the fruitful discussions and valuable experimental suggestions. I thank Dr. Sheeba Vasu, JNCASR, with whom we collaborated for our *Drosophila* work. She helped us a lot with troubleshooting our experimental strategies, and her lab members assisted us with the practical expertise required for fly work. I thank Dr. Jayaram, CHG, Bangalore, for his help with karyotyping of the cell lines used in the study.

I extend my gratitude towards the faculty members of the Molecular Biology and Genetics Unit and Neuroscience Unit: Prof. MRS Rao, Prof. Tapas Kundu, Prof. Udaykumar Ranga, Prof. Namita Surolia, Prof. Hemalatha Balaram, Prof. Maneesha Inamdar, Prof. Kaustuv Sanyal, Dr. Ravi Manjithaya, Dr. James Chelliah and Dr. Kushagra Bansal for their constructive comments, discussions during departmental seminars, and at different instances of interactions I had with them. I thank my Graduate student advisory committee members: Prof. Anuranjan Anand, Dr. Ravi Manjithaya, Dr. James Chelliah, and Dr. Ravi Mudashetty, for their inputs on future directions for the work. I would like to thank Prof Michael Kibler, LMU, Germany, whom I met at a conference, and his suggestions helped shape several of the experiments.

I would like to thank Aishwarya Prakash and Harshdeep Kaur for their contributions to fly work and Sukanya Majumdar for the help with mice work. All three of them were exceptionally sincere and hardworking, and their enthusiasm motivated me to start new experiments. I thank Dr. Mohammad Atif and Utsa Bhaduri for their help with part of the sequencing analysis work. I

thank Aishwarya Iyer and Rutvij Kulkarni for their inputs on fly behavioral experiments and analysis.

I extend my thanks to all the past and present members of the lab: Dr. Manpreet Kaur, Dr. Praveen Raju, Dr. Kalpita Karan, Dr. Shalini Roy Choudhury, Sourav, Chandru, Mohan, Dr. Rammurthy, Shveta, Pooja, Shrilaxmi, Girija, Yashwini, Nazia, Sri Charani, Amrutha, Pallavi, Karan. I especially thank my seniors Shalini and Kalpita, whose projects I inherited. I am incredibly thankful to Shveta, Pooja and Shrilaxmi, for ignoring the fact that I talk to myself constantly, for their constant support when experiments didn't work (which was always), for constructive criticism on work, long hours of non-stop discussions on random things, for all the lovely outings, and for inspiring me with positivity.

I acknowledge support from the confocal, sequencing, FACS, animal house, and central facilities at MBGU, computer lab and library. I especially thank Ms. Anitha, Ms. Greeshma, Ms. Suma, Mr. Sunil, Ms. Keerthana, Dr. Narendra Nala for their help in using central facilities at MBGU. I would like to thank Mr. Jaychandra and Prof. Umesh Waghmare for showing concern about my health and progress on various occasions. I thank all the staff and facilities at JNCASR, especially at the hostel; they made life on the campus safe and comfortable. I acknowledge the financial support provided by JNCASR.

I would like to thank my colleagues and friends at JNCASR who I interacted with personally and professionally: Abhik, Ananya, Utsa, Preeti, Kajal, Arindam, Divyesh, Aksah, Diana, Rajarshi, Arnab, Vikas, Sidharth, Shreshta, Pallabi, Moumita, Smitha, Akash, Vijay, Ranbir, Priya J., Rahul, Priya B., Rohit, Hashim, Krishnendu, Shreyas, Arun, Chavi, Haider, Disha, Sushma, Chitrang, Rutvij, Arijit, Jaymin, Veena, Pradeep, Apoorva, Atreyee, Vybhav, Swati. I would like to thank Prof. Mrinal Bhattacharya, University of Hyderabad, whose guidance and exposure encouraged me to pursue a scientific career.

I am incredibly grateful to my foster family at JNCASR: Atif, Aishwarya, Pooja, and Vijaya, for not killing me every time I came crying to them, which was a lot! I thank my friends who in the past six years had only one question for me, "when am I finishing my Ph.D.?": Shruti, Sonia, Gaurav, Sandeep, Nidhi, Kittu, Soumik, Pallavi, Shalini, Shariq, Bharti, you guys motivated me a lot! I thank my uncle, Dr. Akhilesh Gupta, for being a constant encouragement. Most importantly, I thank my family: Mummy, Papa, Nanu, Nani, Dadi, Bhai, Mausi, Mosu, Jinnu, Myra. Firstly, for making sure I gained at least 5kgs when I visited home, next for not asking me about my marriage plans, and then for their endless love, support, positivity, faith, and motivation. All these people inspire me to do good, to be a better person in life, and I wouldn't have made it so far without them. Thank you for believing in me, and I hope I did all of you proud.

Table of Contents

Chapter 2: Genetic analysis of a potential causal gene for hot water epilepsy in a multi-affected family

Chapter 3: Elucidating cellular functions of R3HCC1

Chapter 4: Genetic and cell biological aspects of *ZGRF1*, a potential hot water epilepsy gene

Appendices:

Appendix 1: Supplementary Results

Appendix 2: Reagents and Protocols

Abbreviations

A	Adenine	DAPI	4',6-diamidino-2-phenylindole
C	Cytosine	DDT	Dithiothreitol
G	Guanine	DEPC	Diethyl pyro carbonate
T	Thymine	DMSO	dimethyl sulfoxide
θ	recombination fraction	DNA	deoxyribose nucleic acid
°C	degree Celsius	dNTP	deoxyribonucleotide triphosphate
cm	Centimeter	EDTA	ethylene diamine tetra-acetate
µg	Microgram	g	unit of acceleration
µl	Microliter	g	gram
µM	Micromolar	GFP	green fluorescence protein
Mb	Megabase	HCl	hydrochloric acid
mg	Milligram	MQ	Milli-Q water
mM	Millimolar	mRNA	messenger
M	Molar	NaCl	Sodium Chloride
ml	Milliliter	N-terminal	amino-terminal
mm	Millimeter	PCR	polymerase chain reaction
ng	Nanogram	pH	power of hydrogen
nm	Nanometer	PVDF	polyvinylidene difluoride
bp	base pair	RNA	ribonucleic acid
kb	Kilobase	Rpm	revolutions per minute
kDa	Kilodalton	SNV	single-nucleotide variation
cDNA	complementary DNA	TAE	tris-acetate-EDTA
CO_2	carbon dioxide	Taq	Thermus aquaticus
C-terminal	carboxy-terminal	U	units

Chapter 1

Introduction

1.1. Epilepsy and its classification

Epilepsy is a common neurological disorder where a population of nerve cells located predominantly in the central nervous system manifests abnormal, excessive, and hyper-synchronous discharges, which lead to occurrence of seizure (Lopes Da Silva et al., 2003). Due to its diverse etiologies and seizure types, epilepsy is considered to be a spectrum disorder. The severity and impact of seizures vary from person to person; for example, some people may have convulsions and lose consciousness, while others may simply have a brief lapse of awareness. Nearly 1% of the population suffers from epilepsy, in which about one-third of the patients have refractory epilepsy, wherein seizures are not controlled by two or more appropriately chosen antiepileptic medications or other therapies. Almost 75% of epilepsy starts during childhood, emphasizing the developing brain's heightened susceptibility to seizures (Hauser and Hesdorffer, 1990; Stafstrom and Carmant, 2015; Devinsky et al., 2018).

The International League Against Epilepsy (ILAE) defines epilepsy as a brain disorder complying with any of the following conditions: (i) at least two unprovoked or reflex seizures occurring more than 24 hours apart; (ii) one unprovoked or reflex seizure and a probability of further seizures similar to the general recurrence risk at least 60%, after two unprovoked seizures, occurring over the next ten years; and (iii) diagnosis of an epilepsy syndrome (Fisher et al., 2014). The ILAE has also recently proposed a multilevel classification paradigm considering the classification of epilepsy in a different clinical environment (Figure 1.1; Scheffer et al. 2017). There are three levels of classification: diagnosis of seizure type, diagnosis of epilepsy type, and diagnosis of epilepsy syndrome, if any. The first level is where a seizure type is classified into focal-onset, generalized-onset, and unknown-onset. The second level of classification takes into account the proposed definition of epilepsy, followed by categorizing it into focal, generalized, combined focal and generalized, and unknown epilepsies. Generalized seizures are a result of abnormal neuronal activity that rapidly emerges on both sides of the brain as evidenced by electroencephalography (EEG) findings. These seizures include absence, tonic, clonic, tonic-clonic, atonic and myoclonic seizures. Focal seizures, on the other hand, originate in just one part of the brain and exhibit non-motor

and motor manifestations. The person with focal seizures may present consciousness but experience sensory, motor, or psychic feelings, or have changes in consciousness, which could lead to a dream-like experience and exhibition of strange, repetitious behaviors such as twitches, blinking, and mouth movements. The combined focal and generalized group is a new addition to the classification, which takes into consideration both the seizure types. Dravet syndrome and Lennox-Gastaut syndrome are examples in which both types of seizures occur. The unknown category is where epilepsy is diagnosed, but the class cannot be determined due to unavailable or uninformative EEG studies. The third level is diagnosis for epilepsy syndromes, which is based on various features that include seizure types, EEG, and imaging features that tend to occur together. Also, it considers the age of onset, seizure triggers, and diurnal variations, and co-morbid association with intellectual and psychiatric dysfunction.

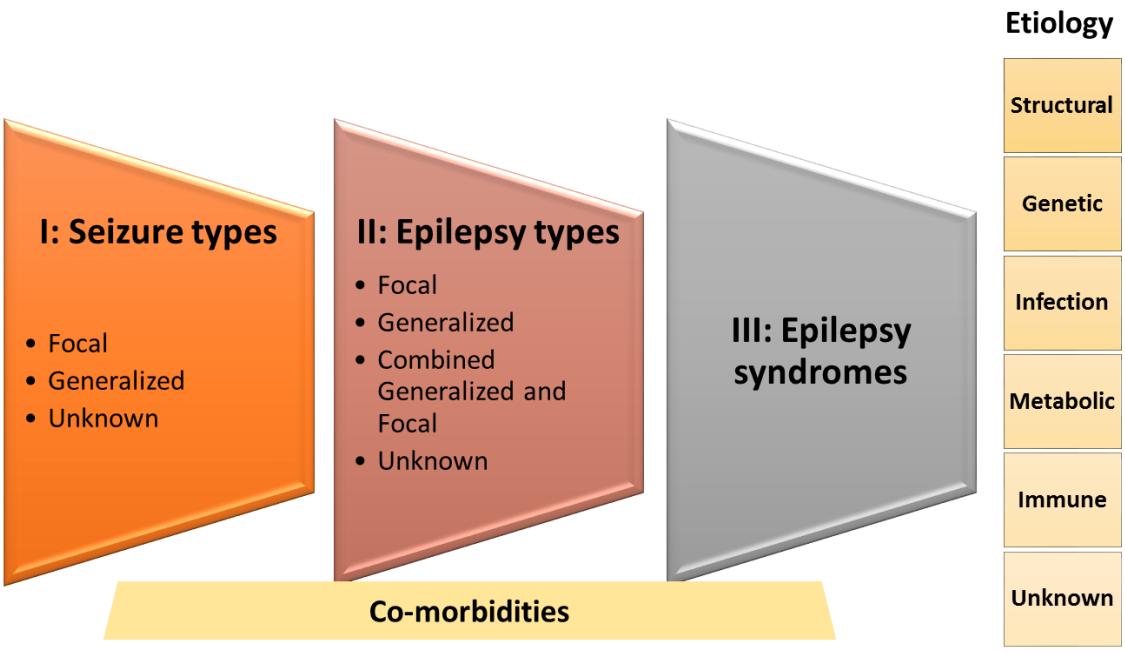


Figure 1.1. ILAE 2017 framework for classification of epilepsies (Adapted from Scheffer et al. 2017)

1.1.1. Etiology

The epilepsies can be categorized into six different etiologies. Anything that leads to disturbance in the regular neuronal activity pattern ranging from illness to abnormal brain development to brain damage can cause seizures. The six categories include genetic factors, structural brain abnormalities, infection, metabolic, immune and, unknown problem (Scheffer et al., 2017).

a) A structural abnormality can be genetic, resulting in brain development malformation or acquired due to brain tumors, head trauma, stroke, inflammation, and infection. Genetic factors contribute to the etiology of at least 35% of all epilepsy cases (McNamara and Puranam, 1998; Stein et al., 2012; Hibar et al., 2015). In 2018, the ENIGMA-Epilepsy consortium analyzed neuroimaging data of epilepsy patients from centers in 14 countries across North and South America, Asia, Europe and Australia. The study aimed at examining structural abnormalities in common epilepsies, including genetic generalized epilepsy, temporal lobe epilepsy and extratemporal epilepsy. Some of the observations included extensive alteration of subcortical volume and reduced cortical grey matter thickness. All epilepsy patient groups in comparison to controls showed lower volume in the right thalamus and lower thickness in the precentral gyri bilaterally. Subclasses of epilepsy syndrome were also studied for brain abnormalities (Whelan et al., 2018).

b) In many cases, epilepsy affects multiple related family members, indicative of a strong inherited component. In other cases, genetic mutations may occur *de novo* and contribute to the development of epilepsy among individuals with no family history of the disorder. The studies in human families and mouse models have identified more than 40 genes associated with epilepsy (Myers et al., 2019; Thakran et al., 2020). Most of these genes encode voltage-gated ion channels, entitling certain classes of epilepsy as channelopathies. In addition, a few non-ion channel genes are known, which provide evidence for the basis of novel pathways and mechanisms underlying seizure activity (Robinson and Gardiner, 2004). These include neuronal maturation and migration factors (*EFHC1, PCDH19, LIS1*) (Lo Nigro et al., 1997; Suzuki et al., 2004a; Dibbens et al., 2008), neurotransmitter release machinery (*STXBP1, SYN1*) (Khaikin and Mercimek-Mahmutoglu, 1993; Fassio et al., 2011a), neurotransmitter receptors (*GABRA1, CHRNA7*), transcription factors (*MECP2, ARX*) (Poduri and Lowenstein, 2011), and RNA-binding proteins (*NOVA2*) (Eom et al., 2013).

c) Epilepsy arising due to infections is the most common cause worldwide and can further lead to structural abnormalities. These infections include HIV, tuberculosis, cerebral malaria, cerebral toxoplasmosis, subacute sclerosing panencephalitis, and congenital infections such as Zika virus and cytomegalovirus (Vezzani et al., 2016).

d) Metabolic defects leading to biochemical changes throughout the body such as uremia, porphyria, aminoacidopathies, or pyridoxine-dependent seizures are also reported (Lee et al., 2018).
e) Autoimmune-mediated central nervous system inflammation also leads to epilepsy, contributing to the immune etiology of the disorder (Quek et al., 2012).

1.2. Reflex epilepsy

One defines an epileptic episode as a reflex seizure and epilepsy when triggered instantaneously by a well-defined motor, sensory or cognitive stimulus that can be controlled. The stimuli for inducing reflex seizures can be light, hot water, sound, which are relatively simple ones, but there can also be complex stimuli like thinking, reading, writing, and solving puzzles. A reflex seizure can be either generalized or focal, with or without impairment of consciousness, and may be genetically inherited or acquired, idiopathic, cryptogenic, or supported by a brain lesion (Striano et al., 2012). Accounting for about 1% of all human epilepsies, reflex epilepsies offer an interesting avenue to understand the basic mechanism of epilepsy.

1.2.1. Classification of reflex epilepsy

The ILAE in the year 2001 defined reflex seizures as seizures evoked by a specific afferent stimulus or by the activity of the patient (Blume et al., 2001). Reflex syndrome, on the other hand, was described as those where all epileptic seizures are precipitated by sensory stimuli (Engel, 2001; Illingworth and Ring, 2013). The proposed ILAE enlists the type of triggers for reflex seizures like visual, somatosensory, proprioceptive, listening to music, praxis, hot water, eating, startle, thinking and reading. It also states types of reflex epilepsies as idiopathic photosensitive occipital lobe epilepsy and other visual sensitive epilepsies, primary reading epilepsy, language-induced, startle, musicogenic, eating, and thinking induced epilepsy (Engel, 2001; Striano et al., 2012).

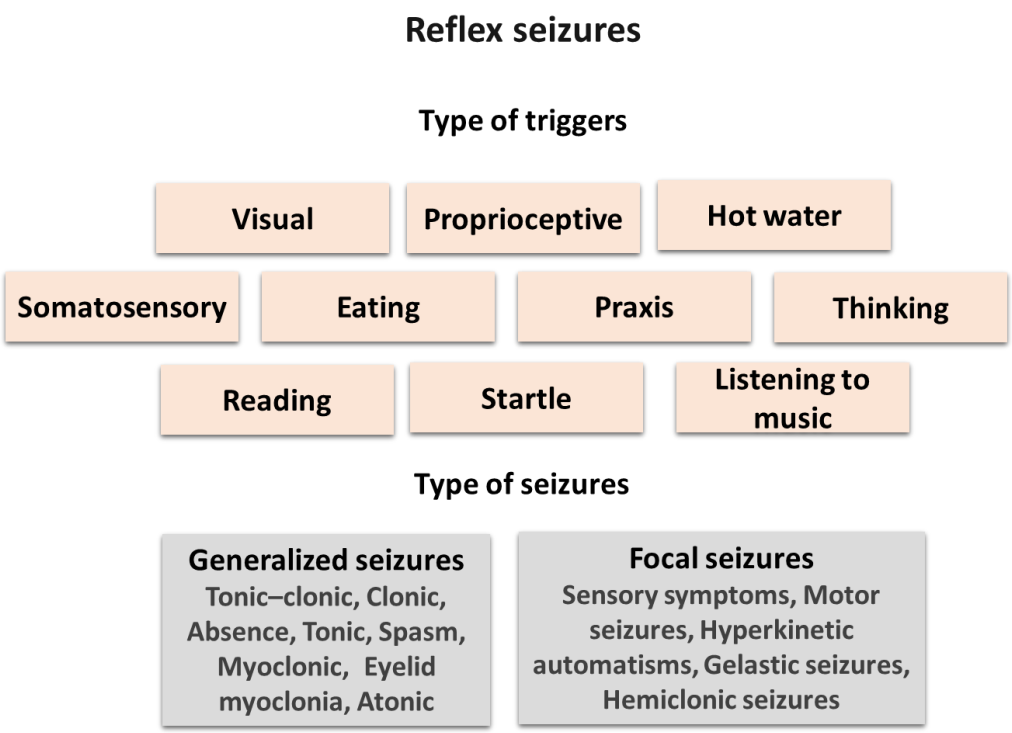


Figure 1.2. Proposed ILAE diagnostics for reflex seizures (Engel, 2001)

1.2.2. Precipitating stimulus for reflex seizures

The precipitating factor for a reflex seizure can be either a single stimulus or a repetitive one. For instance, in startle epilepsy, a single stimulus- sudden noise or surprise, can induce a seizure. These are reported in Down syndrome, Lennox-Gastaut syndrome, and dysplasia. In most cases, seizures are induced by a repetitive stimulus that can be in the form of flickering lights, loud noise, bathing in hot water, playing video games, eating, and rubbing the skin. The stimulus can also be classified as extrinsic, intrinsic, or both (Panayiotopoulos, 2005). An extrinsic stimulus can be flashes of light, elimination of visual fixation, tactile stimuli like hot water, reading, or eating. On the other hand, an intrinsic stimulus can be elementary such as movements or an elaborate one, involving higher brain functions, emotions, and cognition. These stimuli are highly variable, and susceptible individuals can have different latencies from the stimulus of onset to the clinical or EEG response.

1.2.3. Response to stimulus

Clinical and EEG manifestations alone, or in combination, are responses generated after the exposure to the stimulus. Reflex seizures may be generalized, such as myoclonic jerks or generalized tonic-clonic seizures (GTCS), or can be focal, such as visual, motor,

or sensory (Panayiotopoulos, 2005). Reflex generalized seizures are either presented independently or within the broad framework of a certain epileptic syndrome. The patient, in response to a stimulus, may exhibit various forms of generalized seizures, either in combination or alone. Myoclonic jerks, manifested in the limbs and trunk, are the most commonly observed in patients. Absence seizures are common in response to a set of stimuli such as photic, proprioceptive, cognitive, pattern, emotional, or linguistic (Panayiotopoulos, 2008). Reflex focal seizures are primarily observed in certain types of reflex focal or lobular epilepsy, such as complex focal temporal lobe seizures of musicogenic epilepsy or visual seizures of photosensitive occipital lobe epilepsy. The electroclinical event may be restricted to a brain region or spread to other cortical regions, or become generalized. An example of this can be photically- induced EEG optical spikes that eventually spread and become generalized, as in the case of photoparoxysmal responses of idiopathic generalized epilepsy (IGE). On the other hand, reading may evoke electroclinical events strictly confined to the brain regions, as observed in alexia associated with focal ictal EEG paroxysms. There is also variability in response towards the same stimulus among individuals (Panayiotopoulos, 2005).

1.2.4. Self-induction

Self-induction is a mode of seizure precipitation employed by patients to produce seizures for themselves (Ng, 2002). Self-induced seizures are provoked by producing optimal conditions of stimulation by flickering light (Andermann et al., 1962; Ames, 1971), patterns (Matricardi et al., 1990), proprioceptive stimuli (Guerrini et al., 1992), or higher brain functions (Fenwick, 1998). Even though deliberately self-induced seizures are less prevalent, they are most commonly observed in photosensitive individuals (Ames, 1971; Darby et al., 1980; Bettoni et al., 1986; Tassinari CA, Rubboli G, Rizzi R, Gardella E, 1998). Rub epilepsy is another example of self-induction where patients evoke seizures deliberately by repeatedly rubbing an area of the skin (Kanemoto et al., 2001). In India, about 30% of hot water epilepsy cases exhibit self-induction behavior (Satishchandra, 2003a). A comparative Turkish study between hot water epilepsy and photosensitive epilepsy showed a high frequency of self-induction in the HWE group (Bebek et al., 2006). Absences and myoclonic jerks are the common seizures in self-induction. Patients describe the objective of self-induction as to experience relief from anxiety and tension as well as a getaway from a disturbing situation (Panayiotopoulos, 2005).

1.2.5. Seizure suppression

In many instances, the stimulus that provokes seizures can also aid in suppressing seizures. This type of response relies on cortical activation during input; for instance, this input to neurons in their resting stage can precipitate seizure, and when epileptic discharge originates from the activated area, seizure interruption is initiated by a large group of neurons where normal action potential prevails paroxysmal depolarization shift (Wolf, 2005). Intermittent photic stimulation has been shown to inhibit seizures in 30% of focal epilepsy and 6% of generalized epilepsy patients through EEG studies (Stevens, 1962). However, a rebound has been noted after the withdrawal (Stevens, 1962). Suppression of seizure is also observed in photoparoxysmal responses during an intermittent photic stimulation when the patient is presented with pictures, listens to clicking sounds, or performs calculations. Exercising also can provoke or inhibit seizures (Kamel et al., 2014).

1.3. Types of reflex epilepsy

1.3.1. Photosensitive epilepsy

Photosensitive epilepsy (PS) is characterized by seizure generation due to light stimulus as well as more complex stimuli such as video games, television, and visual patterns. Photoparoxysmal response (PPR) is an EEG trait of spikes or spikes-waves induced by intermittent light frequency of 15-18 Hz (Hughes, 2008), and the discharges are majorly seen in the occipital lobe of the cortex (Wolf and Goosses, 1986). The prevalence of PPR is seen in about 1.6% of healthy adults, whereas 7.4–9.9% constitutes that of adult epileptics (Wolf and Goosses, 1986; Quirk et al., 1995). Its occurrence is higher in adolescents and women. The annual incidence of PS among epilepsy cases is 10% in 7- to 19-year old (Quirk et al., 1995). The most common triggers for PPR are television, video games, flickering lights like sunlight through trees, or discotheque lights. Clinical and genetic studies on families with PPR-positive epilepsy and non-epilepsy subjects have provided data for an autosomal dominant inheritance with reduced penetrance for photosensitivity. Linkage studies have identified loci conferring susceptibility to PPR at 5q35.3 (de Kovel et al., 2010a), 6p21.1 (Tauer et al., 2005), 7q32 (Pinto et al., 2005), 8q21.1 (de Kovel et al., 2010a), 13q31.3 (Tauer et al., 2005) and 16p13.3 (Pinto et al., 2005). *BRD2* has emerged as an associated candidate in photosensitive epilepsy, wherein rare SNPs have been found in photosensitive epilepsy patients in Turkey (Yavuz

et al., 2012). A study by Galizia and colleagues proposed *CHD2* as a novel gene for photosensitive epilepsy, with a higher prevalence of unique variants among patients as compared to the control cohort (Galizia et al., 2015). PPR is associated with many types of seizures, including eyelid myoclonus, generalized myoclonic jerks, generalized tonic-clonic seizures, absence seizures, and focal seizures.

The genetically determined propensity to seizures induced by light may be asymptomatic throughout life or manifest with an epileptic seizure. Genetics of PPR without epilepsy, in which one epileptic member is PPR positive, mainly indicates autosomal dominance inheritance. Genetics of PPR with epilepsy is seen in IGE, idiopathic partial, and symptomatic cases. PPR can otherwise also occur in different epilepsy syndromes, such as with juvenile myoclonic epilepsy, childhood epilepsy with occipital paroxysms, absence epilepsy, and progressive myoclonic epilepsies. In IGE, the rate of PPR is about 30%. Under the class of idiopathic partial seizures, children with neonatal seizures exhibit PPR in 44% of the cases (Doose et al., 2002). A high prevalence of photosensitivity is also encountered in about 70% of Dravet syndrome cases (Stephani et al., 2004).

In photosensitive patients, the stimulation of a critical neuronal mass in the occipital cortex is the primary site in pathogenesis. Most of the patients have intrinsic hyperexcitability of the visual cortex, which can predispose to a large-scale neuronal activation (Wilkins et al., 2004; Fisher et al., 2005; Koepp et al., 2016). In some cases, like eyelid myoclonia or absences, frontal lobe origin is noted. In a study with photosensitive baboons, after stimulation of the occipital cortex, the motor cortex was observed to produce generalized myoclonic jerks along with hyperactivity of cortico-subcortical loops, including the reticular formation and the thalamus (Silva-Barrat et al., 1986). Mechanistically, it remains unclear whether it is a defect in the excitatory or inhibitory pathway or both. Still, evidence from many groups suggests a heightened excitatory phenomenon rather than failed inhibition. A dopamine receptor agonist, apomorphine, is shown to block PPR in IGE patients, but opiate antagonists like naloxone fail to do so, suggesting the failure of the dopaminergic system (Mervaala et al., 1990). The first line of therapy considered for patients with photosensitive seizures is avoidance or modification of the stimulus. If these precautions are not enough to control, or PPR are combined with IGE, valproate is the preferred drug for treatment

(Harding et al., 1978; Rimmer et al., 1987; Verrotti et al., 2012). Self-induced seizures are handled with psychiatric or behavioral interventions.

1.3.2. Audiogenic seizures

Audiogenic seizures are induced by an acoustic stimulus. Audiogenic seizure susceptibility many times is a consequence of alterations in noise levels, acoustic deprivation and kanamycin (Bac et al., 1998), alcohol withdrawal (Faingold and Riaz, 1995), and progesterone levels (Hom et al., 1993). The first observation for audiogenic seizures was made by Ivan Pavlov in appetite-conditioning experiments on mice where the bell sound used during food presentation induced seizures in some mice unexpectedly (Brennan et al., 1997). The first genetic evidence for audiogenic susceptibility was reported in a fringe mice strain, wherein a mutation in the gene, *mass1,* was found in a recessive-locus for audiogenic seizure at chromosome 13 (Skradski et al., 2001). *Mass1* codes for a large G-protein coupled receptor (VLGR1), and a single base nonsense mutation V2250X in the gene led to the formation of a truncated protein, which resulted in susceptibility to audiogenic seizure in fringe mice. In humans, *mass1* localizes to chromosome 5q14 (Skradski et al., 2001), which interestingly is close to two critical loci: familial febrile convulsions (*FEB4*) and Usher syndrome type 2c (*USH2C*). Mice knockout for genes- *fmr1* (Chen and Toth, 2001), *interleukin6* (De Luca et al., 2004), 5HT2C receptor (Brennan et al., 1997), and *Lgi1* (Chabrol et al., 2010) result in audiogenic seizures. Triple-knockout mice of PAR bZip proteins, *Dbp*, *Hlf*, and *Tef* (Gachon et al., 2004) and double-knockout mice of *Gd3s* (GD3 synthase) and *GalNAcT* (beta-1, 4-Nacetylgalactosaminyltransferase) (Kawai et al., 2001) exhibit susceptibility to audiogenic seizures. A study has also shown the involvement of extracellular signal-regulated kinases 1 and 2 (ERK1/2) as one of the crucial contributors to audiogenic seizures in seizure-prone rats (Glazova et al., 2015). It is supported by experiments using ERK1/2inhibitor SL327, which blocks seizures. Selenoproteins were also implicated in causing severe neurological dysfunction leading to audiogenic seizures (Byrns et al., 2014). Audiogenic seizure development could be a result of either anomaly in the sensory part of acoustic impulses in the brain pathway or abnormal biochemical and physiological status of central auditory and other structures (Garcia-Cairasco, 2002). Audiogenic epileptic EEG has been demonstrated in the medulla, midbrain, and lateral geniculate bodies (Krushinsky et al., 1970; Ribak et al., 1994; Ribak and Morin, 1995; N'Gouemo and Faingold, 1999).

1.3.3. Other reflex epilepsies

Reading epilepsy is a rare form of reflex epilepsy in which seizures are induced by reading (Bickford et al., 1956). It manifests with jerks in the jaw followed by generalized seizures, with generalized or focal paroxysmal activity (Striano et al., 2012). EEG and neuroimaging studies indicate occipital-temporal regions of the dominant hemisphere predominantly involved in the generation of the seizure (Gavaret et al., 2010). In startle epilepsy, seizures are triggered by a rapid, unexpected stimulus, and response to it is mostly brief (up to 30s) and consists of axial tonic posturing, frequently causing falls. Other symptoms, such as marked automatisms, laughter, autonomic manifestations, and jerks may occur. Rarely will one observe startle-induced atonic or myoclonic jerks. Startle-induced seizures are usually observed in patients that present heterogeneity in localized or diffuse static brain pathology (Alajouanine and Gastaut, 1995). Startle-induced seizures are reported in Down syndrome, focal dysplasia, and Lennox-Gastaut syndrome (Guerrini et al., 1990; Ferlazzo et al., 2009). Eating epilepsy is a rare condition wherein seizures are most often complex-partial. However, in patients with diffused cerebral damage, a generalized seizure may also co-exist (Loreto et al., 2000). Seizures induced by food intake or proprioception were reported in three patients with Rett Syndrome (Martínez et al., 2011). *MECP2* duplication has been reported in a patient presenting eating induced epileptic spasm along with delayed motor development, absent speech, and several stereotyped midline hand movements. (De Palma et al., 2012). Seizures can also be induced by listening to various kinds of music; classical, lyric, religious, popular, military, sad, or gay, varying from one patient to another (Striano et al., 2012). Musicogenic seizures usually occur in patients with focal symptomatic or cryptogenic temporal lobe epilepsy (Italiano et al., 2016). Although spontaneous seizures are documented, it is a complex phenomenon where seizures have a temporal lobe origin with limbic system distribution (Ellis, 2017). Musicogenic seizures were reported in a 7-year-old male with Dravet syndrome who harbored a mutation in the *SCN1A* again emphasizing on comorbid associations of different epilepsy types. Other interesting types of reflex epilepsy are the ones induced by cognition-guided higher mental activities. The trigger factor in these epilepsies can be mental calculation, playing board games, or thinking and making decisions (Striano et al., 1993). Praxis induction is another phenomenon where seizures are induced by complicated movements like calculation, drawing, playing games, construction, writing, and complex finger

manipulations (Striano et al., 2012). These types of epilepsies are idiopathic and generalized in nature, characterized by myoclonic, absence, or tonic-clonic seizures (Ingvar and Nyman, 1962). It is important that reflex seizures are distinguished from provocative precipitants in a certain situation or by general internal precipitants (such as high body temperature, emotional stress, fatigue, lack of sleep, specific stages of normal sleep, and menstrual cycle) or external precipitants (such as hyperventilation, alcohol consumption or particular foods).

Table 1.1. Genetic findings for reflex seizures (adapted from Italiano et al., 2016; Okudan and Özkara 2018)

Reflex seizure type	References	Seizure type	Inheritance	Identified loci or genes
Photosensitivity	Pinto et al., 2005; Tauer et al., 2005; de Kovel et al., 2010a; Yavuz et al., 2012; Galizia et al., 2015	Absence, myoclonia, GTCS, Focal (mainly occipital)	Likely autosomal dominant with reduced, age-dependent penetrance	5q35, 6p21, 7q32, 8q21, 13q31, 16p13 *BRD2, CHD2*
Reading-induced seizures	Matthews and Wright 1967; Radhakrishnan et al., 1995	Jaw jerks that may progress to GTCS Rare: focal seizures, with alexia and variable degree of dysphasia	Likely autosomal dominant with incomplete penetrance	None
Audiogenic seizures	Ottman et al., 2004; Sanchez-Carpintero et al., 2013	Usually temporal lobe seizures	Reported in patients with ADTLE and Dravet Syndrome	*LGI1/Epitempin SCNIA*
Eating seizures	Senanayake 1990; Martínez et al., 2011; De Palma et al., 2012	Focal seizures with or without impairment of awareness	Unknown, Familial cluster in Sri Lanka, Cases associated with Rett syndrome	*MECP2* duplication
Hot water/Bathing epilepsy	Satishchandra 2003a; Ratnapriya et al., 2009b, a; Santos-Silva et al., 2015; Nguyen et al., 2015; Karan et al., 2017, 2018; Roy Choudhury et al., 2019	Focal seizures with or without impairment of awareness	Likely autosomal dominant with incomplete penetrance	10q21.3–q22.3, 4q24–q28, 9p24.3-p23, *SLC1A1, Synapsin 1, GPR56, ZGRF1*

1.4. Hot water epilepsy

One of the types of reflex epilepsy is hot water epilepsy (HWE), which is triggered, as the name suggests, by contact with hot water at 40°-50° C. HWE has been officially categorized as a reflex epilepsy type under the ILAE classification (Engel, 2001). Seizures precipitated by contact with hot water have also been described as water immersion epilepsy (Mofenson et al., 1965), bathing epilepsy (Shaw et al., 1988; Lenoir et al., 1989), and hot water epilepsy (Satishchandra, 2003b). The first case of HWE was reported in New Zealand, where a 10-year-old boy developed a staring expression, stiffened extremities, and loss of consciousness upon bathing (Allen, 1945). While there are HWE cases reported from across the world: Australia (Keipert, 1969), Turkey (Bebek et al., 2001), the United States (Stensman and Ursing, 1971), the United Kingdom (Moran, 1976), Canada (Szymonowicz and Meloff 1978), Japan (Kurata, 1979; Morimoto et al., 1985), it is from India that largest number of patients for this disorder have been reported (Mani et al., 1974b; Satishchandra et al., 1988a; Satishchandra, 2003a). The first case of HWE from India was of an 11-year old boy who presented troubling behavior in the background of impaired consciousness during hot water bathing (Mani et al., 1968). Following this, a series of 42 cases were reported with similar behavioral and clinical patterns while bathing in hot water (Mani et al., 1968). In subsequent years more cases were reported with this particular type of reflex epilepsy: 60 patients in 1972 (Mani et al., 1972) and 108 cases in 1974 (Mani et al., 1974a). This disorder was observed to be more prevalent in those parts of the country where hot water bathing following copious application of warm oil on the head is a cultural custom. In 1988, Satishchandra and colleagues published 279 HWE cases accounting for about 3.6-3.9% of epilepsy patients in south India (Satishchandra et al., 1988a). Apart from India, 21 HWE cases have been reported from Turkey, accounting for 0.6% of all epilepsies in their population (Bebek et al., 2001).

1.4.1. Clinical features and seizure phenomenology of HWE

In hot water epilepsy, it is observed that the frequency of seizures is dependent on the frequency of pouring hot water over the head. A study in 2012 reported that among 45 HWE patients from Karnataka, one-third of the patients display a 1:1 relationship of seizure episodes to hot water bathing (Meghana et al., 2012). In their lifetime, 5-10 % of HWE patients experience seizures only while taking a body bath. The seizures are focal in about two-thirds of the patients with fear, impaired cognition, gaze staring, irrelevant

speech, visual and auditory hallucinations. Other, less frequent cases exhibit generalized tonic-clonic seizures. In India, 67-80% of HWE cases (Mani et al., 1974b; Satishchandra et al., 1988b), and in Turkey, 97% of the patients presented with focal seizures with impaired cognition (Bebek et al., 2001). The seizures are manifested either at the beginning or at the end of the bath and last from 30 sec to 3 min (Satishchandra, 2003a). Even though no specific age group is defined for its onset, HWE is mainly seen in children. The mean age of seizure onset was observed to be 13±11 (mean ± standard deviation) years among Indian patients (Satishchandra, 2003b). A study of 21 HWE cases in Turkey identified the mean age of onset at 12 years (ranging from 1.5 years to 27 years), where 57% of the patients experienced their initial seizures in the first decade of their lives (Bebek et al., 2001). Self-induction is commonly observed in HWE patients, wherein a sense of pleasure provokes these patients to induce seizures by continuously pouring hot water over the head. One-third of HWE patients in Turkey reported 'self-induced' seizures (Bebek et al., 2001), whereas, in India, about 10% of HWE patients showed this tendency (Satishchandra et al., 1988b). It has been observed in several studies that patients in their later stages of lives, develop non-reflex seizures. With Indian HWE cases, 16-38% of patients developed non-reflex seizures subsequently (Mani et al., 1974b; Satishchandra et al., 1988b), while a report from Japan reported a 100% incidence of development of non-reflex seizures (Kurata, 1979). In turkey, 60-62% of HWE patients developed generalized seizures following HWE. History of febrile convulsions (FCs) prior to the development of HWE has been reported in various studies. About 5% of the HWE patients reported from India had histories suggestive of FCs before the development of HWE (Mani et al., 1974b; Satishchandra et al., 1988b). From Turkey, Bebek and colleagues (2001) observed 9.5% of patients developing FCs.

1.4.2. Electroencephalography (EEG) features of HWE

In most HWE cases, the interictal scalp EEG is usually normal with no apparent cerebral lesions or characteristic EEG abnormalities; however, 15-20% of patients show diffused EEG patterns (Mani et al., 1972). In HWE, ictal EEG recordings are difficult to obtain due to technical limitations. In few isolated cases, lateralized or localized spike discharges in the anterior temporal regions have also been reported (Szymonowicz and Meloff, 1978; Miyao et al., 1982). A study of 21 cases from Turkey documented interictal epileptogenic abnormalities in eight patients and normal EEG in another eight. The primary sites of epileptogenic activity were over the unilateral temporal region,

observed in 8 patients. Also, among four patients with non-epileptogenic abnormalities, slow-wave activity within theta range was present either diffusely bilaterally or over bi-frontotemporal, right frontocentral, or right frontotemporal regions (Bebek et al., 2001). Interictal EEG of 70 patients from India showed abnormal EEG in nine out of 45 HWE patients and six out of 25 patients having HWE with spontaneous seizure (Meghana et al., 2012). The abnormalities were majorly located in the frontotemporal region on either or both sides (Meghana et al., 2012). Ictal EEG recordings obtained in few studies during provocation with hot water have demonstrated left temporal rhythmic delta activity (Stensman and Ursing, 1971), sharp and slow waves in the left hemisphere (Moran, 1976), bilateral spikes (Morimoto et al., 1985), and temporal activity (Shaw et al., 1988; Lenoir et al., 1989). EEG recording in a patient (child) showed delta waves with a right hemisphere initial dominance, with rapid secondary generalization (Roos and Van Dijk, 1988). Another patient showed ictal EEG changes of increasing amplitude and decreasing frequency originating from the left centro-temporoparietal electrodes (Lee et al., 2000).

1.4.3. Neuroimaging in HWE

In a study in Indian HWE patients, computed tomography (CT) and magnetic resonance imaging (MRI) revealed no apparent structural abnormalities (Satishchandra et al., 1988b). A few cases reported by Szymonowicz and Meloff 1978, indicated a temporal lobe or frontal lobe abnormality. An interictal and ictal single-photon emission scan in 10 HWE patients stimulated with hot water showed hypermetabolic tracer uptake in the medial temporal structures and hypothalamus in one hemisphere with spread to the opposite hemisphere. This indicated a functional involvement of these structures in prompting HWE (Satishchandra et al., 2000). Another study by Meghana et al., 2012, indicated normal CT in all patients examined, however MRI in one out of three cases studied with HWE and spontaneous seizure for six years duration had right-sided hippocampal sclerosis (Meghana et al., 2012). Ictal single-photon emission computed tomography imaging in a study of 5 patients showed medial and lateral temporal lobe hyperperfusion in two cases (Patel et al., 2014). A case study of brain MRI on an HWE patient from Italy presented left parietal focal cortical dysplasia (Grosso et al., 2004). Another study from Turkey reported intracranial pathologies like hippocampal sclerosis, dysplasia, and a huge cystic lesion in their MRI study of five HWE patients (Tezer et al., 2006). A study intended to study functional brain connectivity in 36 patients with hot-

water epilepsy through functional MRI studies revealed that repeated seizures impact brain connectivity and that patients with frequent seizures have widespread connectivity changes, which involves the default mode network, and recruitment of several seizure-prone areas in the medial temporal lobes bilaterally (Bharath et al., 2015).

1.4.4. Pathophysiology of HWE

Pathogenesis of HWE remains mostly unknown, and a wide range of variability exists across populations regarding the temperature of the water and the way of bathing, which leads to precipitation of seizure. In India, it is observed that hot water of 45-50ºC poured over the head in quick succession induces seizures (Satishchandra et al., 1988b). Certain French patients experienced seizures with water at 37ºC, whereas a study by Auvin and colleagues (Auvin et al., 2006) reported a patient who developed seizures with water at 33ºC. In another case, seizures were reported due to bathing in hot and cold water irrespective of its temperature (Seneviratne, 2001). It has been observed that multiple specific stimuli can trigger seizures, which include heat of water, amount of water, duration of the bath, touching of water onto body or face, touching the water with hand or even having raindrops touch the face, terminating the bath, bathing in one's bathroom and just by hearing the sound of water (Bebek et al., 2001). A study involving 71 HWE patients from India showed that they all exhibited seizures upon hot water head baths. However, for some of them, seizures also precipitated while washing face with hot water, by pouring hot water over the body and for some upon running cold water over their head (Savitha et al., 2007). A history of a 1:1 relationship between epilepsy and hot water bath was obtained in 33-63% of the patients (Satishchandra et al., 1988b; Meghana et al., 2012). In laboratory conditions, it is possible to precipitate seizures by pouring hot water over the head, but inducing seizures with a hot water towel, sauna, or blowing hot air on the head has by far not been successful (Satishchandra et al., 1988b). This suggests that the triggering stimulus is complex and involves a combination of factors such as contact of the scalp with hot water, temperature of the water, and specific cortical area of stimulation. The association of HWE with such complex stimuli and the abundance of the cases are plausible in those parts of the country where hot water bathing following application of warm oil on the head is a cultural custom (Mani et al., 1972; Satishchandra et al., 1988b). Stensman and Ursing (1971) suggested that HWE is precipitated by complex tactile and temperature stimuli wherein tactile stimuli may play

a role in triggering seizures, with temperature cues providing additive effect over somatosensory stimulation.

1.4.5. Differential diagnosis of HWE

Hot water epilepsy has overlapping features with several other conditions, but careful examination along with clinical history can establish the diagnosis of the phenotype. In some cases, non-epileptic events of febrile seizures, startle reflex, or vasovagal syncope can be confused with the diagnosis of HWE. A sudden touch of hot water on the skin or scalp may induce a startling effect in due course of bathing, but this is not considered HWE (Satishchandra et al., 1988b; Bebek et al., 2001). HWE has a more extended latency period for seizure generation, which acts as one factor that differentiates it from other reflex epilepsies. Febrile seizures are common non-epileptic events that occur among infants or children between six months to six years of age (Leviton and Cowan, 1982). Around 11-27% of Indian patients exhibit HWE in association with febrile convulsions (Satishchandra, 2003b); however, so far, there has not been any link between the two. Besides, at least two other human epilepsies, generalized epilepsy with febrile seizure plus (GEFS+) and severe myoclonic epilepsy in infancy (SMEI), are induced by high body temperature. Hot water-induced seizures are also observed in an intractable form of epilepsy SMEI or Dravet syndrome, characterized by febrile convulsions before the age of one year, hemi convulsions, partial seizures, and myoclonic seizures (Dravet et al., 1992). Although hot water-induced seizures coexist with FCs, SMEI and GEFS+, a thorough clinical investigation and questioning the patients, can distinguish HWE from other events.

1.4.6. Hyperthermic kindling in HWE

The kindling effect was first described by Goddard in 1969 when he demonstrated that repeated exposure of constant non-polarizing electrical stimulus, just large enough to trigger a brief burst of epileptiform activity, can eventually lead to fully generalized behavioral convulsions (Meekings and O'Brien, 2004). One of the phenomena associated with HWE is hyperthermic kindling, wherein repeated exposure to hot water leads to progressive recruitment of convulsion in rats (Klauenberg and Sparber, 1984; Ullal et al., 1996; Jiang et al., 1999). Ullal et al., 1996 reported the manifestation of tonic-clonic seizure activity upon exposing groups of freely ambulant Wistar adult rats to a hot water jet on the head (40-55ºC) for 8 to 10 minutes. Bipolar depth electrode

recording from the hippocampus showed seizure discharges during the ictus followed by low voltage indeterminate activity and a quiescent resting phase. The intervention of hyperthermia by cooling after the ictus prevented subsequent seizure activity. They also observed a rise in hippocampal temperature following hot water stimulation in adult rats. Histological examinations of the brain of these rats exhibit degeneration of neurons in the hippocampus, brain stem, and cerebellum and also show mossy fiber sprouting, which was correlated to the number of stimulation of hot water (Jiang et al., 1999; Ullal et al., 2006).

1.4.7. Genetics of HWE

A study on humans compared subjects with no history of seizures or HWE to the ones with HWE to show that a hot water bath on the head resulted in a 0.5-0.6°F rise in oral temperature for non-HWE patients and 2-2.5°F rise in HWE patients (Ullal et al., 1996). Thereby, it was proposed that HWE patients might have an aberrant thermoregulatory system, which is sensitive to a sudden surge in external temperature. Studies have also shown that genetic factors play a major role in the development of HWE, and the identity of genes underlying this disorder is beginning to be elucidated. The study of HWE cases in India suggested familial clustering in up to 15% of the probands (Mani et al., 1974b; Satishchandra et al., 1988b), indicating the presence of a genetic component in the etiology of the disorder. So far, the genetic studies suggest that the disorder is of autosomal dominant inheritance with incomplete penetrance with two loci mapped to 10q21.3-q22.3 (*HWE1*, MIM: 613339; (Ratnapriya et al., 2009b), 4q24-q28 (*HWE2*, MIM: 613340; (Ratnapriya et al., 2009a). Mutations in *SLC1A1* are reported from India exclusively in HWE patients (Karan et al., 2017). SLC1A1 belongs to the family of neuronal high-affinity glutamate transporter. Along with other members of the family, it is shown to play a role in limiting excitotoxicity by terminating the postsynaptic action glutamate to maintain extracellular glutamate concentrations below neurotoxic levels (Kanai et al., 1995). It also plays a key role in glutamate-mediated neuroplasticity (Scimemi et al., 2009). Loss-of-function of SLC1A1 in mice leads to age-dependent neurodegeneration (Aoyama et al., 2006). These mice exhibited ventricular enlargement, cortical thinning, reduced size of the CA1 field of the hippocampus and the corpus callosum, along with some behavioral changes and cognitive impairments (Aoyama et al., 2006; Cao et al., 2012; Lee et al., 2012b). The non-synonymous variants in *SLC1A1* reported in the study affect glutamate uptake and exhibit altered glutamate kinetics and

anion conductance properties of SLC1A1 (Karan et al., 2017). A report of X-linked focal epilepsy along with reflex bathing seizure in a French Canadian family and a Latino patient with reflex bathing seizures along with epilepsy, cognitive impairment, autism spectrum disorder, both reported loss of function mutation in *Synapsin (SYN1)* (Nguyen et al., 2015; Sirsi et al., 2017). *SYN1* is an X-linked gene that codes for neuron-specific phosphoprotein implicated in the regulation of neurotransmitter release and synaptogenesis (Fassio et al., 2011b). *SYN1* knockout mice exhibit seizures from around two months of age, and seizures are typically elicited during handling procedures like cage transfer. The seizure precipitation for these mice involves variable triggers, including sensory stimuli like sudden sounds, vestibular activation, and somatosensory stimuli, and motor program activation, as well as emotional stress (Etholm et al., 2013). Another case report of a 5-year-old Portuguese boy with bilateral frontoparietal polymicrogyria who presented HWE exhibited a novel mutation in *GPR56* (Santos-Silva et al., 2015). GPR56 belongs to the family of adhesion G protein-coupled receptors (Bjarnadóttir et al., 2004). GPR56 mRNA shows selective expression in neural progenitors and hematopoietic stem cells, indicating its function in multipotent cell identity and tissue development. The GPR56 protein undergoes an autoproteolytic process that results in the formation of an N- and a C-terminal fragment named $GPR56^{N}$ and $GPR56^{C}$, respectively. Mutations in GPR56 bilateral frontoparietal polymicrogyria interferes with intracellular trafficking and higher order of glycosylation of the $GPR56^{N}$ and surface expression of $GPR56^{C}$ (Jin et al., 2007; Luo et al., 2011). Another recent study has reported mutations in *ZGRF1* in HWE patients (Roy Choudhury et al., 2019). *ZGRF1* codes for Zinc finger GRF protein 1 and is implicated in the double-stranded break repair mechanism (Brannvoll et al., 2020; Yan et al., 2021). The presence of mutations in genes involved in different cellular and molecular pathways for HWE gives insights into how different mechanisms can play a role in eliciting similar responses and the complexity of the biological system, in general.

While most seizures are apparently spontaneous in nature, reflex epilepsies are a set of intriguing neurological disorders, studies of which can provide answers to long-standing questions in the field about how a particular stimulus can rapidly precipitate seizure phenomenon. One of the obvious possibilities is that triggers reduce the threshold for more specific stimuli and cause seizures precipitation in certain people. Also, an initial incident can alter a particular network in the brain, predisposing it to hypersynchrony,

which allows specific stimuli to prevail over the threshold for seizure activity. In addition, it is more likely in some cases that an underlying genetic susceptibility has led to an engraved circuit defect, and the stimuli acts as a facilitating factor in lowering seizure threshold, thus causing seizures (Irmen et al., 2015).

Recent advancements in technologies have been a boon in the identification of genetic factors underlying different types of epilepsy. Genome-wide association studies and next-generation sequencing approaches are major contributors to this. In 2012, the National Institute of Neurological Disorders and Stroke sequenced 4000 epilepsy patients intending to find genetic mechanisms that cause epilepsy in the majority of patients. This Epi4K project centers around two areas of epilepsy, one of them being epileptic encephalopathies which are refractory and accompanied by cognitive and behavioural dysfunctions (The Epi4K Consortium, 2012). The study here mainly focused on two common types: Lennox-Gastaut syndrome and infantile spasms. These efforts identified nine genes with de novo mutations in two or more probands, among which five genes were already reported for epileptic encephalopathy. The known epileptic encephalopathy genes with mutations in several subjects were: *SCN1A*, *SCN2A*, *SCN8A*, *STXBP1*, and *CDKL5*, while other novel candidates identified included: *CACNA1A*, *CHD2*, *FLNA*, *GABRA1*, *GABRB3*, *ALG13*, *GRIN1*, *GRIN2B*, *IQSEC2*, *MTOR*, *HNRNPU*, and *NEDD4L* (Allen et al., 2013). Following this, another collaborative Epi25, comprised whole exome sequencing data sets of 13,487 epilepsy-affected and 15,678 control individuals, helped find ultra-rare deleterious variants and candidates for genetic generalized epilepsy, non-acquired focal epilepsy, and severe developmental and epileptic encephalopathies (Motelow et al., 2021).

Sadly, these studies did not include relatively less frequent cases of reflex epilepsies. However, as discussed earlier, certain reflex epilepsies do co-occur with epilepsy/epilepsy syndromes, such as Lennox-Gastaut syndrome, Dravet syndrome, and Juvenile myoclonic epilepsy, examined by these datasets. The genetic cues obtained from the large number of individuals for various epilepsies may provide insights into molecular-clinical mechanisms which underly these fascinating complex neurobehavioral disorders.

1.6. Objective of my research work

My thesis work focuses on identifying candidate genes for hot water epilepsy by employing whole genome sequencing of two multi-generation and multi-affected families from south India. Previous studies in the laboratory have suggested the potential involvement of the *R3HCC1* and *ZGRF1* genes. In addition to taking the genetics studies further, this work focuses on evaluating molecular and cell biological aspects of both the genes and some of their functional aspects.

The main objectives of the study are as follows:

1. Genetic analysis of a potential causal gene for hot water epilepsy in a multi-affected family.
2. Elucidating cellular functions of R3HCC1.
3. Genetic and cell biological aspects of *ZGRF1*, a potential hot water epilepsy gene.

Chapter 2

Genetic analysis of a potential causal gene for hot water epilepsy in a multi-affected family

A previously conducted linkage-first, exome-later - study of a multi-generation family with several of its members affected with hot water epilepsy (HWE 244) helped identify a genetic locus at chromosome 8p23-p12 and a low frequency variant, c.1292G>A in the gene R3HCC1, indicating a possible role of the gene behind HWE in the family. For a comprehensive coverage of the critical genomic interval including for the non-coding regulatory regions, I performed whole genome sequencing for one affected individual in the family. I identified a total 73,943 variants in the region 8p23-p12 and among these, only previously reported variant c.1292G>A in R3HCC1 was found to occur at a low frequency and co-segregate with the disorder in the family. To find additional HWE associated variants in R3HCC1, the complete transcript of the gene was sequenced in 288 HWE patient samples and two rare variants c.1297G>A and c.424G>A were identified. The variants in R3HCC1 were either absent or present with minor allele frequency (MAF)<0.005 in 480 control individuals from southern parts of India. In this chapter, I describe genetic studies conducted to indicate R3HCC1 as a potential candidate gene for HWE.

2.1. Introduction

A previous genetic study conducted in the laboratory examined a four-generation hot water epilepsy family (HWE244) consisting of 6 affected, 11 unaffected, and 2 asymptomatic individuals with disease phenotype segregating in an autosomal dominant manner (Figure 2.1). A parametric genome-wide linkage analysis on the family was carried out, and analysis was taken up for sub-genomic regions with LOD score values 1.0 or greater. The highest two-point LOD score obtained was 1.36 for the marker D8S277 (Chr 8p23-12) at $\theta = 0$ and 60% penetrance value. The multipoint LOD score obtained was 1.7 for the region encompassing D8S277. A 13-marker haplotype was found to be shared among all affected members and asymptomatic carriers. The recombination events were defined between D8S264 and D8S1781 (centromere-proximal) and between D8S1809 and D8S1769 (centromere-distal) in individuals III:8 and III:7, respectively. In addition to the 8p23-12 region, we also found a signal of

two-point LOD score of 1.6 for D11S908 at 11q21-23. However, this marker was not informative in individuals II:2, III:1, and III:2, and flanking marker D11S898 was also uninformative. For no other marker in the whole genome, a LOD score of 1.0 or more was obtained. In addition to the markers at 8p23-12 and 11q21-q23, 10 microsatellite markers across the genome had positive LOD score, with the highest score being 0.62 for D3S1311 and the lowest, 0.14 for D4S412 at $\theta=0$. These markers segregated among four affected individuals but not among two affected individuals and were present in the unaffected individual, II:2, belonging to a separate branch marrying-in the family. Taken together, these observations suggested a region of about 29 Mb at chromosome 8p23-12 to be critical for further analysis. Whole exome sequencing of four individuals- three affected individuals: III:8, III:9, III:2, and one asymptomatic carrier II:2 was carried out. The total number of variants common in these four individuals in the critical region were 1145, out of which 516 were heterozygous. From among these, novel and rare variants (MAF< 0.005) were taken up for further analysis of segregation with the disorder in the family and for their frequency in 200 ethnically matched control individuals. The data revealed one non-synonymous variant, c.1292G>A in the gene, *R3HCC1* (R3H and coiled-coil domain containing 1), co-segregating with the hot water epilepsy in HWE244 (Kalpita R. Karan, Ph.D. Thesis, 2015).

To analyze potential causative variants in other regions of the genome such as UTRs, enhancers, promoter region, non-coding RNA regions, I performed whole genome sequencing (WGS) for the patient, III:4. In this chapter, I provide details of the WGS and Sanger-based sequencing studies and propose *R3HCC1* as a potential candidate gene for HWE.

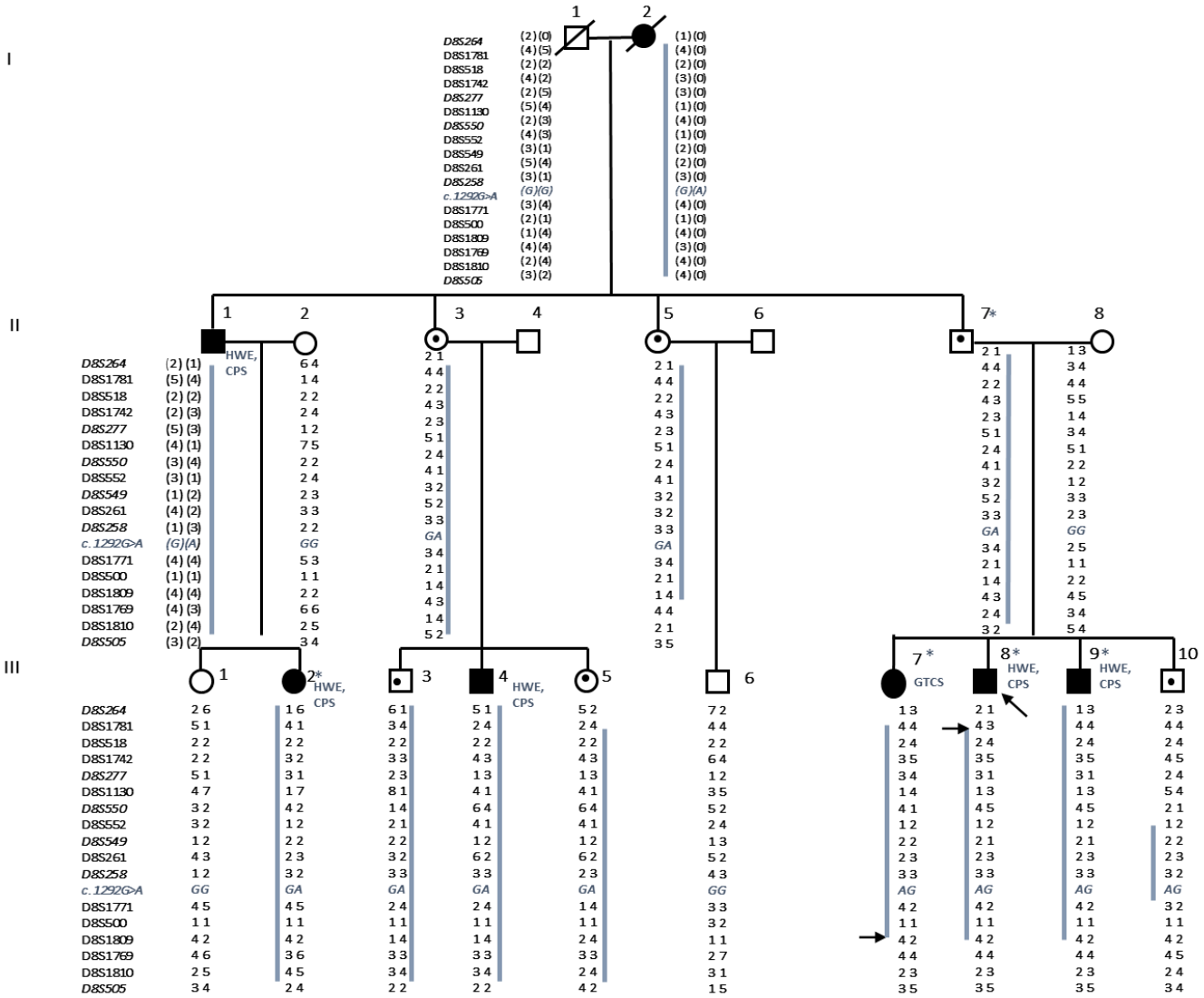


Figure 2.1. A 13-marker haplotype segregation in HWE244: Males are denoted by squares, circles denote females, affected individuals are shown with filled symbols, and empty symbols represent unaffected members. A symbol with a dot depicts an asymptomatic carrier. Proband is denoted by an arrow pointing towards the symbol. The microsatellite markers shown towards the left, span the 8p23-p12 region from the telomere to centromere. The haplotype segregating (solid blue bar) including the *R3HCC1* variant, c.1292G>A, is shown for the members. CPS, complex partial seizure; GTCS, generalized tonic-clonic seizure, HWE, hot water epilepsy. *denotes individuals whose exomes were examined.

2.2. Material and methods

2.2.1. Sequencing of the uncovered exons in the region 8p23-p12

To identify any potentially 'meaningful' variants which may have been missed due to the lack of probes in the exome enrichment experiments, 90 coding exons in different genes in the region 8p23-p12 which remained uncovered or had low coverage in exome sequencing data, were examined. The sequences of the uncovered exons were obtained from Genbank. Primers were designed using Primer 3 and Oligonucleotide Properties Calculator (OligoCalc) and obtained from Sigma Aldrich (Missouri, USA) (A.2.1). PCR amplification was done for DNA of proband (III:8) of HWE 244 family. The amplified fragments were purified using 96 multiscreen filter plates and sequenced on an ABI 3730 Genetic Analyzer (Thermo Fisher Scientific, Massachusetts, USA). The protocol details for PCR and Sanger sequencing are given in A.2.13-A.2.16. To find variants in the proband, the sequences obtained were compared with their respective reference sequence obtained from Genbank using SeqMan II v5.01. The variants were checked for their MAF in the following databases: dbSNP244, 1000 Genomes, Exome variant server, gnomAD v2.1.1 and v3.1.1. The novel and rare heterozygous variants identified for uncovered exons were screened for segregation in the family.

2.2.2. Whole genome sequencing

Whole-genome sequencing was done for the affected individual III-4 using NEB NEXT Ultra II DNA library prep kit (New England Biolabs, Massachusetts, USA). For this, 1000ng of intact DNA was enzymatically fragmented by targeting 200bp fragments size. The DNA fragments were subjected to end repair; wherein the overhangs were converted to blunt ends. To the blunt-ended fragments, adenylation was performed at the 3' ends. The adenylated fragments were ligated to adapters and enriched with PCR using NEBNext Ultra II Q5 master mix. The amplified products were cleaned up by using AMPure beads, and the final DNA library was eluted in 15µl of 0.1X Tris EDTA pH 8.0 buffer. The sequencing libraries obtained had an average insert size of 350bp (A.1.1). The library quality and quantity were checked using the Qubit HS dsDNA kit (Thermo Fisher Scientific) and Agilent Bioanalyzer DNA 1000 kit (Agilent Technologies, California, USA). The quality control (QC) passed libraries were sequenced for 150 bp paired-end reads using the Illumina HiSeqX platform. The sequence data quality was checked using FastQC and MultiQC. The data were also checked for base call quality

distribution, percentage of bases above Q20, Q30, percentage of GC content, and sequencing adapter contamination (A.1.2). The sequence data was processed using trimgalore to remove adapter sequences and low-quality reads. The QC passed reads were mapped to human reference genome build GRCh38 that was provided in GATK Resource Bundle using BWA-MEM algorithm (Li and Durbin, 2009). The alignments were sorted, indexed and PCR duplicates were marked and removed using Picard tools. GATK (McKenna et al., 2010) work-flow for short variant discovery (SNPs + INDELs) using GATK v4 was followed for variant calling. The variants were annotated using the Variant Effect predictor (McLaren et al., 2016). The variants were also annotated separately using the gnomAD v3.1.1 data release.

2.2.3. Sequence validation and segregation

Primers were designed to validate the rare variants obtained from the whole genome sequencing study (A.2.2). All amplicons were sequenced in the affected individual III-4 to check if the variant is a true variant, followed by checking for co-segregation in all the members of the family. The segregating variants were examined in ethnically matched control individuals to examine their allele frequencies.

2.2.4. Screening for *R3HCC1* variants in HWE patients and controls

The candidate variant, c.1292G>A in *R3HCC1* was screened in 480 ethnically-matched control individuals to check for its minor allele frequency. To gather further genetic evidence for the potential role of *R3HCC1* in HWE, all exons, flanking intronic boundaries, and the 5'- and 3'- untranslated regions of the gene were screened for additional variants in a set of 288 unrelated HWE patients. These patients were ascertained at the National Institute of Mental Health and Neurosciences and were diagnosed according to ILAE guidelines for epilepsy classification (Engel, 2001; Fisher et al., 2014). The same regions for *R3HCC1* were also examined in the above-mentioned 480 ethnically-matched individuals. DNA was isolated from peripheral blood samples of the patients and the controls using the phenol-chloroform method (Sambrook and Russel 2001). Primers were designed using Primer 3 and OligoCalc for the seven coding exons, flanking intronic segments covering splice junctions, and 5'- and 3'- untranslated region (UTRs) (A.2.3). The PCR amplifications and sequencing were performed, and sequences were compared with the *R3HCC1* reference sequence (ENST00000265806) using SeqMan II v5.01 and analyzed for variations.

2.2.5. Bioinformatic analysis

The nomenclature of *R3HCC1* variations obtained in the WGS study corresponds to Ensembl transcript ENST00000265806. NCBI reference codes for the R3HCC1 gene, mRNA, and protein are NC_000008.11, NM_001136108.3, and NP_001129580.2, respectively (NCBI assembly GRCh38, annotation release 108). The complete nucleotide and protein sequence of *R3HCC1* was downloaded from NCBI, GRch38 assembly, annotation release 107. NCBI conserved domain prediction software was used to identify the presence of any conserved domains in the protein. SMART tool was used to predict low complexity regions in the protein. Multiple sequence alignment using Clustal Omega (EMBL-EBI, https://www.ebi.ac.uk/Tools/msa/clustalo/) was performed for R3HCC1 protein sequences from various orthologs to study the conservation of the protein; its domain, and the variants present in it across species. To gain further insights into the potential effects of the variants on the function protein, two prediction tools, Polyphen-2 and SIFT, were used. PolyPhen-2 (Polymorphism Phenotyping v2) predicts the possible impact of amino acid substitutions on the bases of sequence-, phylogenetic- and structural- information characterizing the substitution. SIFT (Sorting Intolerant From Tolerant), on the other hand, considers sequence homology and the physical properties of the amino acid change.

2.3. Results

2.3.1. Whole genome analysis

In the WGS study, we obtained an average read depth coverage of 33.8X and for the 8p23-p12 critical region, coverage was 34.67X (Table 2.1). The variants in the critical region were taken up for further analysis. The region harbored 73,943 variants. These variants were filtered out based on various parameters, such as removal of duplicate reads, variants in predicted transcripts, and exclusion of heterozygous variants with MAF>0.005. The filtered rare variants were segregated into two categories: variants that are present in protein-coding genes and variants in the non-coding/regulatory regions. Among the 547 variants in the protein-coding genes, I obtained the following: 8 exonic variants, 8 UTR variants, and 10 intronic variants within 100bp of intron-exon boundaries (Figure 2.2). These 26 variants were validated by Sanger sequencing in the patient III:4 and checked for their co-segregation with HWE in the family (Table 2.2). Among these, two variants were found to segregate: one c.1292G>A (p.Arg431Gln) in

R3HCC1, and c.879+42G>C in *C8orf58*. Both the variants were examined in the control set of 480 ethnically matched individuals. MAF for the *R3HCC1*, c.1292G>A was 0.001, and that for the C8orf58, c.879+42G>C was 0.01.

2.3.2. Analysis of non-coding regulatory variants

I found 482 non-coding regulatory variants with MAF<0.005 (Figure 2.2). Analysis of these regulatory regions is important since both *cis*-regulatory and non-coding RNA may have a functional impact as they regulate chromatin organization, epigenetic modulation, transcriptional, post-transcriptional, translational, and protein localization processes. The variants situated in the promoter regions, promoter-flanking regions, CCCTC-binding factor (CTCF) binding sites, enhancer regions, open chromatin sites, and transcription factor (TF) binding sites were checked for conservation across species. The regulatory variants in long non-coding RNA (lncRNA), micro-RNA, small nucleolar RNA were further filtered by selecting variants present in the non-coding exonic regions and intronic regions within 100bp. A total of 202 variants were checked for conservation using genome Asia UCSC for Phylo P scores in 30 mammals and 100 vertebrates. Twenty three conserved variants included three variants in the enhancer region, two in CTCF binding site, one in open chromatin region, seven in the promoter region, three in promotor flanking region, one in TF binding, six in lncRNA and one in miscellaneous RNA (misc RNA). Upon analysis for their pathogenic effect using FATHMM-XF prediction software (http://fathmm.biocompute.org.uk/fathmm-xf/) (Rogers et al., 2018) (Table 2.3), two variants were predicted as pathogenic but did not co-segregate with HWE in the family.

Table 2.1. Whole genome sequencing coverage

Alignment details	Whole genome	8p23-p12 region
Total number of reads	731,340,130	
% Reads aligned to complete genome	99.96	
Target sequence length	3.2GB	28MB
Number of mapped reads	731,076,301	6,754,254
% of target covered with at least 10X read depth	91.74	94.60
% of target covered with at least 20X read depth	85.28	92.53
% of target covered with at least 30X read depth	68.02	76.50
Average read depth (X)	33.80	34.67

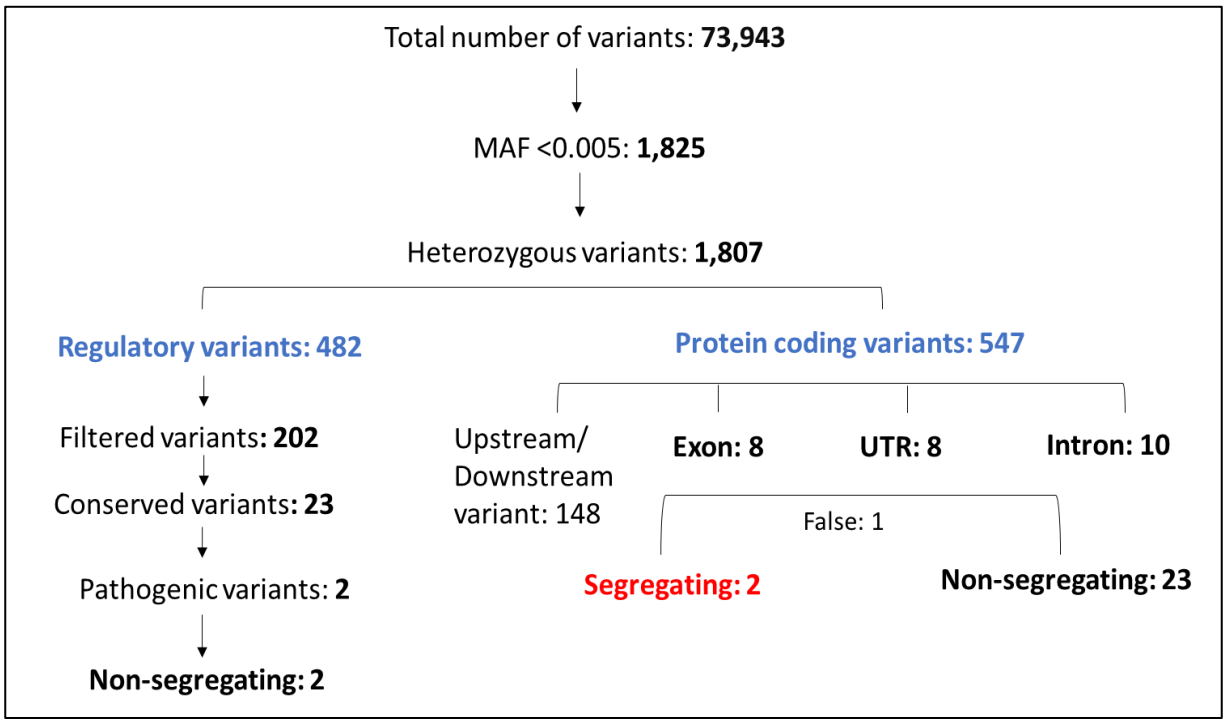


Figure 2.2. Analysis of variants in the 8p23-p12 region: The flow-chart represents the number of variants present at each step during the filtration process. The variants with MAF less than 0.005 are further selected for heterozygosity and then filtered based on their annotation and segregation with the disease phenotype in the family.

Table 2.2. Twenty six heterozygous variants identified from WGS analysis with MAF <0.005.

Nucleotide position	Gene	Reference	Amino acid change	Biotype of SNV	True Variant	Family segregation	rs ID	MAF	MAF Inhouse controls
28780956	INTS9	c.1074G>A	p.I358I	synonymous_variant	Yes	-	rs575319275	0.0004	-
6757159	AGPAT5	c.870-4A>G	-	splice_region_variant & intron_variant	No	-	rs200007113	0.0002	-
10468964	RP1L1	c.2644G>A	p.R882W	missense_variant	Yes	No	rs148936402	0.0008	-
8703038	CLDN23	c.640G>A	p.E214K	missense_variant	Yes	No	rs541030083	0.0001	-
12133733	USP17L7	c.277C>T	p.V93I	missense_variant	Yes	No	-		-
19961052	LPL	c.1291A>T	p.I431F	missense_variant	Yes	No	rs775665524	0.000008	-
22122506	HR	c.2108T>A	p.D703V	missense_variant	Yes	No	rs755106491	8.315E-06	-
22148921	LGI3	c.886C>T	p.V296M	missense_variant	Yes	No	rs141783437	0.000005	-
23296066	**R3HCC1**	**c.1292G>A**	**p.R431Q**	**missense_variant**	Yes	**Yes**	**rs1051804607**	**0.00001**	**0.001**
10538466	PRSS55	c.742-10C>G	-	intron_variant	Yes	No	-		-
11555568	BLK	c.772+84G>A	-	intron_variant	Yes	No	rs1295964815	3.188E-05	-
11867953	CTSB	c.-145+48G>A	-	intron_variant	Yes	No	rs867826953	0.00006	-
21971859	XPO7	c.427-17A>T	-	intron_variant	Yes	No	rs752685991	0.00001	-
22069564	DMTN	c.394+46A>G	-	intron_variant	Yes	No	rs376609490	0.0002	-
22101426	FAM160B2	c.1617-12del	-	intron_variant	Yes	No	rs772161842	0.0003	-
22602354	C8orf58	c.879+42G>C	-	intron_variant	Yes	Yes	rs760433506	0.00001	0.01
24955413	NEFL	c.1044+52_1044+58del	-	intron_variant	Yes	No	-		-
27454136	PTK2B	c.2596-18C>T	-	intron_variant	Yes	No	rs369795324	0.0002	-
22588859	PDLIM2	c.-517C>T	-	5_prime_UTR_variant	Yes	No	-		-
27479248	CHRNA2	c.-563_-562del	-	5_prime_UTR_variant	Yes	No	-		-
8785274	MFHAS1	c.*748C>G	-	3_prime_UTR_variant	Yes	No	rs139552636	0.001	-
11844603	CTSB	c.*522G>A	-	3_prime_UTR_variant	Yes	No	-		-
16109444	MSR1	c.*641G>C	-	3_prime_UTR_variant	Yes	No	-		-
24951539	NEFL	c.*1271A>G	-	3_prime_UTR_variant	Yes	No	-		-
29070394	KIF13B	c.*110C>T	-	3_prime_UTR_variant	Yes	No	rs142859871	0.0008	-
31174737	WRN	c.*1635C>G	-	3_prime_UTR_variant	Yes	No	-		-

Minor allele frequency (MAF) is cumulative of data in dbSNP 224, Ensembl, TOPMED, ExAC, gnomAD v2.1.1 and v3.1.1, NR: Not reported.

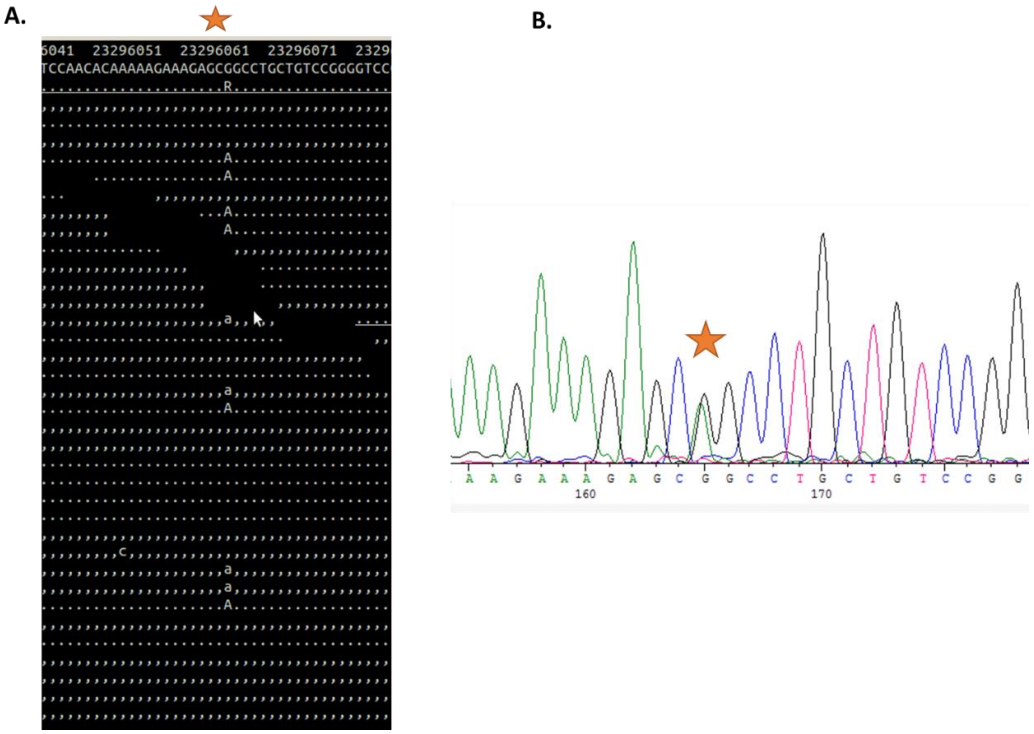


Figure 2.3. Segregating rare variant in family HWE 244: (A) Alignment snapshots of the c.1292G>A variation in individual III:4. The heterozygous variation is seen in paired-end reads (a combination of lower case and uppercase variants) at high read depths. The reference allele is G, and the minor allele is A, at the coordinate chromosome 8-23296066. The allele position is marked by an asterisk.(B) Representative image of electropherogram of c.1292G>A change in *R3HCC1* in the patient.

Table 2.3. Twenty three conserved non-coding regulatory heterozygous variants identified from WGS analysis with MAF<0.005

Nucleotide Position	rs ID	Reference	Change	Symbol	Biotype of SNV	MAF	FATHMM-XF prediction	Family Segregation
16478055	rs568887146	A	T	-	enhancer	0.0006	benign	-
17941675	rs566732276	G	A	-	enhancer	0.0004	pathogenic	No
27419030	-	G	GA	-	enhancer	NR	neutral	-
8667290	-	G	A	-	CTCF_binding_site	NR	benign	-
29339910	rs571540595	G	A	-	CTCF_binding_site	0.0002	benign	-
16383224	-	T	C	-	open_chromatin_region	NR	benign	-
8703038	rs541030083	G	A	-	promoter	NR	benign	-
12665696	-	C	T	-	promoter	NR	benign	-
15540007	-	G	C	-	promoter	NR	benign	-
30156271	-	G	C	-	promoter	NR	benign	-
26046335	rs567142733	G	C	-	promoter	0.0002	benign	-
30156281	-	G	C	-	promoter	NR	benign	-
19003840	rs533533696	T	C	-	promoter_flanking_region	0.0004	benign	-
28875941	-	A	G	-	promoter_flanking_region	NR	benign	-
15701816	-	TTAAAG	T	-	promoter_flanking_region	NR	neutral	-
12543053	rs532583533	C	T	-	TF_binding_site	0.0008	benign	-
4325009	rs554907435	C	G	AC027251.1	lncRNA	0.0002	Benign	-
10037982	-	T	C	LOC105379234	lncRNA	NR	Benign	-
10276457	-	T	A	AC079200.1	lncRNA	NR	pathogenic	No
11649152	rs774077150	C	G	LOC105379242	lncRNA	NR	benign	-
16383224	-	T	C	LOC101929028	lncRNA	NR	benign	-
15900183	-	T	TA	AC018437.3	lncRNA	NR	neutral	-
7216421	-	A	C	LOC101928095	misc_RNA	NR	benign	-

Minor allele frequency (MAF) is cumulative of data in dbSNP 224, Ensembl, TOPMED, ExAC, gnomAD v2.1.1 and v3.1.1, NR: Not reported

2.3.3. *R3HCC1* variants among additional HWE patients

The *R3HCC1* transcript was examined in additional 288 HWE probands and 18 variants were found. Since the phenotype is inherited in an autosomal dominant manner, only heterozygous variants were considered for further analysis. Out of 10 heterozygous variants, two novel/rare variants were found: c.424G>A (p.Val142Met), c.1297G>A (p. Ala433Thr), and the HWE244 variant, c.1292G>A (p.Arg431Gln) in another epilepsy patient (Figure 2.4 A, B). These three non-synonymous variations were either absent or rare among 480 ethnically matched control individuals as well in databases (Table 2.5). Fifteen polymorphisms were found in the gene (Table 2.5). The *R3HCC1* transcript was also sequenced in 480 control individuals and the list of variants found are tabulated in A1.3. The residues Val142, Arg431, and Ala433 were conserved in higher mammals (Figure 2.5 C). Polyphen-2 and SIFT predict the p.Arg431Gln and p. Ala433Thr to be tolerated, whereas p. Val142Met is predicted to be damaging by Polphen 2.

2.3.4. R3HCC1: a potential RNA-binding protein

R3HCC1 encodes a transcript of 1573 bases and protein of 440 amino acids (Gene ID: 203069). It comprises seven coding exons (Figure 2.4 A, B). It has another isoform which is predicted to code for a protein of 398 amino acids. Bioinformatics tools predicted that R3HCC1 consists of two prominent domains: an N-terminal, R3H domain, and a C-terminal, RNA recognition motif (RRM). R3H domain comprises conserved RXXXH motif along with a stretch of hydrophobic residues, proline, and glycine, and this motif folds into a three-stranded antiparallel beta-sheet, against which two alpha-helices are packed (Liepinsh et al., 2003). RRM belongs to an RNA -recognition motif super-family, containing 90 amino acids, which form a sandwich structure of four anti-parallel beta-strands and two alpha-helices (Kenan et al., 1991). Both these domains are known to bind to RNA (Wu et al., 2005b). R3HCC1 also consists of three low complexity domains, predicted by SMART analysis tool. Low complexity domains are known to be present in RNA-binding proteins that are part of ribonucleotide protein complexes and are involved in protein-protein or protein-RNA interactions (Jonas and Izaurralde, 2013; Hennig et al., 2015; Protter and Parker, 2016). These domain predictions suggested that R3HCC1 is an RNA-binding protein.

Table 2.4: Novel and rare (MAF<0.005) variants in 288 HWE patients

Exon	Nucleotide position	Reference	Het/Hom	Biotype of SNV	Amino acid change	rs ID	MAF	MAF Inhouse Control	Polyphen 2	SIFT
2	23288582	c.59A>G	Het	missense_variant	p.H20R	rs11546682	0.13			
4	23290051	c.434G>A	Hom/Het	missense_variant	p.R145K	rs3808536	0.53			
4	**23290041**	**c.424 G>A**	**Het**	missense_variant	**p.V142M**	**NR**	**NR**		**Probably Damaging**	**Tolerated**
5	23291417	c.909A>G	Hom/Het	synonymous_variant	p.T303T	rs2272762	0.44			
5	23291427	c.919G>A	Hom/Het	missense_variant	p.V307M	rs2272761	0.45			
5	23291576	c.1025+43A>G	Hom/Het	intron_variant		rs2272759	0.44			
5	23291556	c.1015+23C>T	Het	intron_variant		NR	NR			
5	23291465	c.957G>T	Het	synonymous_variant	p.V319V	rs2272760	0.13			
6	23293487	c.1096+114A>G	Hom/Het	intron_variant		rs3860836	0.55			
6	23293265	c.1026-	Hom	intron_variant		rs375807313	NR			
6	23293364	c.1087C>T	Het	synonymous_variant	p.L363L	rs545985863	T<0.01			
6	23293365	c.1088T>G	Hom/Het	missense_variant	p.L363R	rs13530	0.55			
7	23294821	c.1149A>G	Hom/Het	synonymous_variant	p.T383T	rs15946	0.78			
7	23294920	c.1192+56A>C	Het	intron_variant		rs11779438	0.0196			
7	23294758	c.1097-11T>G	Het	intron_variant		rs190502477	0.004			
8	**23296066**	**c.1292 G>A**	**Het**	missense_variant	**p.R431Q**	**rs1051804607**	**0.000007**	**0.001**	**Benign**	**Tolerated**
8	**23296071**	**c.1297 G>A**	**Het**	missense_variant	**p.A433T**	**rs773005137**	**0.0001**		**Benign**	**Tolerated**
8	23296145	c.*48G>A	Het	3_prime_UTR_variant		rs144495060	0.01			

MAF is cumulative of data in dbSNP224, Ensembl, TOPMED, ExAC, gnomAD. NR: Not reported

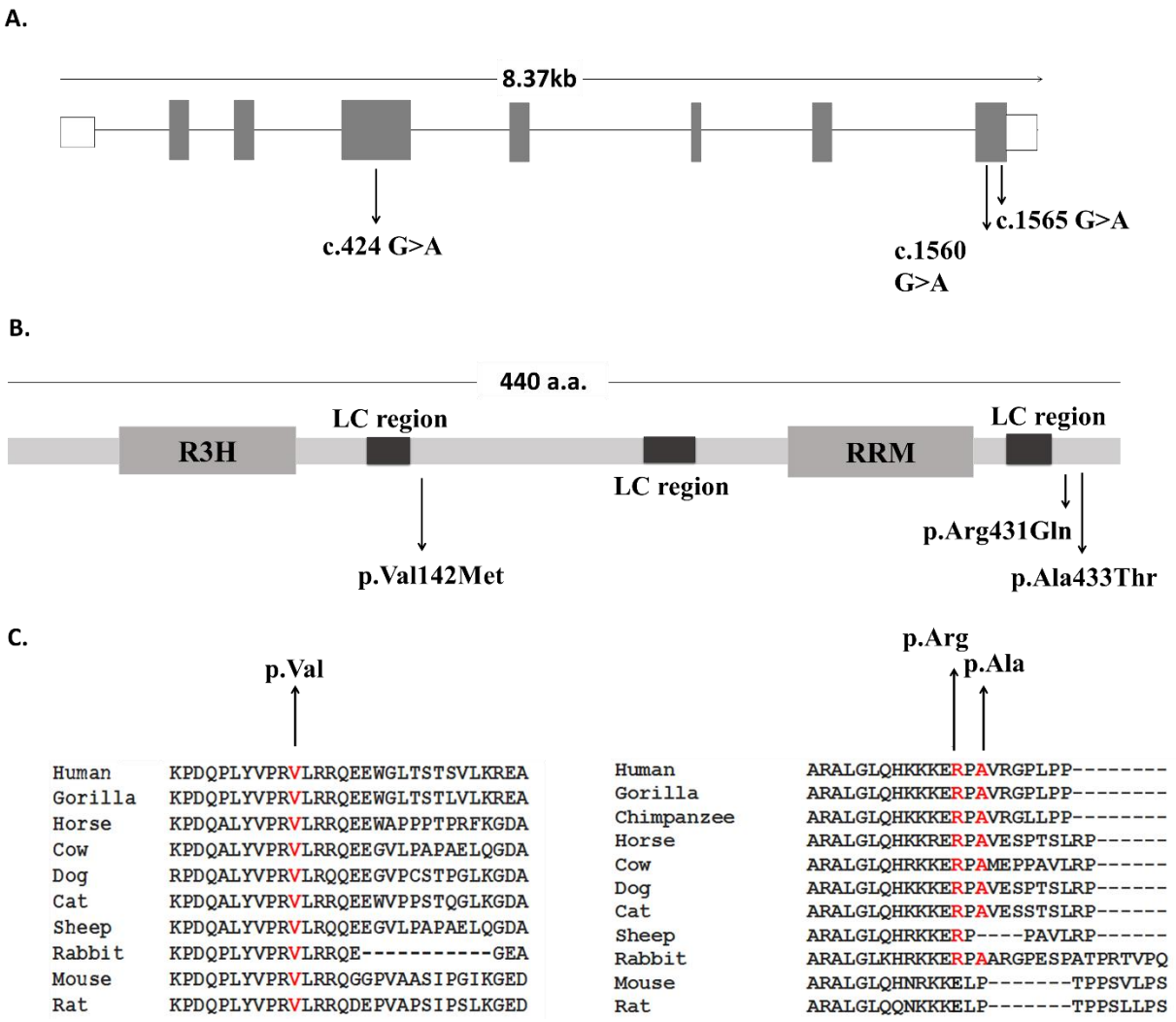


Figure 2.4. Potential causative variants in R3HCC1: (A) Schematic of *R3HCC1* gene structure for longest isoform comprising eight exons. Marked with an arrow are locations of the variants found in the patients at the genome level: c.424G>A, c.1560G>A, c.1565G>A, (B) R3HCC1 protein schematic for 440 amino acids (a.a.) showing domains present, R3H domain (19-80 a.a.), low complexity region (114-131 a.a., 257-268 a.a., 413-422 a.a.) and RRM domain (318-382 a.a.) arrow highlights locations of the variants found in the patients at the protein level and (C) Conservation across ten species, of the key amino acids mutated: p.Val142, p.Arg431 and p.Ala433 (highlighted in red).

2.4. Discussion

Sensory reflex epilepsies are an intriguing class of neurological disorders wherein seizures are provoked by a wide range of stimuli, including bathing in hot water (Satishchandra, 2003b; Striano et al., 2012; Italiano et al., 2016). Hot water epilepsy was first reported in a 10-year-old boy who manifested stiffened extremities and loss of consciousness during a hot water bath (Allen, 1945). Since then, several patients/families from across the world -- Australia (Keipert 1969), Turkey (Bebek et al., 2001), the United States (Stensman and Ursing, 1971), the United Kingdom (Moran, 1976), Japan (Kurata, 1979; Morimoto et al., 1985) and Canada (Szymonowicz and Meloff 1978) have been published. However, the majority of cases have been reported from India (Mani et al., 1974b; Satishchandra et al., 1988a; Satishchandra, 2003a). While molecular mechanisms underlying some of the reflex epilepsies; photosensitive and audiogenic seizures are well studied, for HWE, very little is known. In photosensitive epilepsy, stimulation of a critical neuronal mass in the occipital cortex is the primary site in the pathogenesis (Fisher et al., 2005). Audiogenic seizure development is proposed to be either a result of an anomaly in the sensory part of acoustic impulses in the brain pathway or abnormal biochemical and physiological status of central auditory and other structures (Garcia-Cairasco, 2002). For HWE, it is suggested that an interplay of complex tactile, temperature-dependent stimuli and aberration in the somatosensory network may be involved in eliciting seizure response (Stensman and Ursing, 1971).

Studies of HWE patients in India have suggested familial clustering in up to 15% of the probands (Mani et al., 1974b; Satishchandra et al., 1988a). Another report from Turkey suggested positive family histories in about 10% of the patients (Bebek et al., 2001). These reports indicate that HWE has a genetic predisposition to its etiology, which remained largely unknown until genetic studies in the past few years found three loci associated with the disorder mapping to 10q21.3-q22.3 (*HWE1*, MIM: 613339, (Ratnapriya et al., 2009b); 4q24-q28 (*HWE2*, MIM: 613340; (Ratnapriya et al., 2009a) and 9p24-p23 (Karan et al., 2018). These studies conducted on families with multiple affected members suggest that the disorder is inherited in an autosomal dominant fashion with incomplete penetrance values of 60-90%. In our study, the highest LOD score for the marker D8S277 at 8p23-p12 obtained was at a 60% penetrance value, suggesting weak penetrance of the underlying susceptible allele. Such an observation of incomplete

penetrance is not uncommon in epilepsy syndromes (Thomas and Berkovic, 2014). In the case of autosomal dominant partial epilepsy with an auditory feature, mutations in *LGI1* have a disease penetrance of 67% (Ottman et al., 2004). Another study of a four-generation Caucasian family affected with familial mesial temporal lobe epilepsy reported a maximum LOD score obtained at 80% penetrance (Hedera et al., 2007). Also, *EFHC1* mutations that cause JME have been reported to exhibit variable phenotypic penetrance ranging from 65% to 78% (Suzuki et al., 2004a). Interestingly, the locus found in this study also overlaps with the GEFS6+ locus known in febrile seizures; however, no gene was identified at this position (Baulac et al., 2008). In HWE 244, the centromere-proximal and -distal recombination boundaries comprise a larger critical interval of 29Mb as compared to the GEFS6+ locus defined by a shorter 13Mb interval. Upon analysis of genes in the critical interval at 8p23-p12, it was observed that it does not harbor ion channel genes, which indicated the possibility of other pathways being important for the phenotype. Following the analysis by whole genome sequencing, a low frequency variant, c.1292G>A (p.Arg431Gln) in the *R3HCC1* was identified. I also found two additional R3HCC1 variants, p.Ala433Thr, and p.Val142Met among HWE patients. In our previous study, we reported *SLC1A1* mutations for HWE which affected glutamate uptake, glutamate kinetics and anion conductance properties of SLC1A1 (Karan et al., 2017). SLC1A1 is a neuronal glutamate transporter that plays an important role in limiting excitotoxicity (Kanai et al., 1995) and its loss-of-function in mice is shown to cause age-dependent neurodegeneration along with some behavioral changes and cognitive impairments (Aoyama et al., 2006). Interestingly, functionally relevant *SLC1A1* mutations have been associated with the pathophysiology of psychotic disorders, schizophrenia, and bipolar disorder as well (Myles-Worsley et al., 2013; Afshari et al., 2015; Li et al., 2020). I further investigated *R3HCC1* mutations in publicly available databases from Broad Institute for Bipolar disorder and Schizophrenia. The BipEx (Bipolar dataset) consists of 14,210 cases and 14,422 controls. For *R3HCC1* it reports 306 variants, of which 17 variants (A.1.4) exclusive to patients show pathogenic effect upon analysis with two insilico prediction software; Polyphen-2 and SIFT (https://bipex.broadinstitute.org/gene/ENSG00000104679). A similar analysis was done for schizophrenia patients using SCHEMA: Schizophrenia exome meta-analysis consortium that comprised exome data of 24,248 cases and 97,322 controls. The data shows that out of 95 variants in *R3HCC1* for both patients and controls, seven

pathogenic variants reported (A.1.4) are exclusive to schizophrenic patients (https://schema.broadinstitute.org/gene/ENSG00000104679). Chromosome 8p has also been reported as a potential hub for genetic susceptibility to neuropsychiatric disorders (Tabarés-Seisdedos and Rubenstein, 2009), and since a continuum of phenotype exists between HWE and neuropsychiatric disorders, there is a possibility of involvement of *R3HCC1* in such disorders as well.

R3HCC1, as described earlier, codes for a protein of 440 amino acids. The protein has orthologs in *Bos taurus, Mus musculus, Canis familiaris, Danio rerio, Gallus gallus, Rattus norvegicus* and *Drosophila melanogaster*. While the biological function of the protein remains uncharacterized, its domain architecture suggests that its cellular roles may involve RNA-binding. It is reported that *R3HCC1* makes use of a non-AUG (CUG) start codon with strong conservation and abundance of ribosome occupancy near to the coding sequence (Lee et al., 2012a). A high throughput proteomic study by Stes and colleagues reported polyubiquitin-C as an interacting partner of this protein (Stes et al., 2014). Another study by Kashima et al., 2010, bioinformatically predicts the presence of exon- junction complex binding motif in R3HCC1, which would facilitate its association with exon- junction complex proteins- Y14, MAGOH, and EIF4A3 (Kashima et al., 2010). The study further shows that a paralogue of R3HCC1, R3HCC1L interacts with exon-junction complex proteins Y14 and MAGOH, which are important for the stability and translation of mRNA. In the subsequent chapter, I discuss potential cellular and molecular functions of R3HCC1 in RNA biology using localization and co-immunoprecipitation studies. Further, I examine the effects of patient mutations on this function and discuss their impact in the context of the neurological disorder. I also discuss the potential role of R3HCC1 at organism level in *Drosophila melanogaster*.

Chapter 3

Elucidating cellular functions of R3HCC1

As described in the previous chapter, our studies comprising a genome-wide sequencing and - linkage analysis of HWE244 and additional HWE patients suggested a role for R3HCC1 (R3H Domain and Coiled-Coil Containing 1) in predisposition to HWE. R3HCC1 protein has two predicted RNA-binding domains: the R3H domain and RNA recognition motif (RRM). R3HCC1 is synthesized in cultured neuroblastoma cells, SH-SY5Y, and localizes to the cytoplasm and nucleus. In primary rat hippocampal neurons, under basal conditions, R3HCC1 localized to neuronal cell bodies and proximal dendrites. In a cell, R3HCC1 localizes to stress granules, and overexpression of the rare variants identified leads to increased stress granule size in cultured mammalian cells. I found that R3HCC1 interacts with the RNA binding proteins - G3BP1 (RasGTPase-activating protein-binding protein 1) and UPF1 (Up-Frameshift Suppressor 1 Homolog)- known regulators of translational arrest and RNA decay pathways. The findings presented in this chapter indicate the involvement of R3HCC1 in the mRNA surveillance pathways. In addition, I describe a preliminary study in Drosophila melanogaster, which may in future help understand the role of R3HCC1. The fly protein CG2162 has about 36% similarity with human R3HCC1 and a high domain similarity of 77% and 64% for R3H and RRM, respectively. I noticed that the mutant CG2162 flies are sensitive to heat stress at the adult- and larval- stages. The adult flies exhibited a temperature-induced, seizure-like phenotype. Further studies involving examining R3HCC1's role in RNA biology as well as characterization of the protein in Drosophila shall provide insights into cellular-molecular mechanisms involved.

3.1. Introduction

RNA-binding proteins are key regulators of the post-transcriptional fate of a messenger-RNA (mRNA). Their association with mRNA leads to the formation of messenger ribonucleoprotein particle complex (mRNP), facilitating processes such as mRNA transport, translation, storage, and degradation. Various studies on translationally inactive mRNP have shown that they can assemble into high order complexes, mainly

stress granules, P bodies, and neuronal transport granules (reviewed in, Kedersha and Anderson, 2002; Kiebler and Bassell, 2006; Jain and Parker, 2013). Stress granules (SGs) are formed upon initiation of translational inhibition under various cellular stress conditions (Wek et al., 2006). These stress conditions can arise due to heat, oxidative, radiation, or hypoxia. During such conditions, cells halt the translation of the majority of the proteins and switch to the production of specific cytoprotective proteins, such as molecular chaperones and DNA-repair enzymes. Cells accomplish this goal by forming SGs, which are not only involved in silencing non-essential transcripts but also facilitate sequestering of signaling proteins that are important in cell viability (Anderson and Kedersha, 2002; Arimoto et al., 2008). SGs are dynamic in nature, which means there is a constant shuttling of proteins in and out of SGs (Kedersha et al., 2005). These also interact with other mRNP granules; P bodies, and neuronal granules. SGs assembly is mainly regulated by one or more RNA-binding proteins, which include TIA-1 (Tia cytotoxic granule associated protein) (Gilks et al., 2004), G3BP1 (Ras GTPase-activating protein binding protein 1) (Tourrière et al., 2003), Fragile X Mental Retardation protein (FMRP) (Mazroui et al., 2002), and survival of motor neuron (SMN) protein (Hua and Zhou, 2004). Once the cell has overcome its stress, these granules are disassembled either via ATPase-driven remodeling events or autophagy (Mazroui et al., 2007; Buchan et al., 2013; Jain et al., 2016). P bodies are sites for translational repression, comprising proteins that play a crucial role in mRNA degradation, nonsense-mediated mRNA decay (NMD), and RNA-mediated gene silencing. P bodies are present in cells irrespective of cellular stress, but their assembly can be enhanced in such conditions. The key components of P bodies include mRNA deadenylation and 5′-to-3′ decay complex including the deadenylation complex Ccr4-Not, Lsm1-7, the decapping coactivator and enzyme Dcp1/Dcp2, various decapping activators such as Edc3, DDX6 and EDC4, and the 5′-to-3′ exoribonuclease Xrn1 (Ingelfinger et al., 2002; Fenger-Grøn et al., 2005; Parker and Sheth, 2007). Neuronal transport granules are another class of mRNP granules that have been studied in recent years. They play a critical role in mRNA localization and local translation, which are crucial processes in the development of the central nervous system as they support axonal guidance and the formation of a stable synapse (Bassell et al., 1998; Wu et al., 2005a; Hengst and Jaffrey, 2007). In the mature central nervous system, neuronal transport granules in dendrites are sites for activity-

dependent local protein synthesis, which in turn regulates synaptic plasticity (Doyle and Kiebler, 2011; Kandel et al., 2014).

Interestingly, mutations in RNA-binding proteins which regulate mRNP granules dynamics have been implicated in several neurodegenerative disorders. In disorders like Amyotrophic Lateral Sclerosis and frontotemporal lobar degeneration, mutations in proteins FUS (Aulas and Velde, 2015; Shang and Huang, 2016), TDP43 (Liu-Yesucevitz et al., 2010; Khalfallah et al., 2018), TIA1 (Hirsch-Reinshagen et al., 2017), are known to cause the formation of stress granules-like aggregates which are pathological and cause neuronal death. Fragile X-Syndrome is another example, wherein mutation in FMRP protein leads to a change in dendritic spine morphology and aberrant relation of protein synthesis at the synapse (Dictenberg et al., 2008).

As discussed in the previous chapter, R3HCC1's biological function remains largely unexplored; however, its domain architecture suggests that it is an RNA-binding protein. Here, I present the molecular and cellular characterization of R3HCC1, suggesting its potential role in mRNA regulation. A preliminary study to characterize the role of R3HCC1 in *Drosophila melanogaster* is also presented.

3.2. Material and methods

3.2.1. Plasmids constructs

The open reading frame of R3HCC1 was PCR-amplified from cDNA of HEK293 cells using gene-specific primers incorporating restriction sites: NheI-forward and XhoI-reverse (A.2.5) and cloned into pCDNA-3.1+. R3HCC1 was further sub-cloned into two vectors: p3X-FLAG and pEGFP-C2 using specific primers. Mutations and domain truncations for R3HCC1 were introduced in wildtype constructs using site-directed mutagenesis primers (A.2.5) and QuikChange Site-Directed Mutagenesis reagents (Agilent Technologies, California, USA). The open reading frame of PABPC1, Y14, MAGOH, EIF4A3 was also PCR-amplified from HEK293 cells cDNA using gene-specific primers (A.2.4) and cloned p3X-FLAG vector. The open reading frame of Dcp1A was cloned in pEGFP-N1 (A.2.4). All the constructs used were sequence confirmed.

3.2.2. Cell culture and transfections

The cell lines HEK293, U87-MG, SH-SY5Y, HeLa were maintained in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% heat-inactivated fetal bovine serum (FBS), 2 mM L-glutamine, 100 U/ml penicillin, and 0.1 mg/ml streptomycin (Sigma-Aldrich), in a humidified atmosphere of 5% CO_2 and 37ºC. For transfection, cells were seeded either onto coverslips coated with poly-L lysine in 35mm dishes or on 60/90mm dishes (Eppendorf, Hamburg, Germany). Upon attaining 60% confluence, cells were subjected to antibiotic and serum-free DMEM, and transfections were done using Lipofectemine^{TM}2000 (Invitrogen-Life Technologies, Massachusetts, USA). After 5-6 hours of transfection, the medium was replaced by a complete medium (DMEM with 10% FBS), and 24-hours post-transfection, cells were processed for immunofluorescence and immunoblot assays.

3.2.3. Primary hippocampal cultures

Primary hippocampal neuron cultures were prepared from Sprague Dawley rats at embryonic day 18 (Kaech and Banker, 2006). The cells were plated at the density of 20,000-30,000 cells/cm^2 on poly-L-lysine (Sigma Aldrich) in Minimum Essential Media (Thermo Fisher Scientific) supplemented 10% FBS. The media was changed to a neurobasal medium (Thermo Fisher Scientific) containing B27 (Thermo Fisher Scientific) and Glutamax (Invitrogen-Life technologies) after 3 hours. Neurons were cultured at 37°C in a humidified atmosphere with 5% CO_2. The cells were processed for treatments and staining on Day 14 (DIV14). The animal work was carried out in accordance with procedures approved by the Institutional Animal Ethics Committee and the Institutional Biosafety Committee, inStem, Bangalore, India. The animals were maintained at 20–22°C temperature, 50–60 relative humidity, 0.3 µm HEPA-filtered air supplied at 15–20 ACPH, and 14-h/10-h light/dark cycle and were freely fed with food and water.

3.2.4. Stress experiments

For induction of stress, the cells were subjected to two conditions: (a) heat shock at 44°C for one hour and (b) oxidative stress by treatment with 600µM Sodium arsenite ($NaAs0_2$, Sigma-Aldrich) for 45 minutes, in a humidified atmosphere containing 5% CO_2 and 95%

air (Kedersha and Anderson, 2007). G3BP1 and PABPC1 were used as a marker for stress granules. Dcp1a was used as a marker for the P body.

3.2.5. Nuclear and cytoplasmic fractionation

SH-SY5Y cells were seeded onto a 90mm dish, and upon 90% confluence, cells were pelleted and washed thrice in ice-cold 1X phosphate buffer saline (PBS). Cells were lysed on ice in cytoplasmic lysis buffer (10mM Tris pH 7.9, 1.5mM $MgCl_2$, 10mM KCl, 0.2 % NP40, and protease inhibitor cocktail) for 10 minutes. The nucleus was sedimented by centrifugation at 2500g for 5 minutes, and the supernatant containing cytoplasm was removed. The nuclear pellet was rewashed in cytoplasmic lysis buffer to remove any contaminating cytoplasm and resuspended in nuclear lysis buffer (20mM Tris pH 7.9, 1.5mM $MgCl_2$, 400mM NaCl, 10mM EDTA, 25% glycerol, and protease inhibitor cocktail) for 30 minutes on ice, followed by centrifugation at 20,000g for 30 minutes. The supernatant was collected as a nuclear fraction. For immunoblot analysis, histone H2A and α-tubulin were used as nuclear and cytoplasm controls, respectively.

3.2.6. Microtubule fractionation

SH-SY5Y were treated with 10nM Taxol and DMSO mock control for 1 hour in a humidified atmosphere of 5% CO_2 and 37ºC. The cells were washed in 1X PBS and lysed in gradient lysis buffer (50mM Tris pH 7.5, 150mM KCl, 2mM $MgCl_2$, 2mM EGTA, 2% glycerol, 0.125% Triton X-100, 10nM Taxol, 1mM DTT, 5% NP40) supplemented with protease inhibitor and RNase Inhibitor (100U/ml) or RNase (0.2mg/ml) for 5 minutes. 50µl of lysate was taken as input and centrifuged at 16,000g at 4°C for 30 minutes. The remaining lysate was centrifuged at 700g for 5 minutes at room temperature. The supernatant obtained was further centrifuged at 16,000g for 20 minutes at room temperature. The microtubule pellet obtained was resuspended in 40µl 2X loading dye, and the supernatant was stored as a cytoplasmic fraction. For microtubule disassembly, cells were pre-treated with 10µM of nocodazole for 30 minutes and processed similarly.

3.2.7. Immunocytochemistry

For immunofluorescence assay, cells were washed thrice in 1X PBS and fixed with 2% paraformaldehyde (PFA) for 15 minutes at room temperature. The cells were permeabilized using 0.1% Triton X-100 for 10 minutes, followed by incubation in 3%

bovine serum albumin (BSA) (made in 1X PBS) blocking solution for 1 hour at room temperature. The cells were treated with the required primary antibody for 1 hour at room temperature followed by incubation with 1:500 secondary antibody conjugated with Alexa 488/Alexa 568/Alexa 633 probe (Invitrogen-Life Technologies) for 1 hour at room temperature. The cells were washed in 1X PBS and stained for nucleus using DAPI (Sigma-Aldrich) for 15 minutes. The coverslips were mounted on a glass slide using PVA-DABCO mounting media (Sigma-Aldrich). The antibodies were mixed in a 1:1 ratio for dual antibody staining procedures while treating the cells for primary and secondary antibodies. For co-staining with UPF1, cells were incubated with anti-hUPF1 Alexa Fluor® 488 overnight at 4ºC after adding primary and secondary antibodies against R3HCC1. The imaging was done with LSM 880 Meta confocal laser scanning microscope (Carl Zeiss, Germany) under a 63X/1.4 oil immersion objective.

For staining rat neuronal cultures, cells were fixed with 2% PFA for 20 minutes at room temperature. The cells were permeabilized using 0.1% Triton X-100 for 10 minutes, followed by incubation in blocking solution (2% FBS, 2% BSA, 0.1% TritonX-100 in 1X PBS) for 1 hour at room temperature. The cells were treated with primary and secondary antibodies (made in blocking solution) for 1 hour at room temperature. The cells were washed, stained for the nucleus using DAPI for 15 minutes, and the coverslips were mounted on a glass slide using PVA-DABCO mounting media.

Antibodies used were anti-R3HCC1 rabbit raised (1:500, ab121740), anti-FLAG mouse raised (1:1000, F1804), anti-G3BP1 mouse raised (1:1000, ab56574), anti- hUPF1 Alexa Fluor® 488(1:100, ab201761), anti-MAP2 (1:1000, M9942). All the antibody solutions were made in 1% BSA (in 1X PBS) except mentioned otherwise.

For stress granules quantification, G3BP1 was used as a marker, and cells with at least two foci were marked as positive for stress granule. 80-100 cells overexpressing R3HCC1 wildtype, R3HCC1 patient variants, and R3HCC1 truncations were examined. For quantifying the size and number of SG, ten cells from each experiment were randomly selected, and SG within range 0.2-15μm^2 were analyzed using ImageJ function: Analyze particle (Khalfallah et al., 2018). All the analysis was done for four independent experiments. For statistical analysis, one-way ANOVA followed by post hoc Dunnett's test was done.

3.2.8. SDS-PAGE and immunoblotting

Cultured mammalian cells were pelleted down and washed thrice in 1X PBS. The cells were lysed in RIPA buffer (150 mM NaCl, 10 mM Tris-pH 7.5, 0.1 % SDS, 1 % Triton X-100, 1% deoxycholate and 5 mM EDTA supplemented with protease inhibitor cocktail) by re-suspension on ice for 30 minutes. The lysate was passed through a 26G syringe four to five times at 10 minutes intervals and centrifuged at 15,000 rpm for 30 minutes at 4ºC. The supernatant was collected, and the total protein concentration in the cell lysate was estimated using the bicinchoninic acid assay. 50µg of protein boiled in 6X SDS loading dye for 5 minutes and was loaded and resolved in 8-12% SDS-polyacrylamide gel depending on the desired molecular weight of the protein of interest (Sambrook and Russell, 2001). The proteins were transferred onto a nitrocellulose membrane (GE Healthcare, Chicago, USA) using semi-dry transfer apparatus (Amershan biosciences, Buckinghamshire, UK) at 20V for 1 hour. The membrane was washed with 1X PBS and kept in the blocking solution (5% skim milk in 1X PBS) for 4 hours at 4ºC with constant shaking followed by overnight incubation in the required primary antibody solution at 4ºC. The membrane was washed with 0.1% PBST (0.1% Tween in 1X PBS) twice for 10 minutes each and incubated in a secondary antibody solution conjugated with HRP substrate (1:5000 dilution) for 4 hours at 4ºC. The membrane was washed with 0.1% PBST twice for 10 minutes each, and the protein bands were detected using enhanced chemiluminescence substrate (Thermo Fischer Scientific). α-tubulin was used as a loading control in the same blot wherever required. The same protocol was followed for all immunoblotting experiments unless specified otherwise.

Antibodies used were anti-R3HCC1 rabbit raised (1:1000, ab121740) and anti-α-tubulin mouse raised (1:5000, T6074), anti-Histone H2A rabbit raised (1:10000, ab8580), anti-FLAG mouse raised (1:2500, F1804), anti-G3BP1 mouse raised (1:2500, ab56574), anti-hUPF1 (1:1000, ab109363), and anti-GFP rabbit raised (1:3000, AB10145). All the antibody solutions were made in 1% BSA (in 1X PBS).

3.2.9. Co-immunoprecipitation assay

For co-immunoprecipitation assays, Dynabeads Protein G Immunoprecipitation kit supplemented with buffers (Invitrogen-Life Technologies) was used. 1µg of target antibody in antibody-binding solution was added to 20µl of Protein-G Dynabeads, and the mix was rotated for 8 hours at 4°C. The beads were washed in antibody binding-

washing buffer to remove any unbound antibody. Cultured cells were lysed for 30 minutes on ice in IP lysis buffer (25mm Tris pH 7.4, 150mM KCl, 5mM EDTA, 0.5% NP-40) supplemented with protease inhibitor (Roche, Basel, Switzerland) and spun at 20,000g for 25 minutes at 4°C. The supernatant was added to the antibody complexed beads and incubated overnight with constant rotation at 4°C. The beads were washed four times in the washing buffer for one minute each. The bound proteins were then eluted with 2X Laemmli buffer and analyzed by immunoblotting.

3.2.10. Fly lines

All genotypes were reared on standard cornmeal medium under Light: Dark (12:12-hr cycle) at 25ºC. Dmel\P{GawB}CG2162^{NP3333} (NP104459) enhancer trap line was obtained from the Kyoto stock center. w^{1118} control fly line was obtained from Bloomington (BL5905), and it served as a control for all the behavioral assays. UAS-GFP reporter line (BL6874) was used to study the expression of *CG2162* by immunohistochemistry.

3.2.11. Bioinformatic analysis

The transcript structure and isoforms of *CG2162* were obtained from NCBI; Gene ID: 38360. Multiple sequence alignment (https://www.ebi.ac.uk/Tools/msa/clustalo/) was done to predict conservation between human R3HCC1 and *Drosophila* CG2162 protein as well as for conservation of domains and variations found in the patient samples for HWE.

3.2.12. Genotyping Dmel\P{GawB}CG2162^{NP3333}

To confirm the presence of the P element, genomic DNA was isolated (A.2.19). PCR amplification was done using two primer pairs, one spanning the start sequence of the P element and the other at the end of the P element insertion in the NP3333 line. Primer pair across the P element sequence was used to confirm its absence from control flies (A.2.7). The PCR amplified products were verified by gel electrophoresis and sequence confirmed. The primers used are mentioned in A.2.7.

3.2.13. Expression study for *CG2162*

RNA isolation was done from the head of control w^{1118} and NP3333 lines using the Trizol method (Bogart and Andrews, 2006), followed by cDNA synthesis using the

SuperScript III First-strand synthesis kit (Invitrogen-Life Technologies). To confirm the presence of the transcript, PCR reactions were performed on cDNA using different sets of primers spanning the length of the longest isoform for *CG2162* (A.2.8). All amplicons were sequence confirmed.

For tissue-specific expression analysis for *CG2162*, the NP3333 line was crossed with a UAS-GFP. Dissection of the brain was done from F1 generation third instar larvae and 2-3 day old adult flies. The brains were fixed in 4% PFA for 30 minutes and incubated in 10% horse serum blocking solution made in PBT (Phosphate buffer saline supplemented with 0.5% Triton X-100) at room temperature for 1 hour. The brains were treated with the anti-GFP antibody for 24 hours at room temperature, followed by five washes in 0.5% PBT for 10 minutes each. Next, brains were incubated with a secondary antibody (1:3,000) conjugated with Alexa 488 probe for 12 hours at room temperature followed by washes. The sample was cleaned and mounted on a glass slide. The imaging was done with LSM 880 Meta confocal laser scanning microscope (Carl Zeiss) under 63X/1.4 oil immersion objective at 20X and 40X. The primary antibody used: anti-GFP antibody chicken-raised (1:2000, Invitrogen-Life technologies). The antibodies were prepared in the blocking solution.

3.2.14. Heat-shock assay

Drosophila males and females were separated immediately after eclosion into batches of 15, placed in feeding vials containing standard cornmeal medium, and kept in standard rearing conditions. Three days post-eclosion, the flies were flipped into labeled empty vials and incubated for 15 minutes for acclimatization at 25ºC before starting the experiment. Each of these vials containing 15 flies were immersed in a water bath at 42 ± 0.5ºC for two minutes. The temperature of the water was monitored using a thermometer at the start and end of two minutes. The number of flies "paralyzed" were counted at 15-second intervals in a span of two minutes. Following this, recovery from "paralytic-shock" was recorded by counting flies that could position themselves upright or climb the vial wall at an interval of 3 minutes for 30 minutes. As mentioned earlier, the number of flies used per experiment is 15, and 13 such experiments were conducted for control w^{1118} and NP3333 genotypes. Males and females were analyzed separately.

Drosophila wandering phase third instar larvae of both control and NP3333 flies were kept on 2% agar plates at 37ºC for two hours. These larvae were shifted to fresh food

vials, and the number of flies that emerged was counted. For every experiment, 20 larvae were taken from w^{1118} and NP3333 genotypes, and 10 such experiments were conducted. The percentage of fly emergence was also recorded without giving any heat shock treatment. For statistical analysis, the percentage of fly emergence was plotted, and a two-tailed unpaired Student's t-test was used.

3.3. Results

3.3.1. Expression and localization of R3HCC1

The expression of *R3HCC1* transcripts in human brain regions was analyzed using available datasets for transcript and protein expression. The human protein atlas reports *R3HCC1* transcripts in the cerebral cortex, hippocampus, amygdala, midbrain, and other brain regions (Figure 3.1A). The transcript expression was confirmed by amplifying *R3HCC1* cDNA from Marathon-Ready TM full-length brain cDNAs from the human cerebral cortex, hippocampus, cerebellum, and hypothalamus (Figure 3.1B). The R3HCC1 protein synthesis in the brain and kidney, lung, retina, and other tissues was reported in ProteomicsDB (Figure 3.1C).

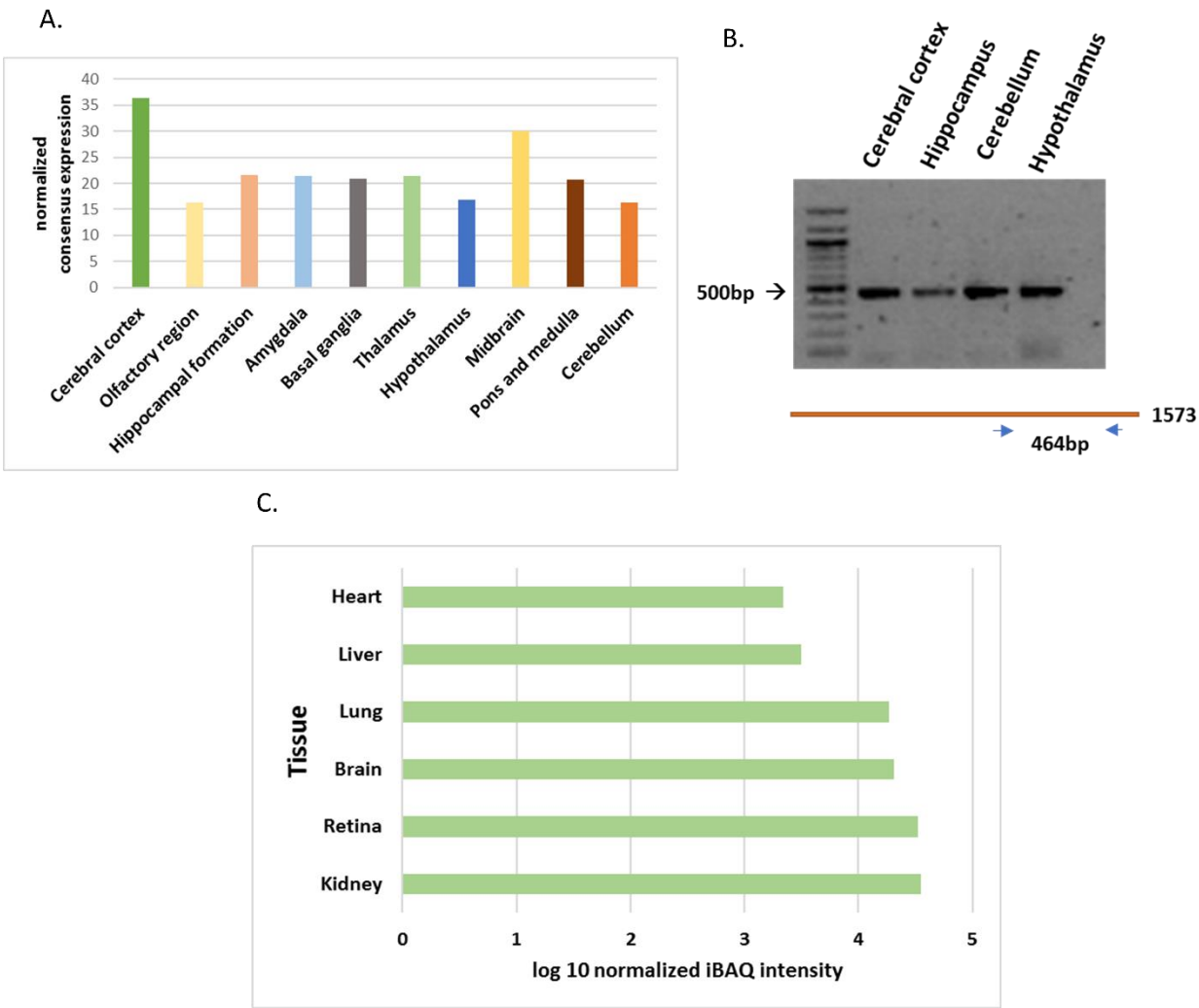


Figure 3.1. *R3HCC1* transcript expression: (A) Data plotted depicts normalized consensus expression levels for ten brain regions based on the data from two transcriptomics resources, GTEx and FANTOM5. (https://www.proteinatlas.org /ENSG00000104679-R3HCC1/brain). (B) R3HCC1-specific, PCR-amplified product from the brain cDNA for four regions corresponds to 464bp at the 3'end of the transcript, sequence confirmed. (C) Data plotted is median protein expression from various tissues, including the brain using proteomicsdb dataset (https://www.proteomicsdb.org/ proteomicsdb/#human/proteinDetails/Q9Y3T6/expression).

I examined R3HCC1 protein expression in various cell lines. A ubiquitous granulated expression pattern was detected in SH-SY5Y, HEK293, U87-MG, and HeLa cells (Figure 3.2 A, B). Cultured neuroblastoma SH-SY5Y cells were used in further experiments, wherein R3HCC1 expression was granulated and seen mainly in the cytoplasm. A somewhat lower level of punctate R3HCC1 staining was also seen in the nucleus, which was confirmed by immunoblot expression analysis in nuclear fractionation assay (Figure 3.2 C, D). R3HCC1 endogenous expression in primary rat

hippocampal neurons was examined, and a similar punctate expression pattern was observed in the soma and dendritic compartments (Figure 3.3A, B).

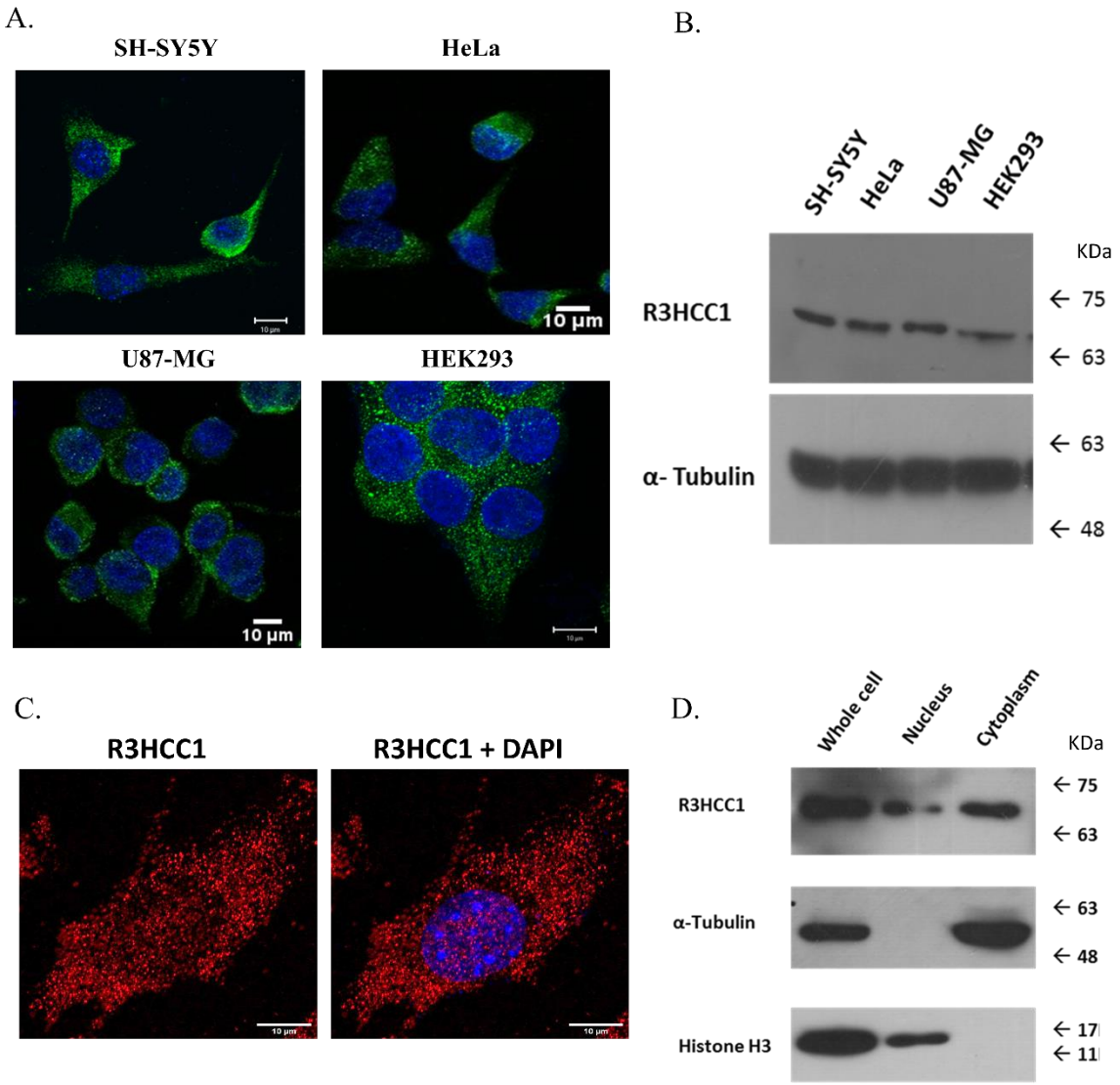
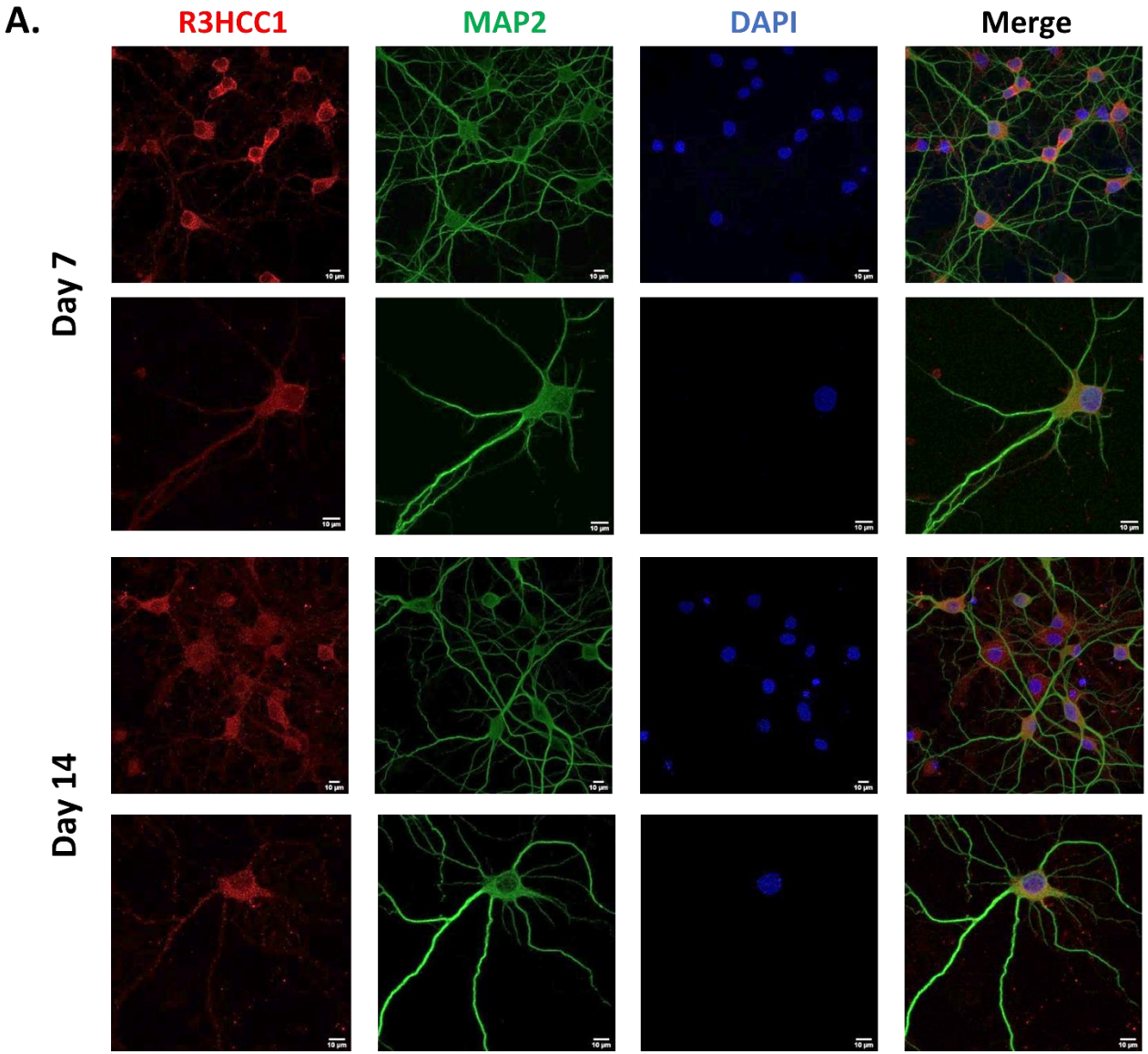


Figure 3.2. Expression of R3HCC1 in cultured mammalian cells: (A) Immunofluorescence experiments to check the endogenous expression of R3HCC1 in cell lines. Green indicates the R3HCC1 protein stained using anti-R3HCC1 antibody. Blue indicates DAPI staining; (B) Western blot indicates R3HCC1 staining at 63kDa in various cell lines; the predicted molecular weight is 50kDa; (C) Punctate expression of R3HCC1(red) in the nucleus and cytoplasm in SH-SY5Y cells. (D) Western analysis for subcellular fractionation of SH-SY5Y cells; stained with anti-R3HCC1, α-tubulin, and Histone H3 antibody for cytoplasmic and nuclear controls, respectively.


A.
R3HCC1
MAP2
DAPI
Merge
Day 7
Day 14

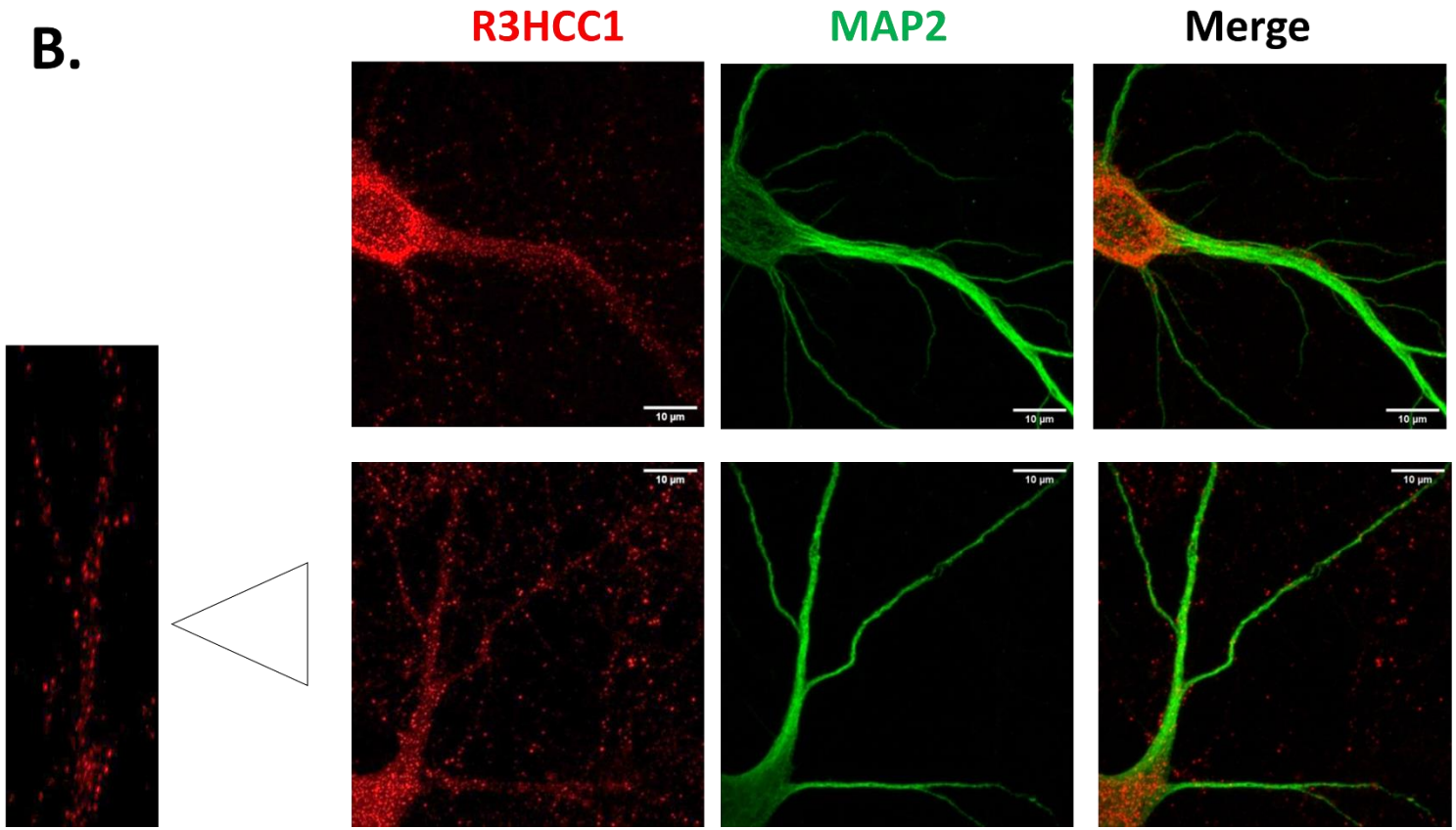


Figure 3.3. Expression of R3HCC1 in rat primary hippocampal neurons: (A) Immunofluorescence images depicting endogenous expression of R3HCC1 at Day 7 and 14 of the culture stained with anti-R3HCC1 antibody (red) co-stained with neuronal marker MAP2 (green); (B) Zoomed-in image of a dendritic shaft of the neuron depicting punctate expression pattern for R3HCC1.

The impact of variants found in the HWE patients on the R3HCC1 was checked by overexpression of eGFPC2-R3HCC1-wild-type and -variants in SH-SY5Y cells. It was observed that R3HCC1 wildtype and variants localize to the cytoplasm, and there is no apparent difference in their localization patterns. Upon immunoblot analysis for their abundance, no apparent change in the steady-state expression for variants was observed in comparison to wildtype (Figure 3.4).

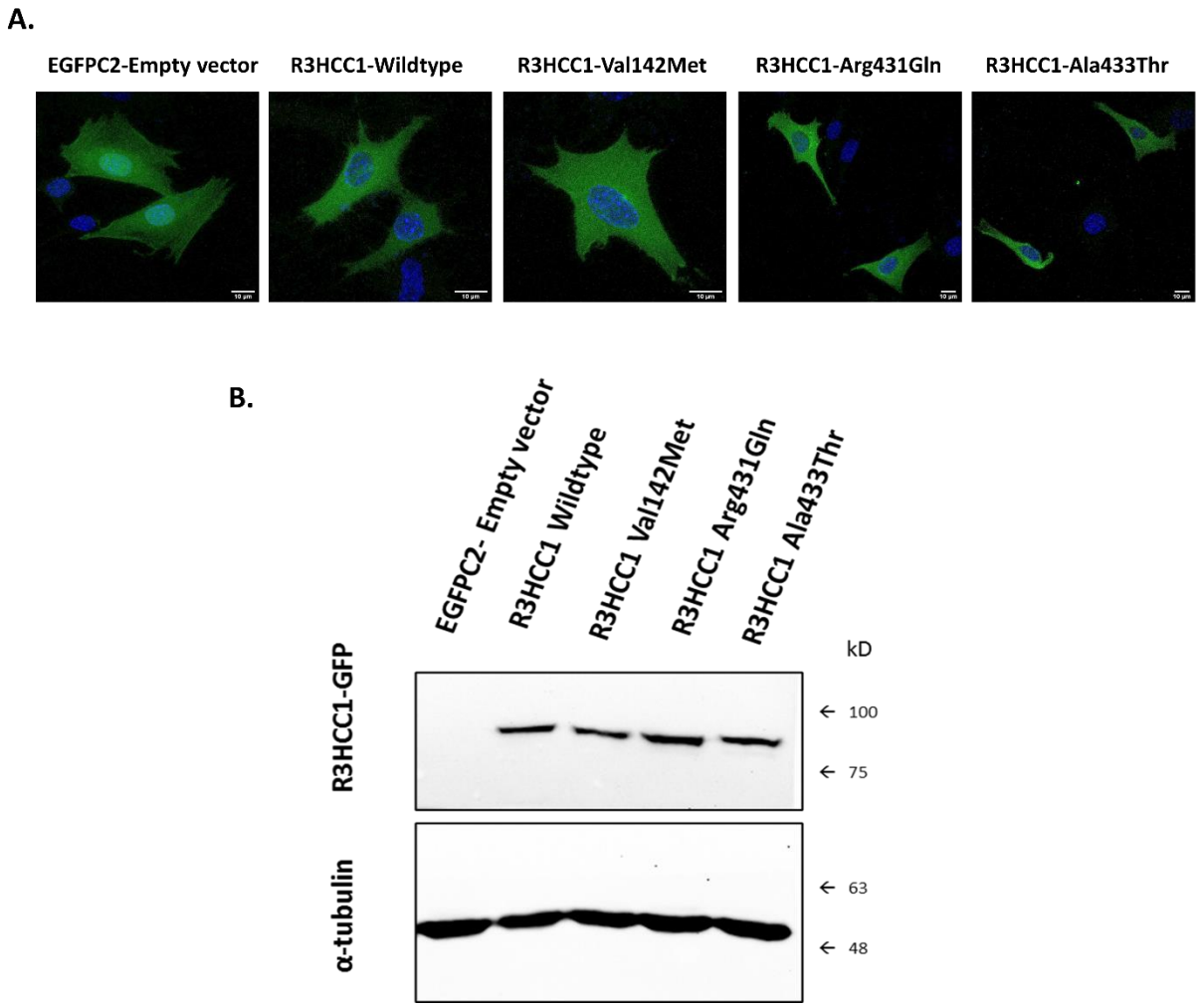
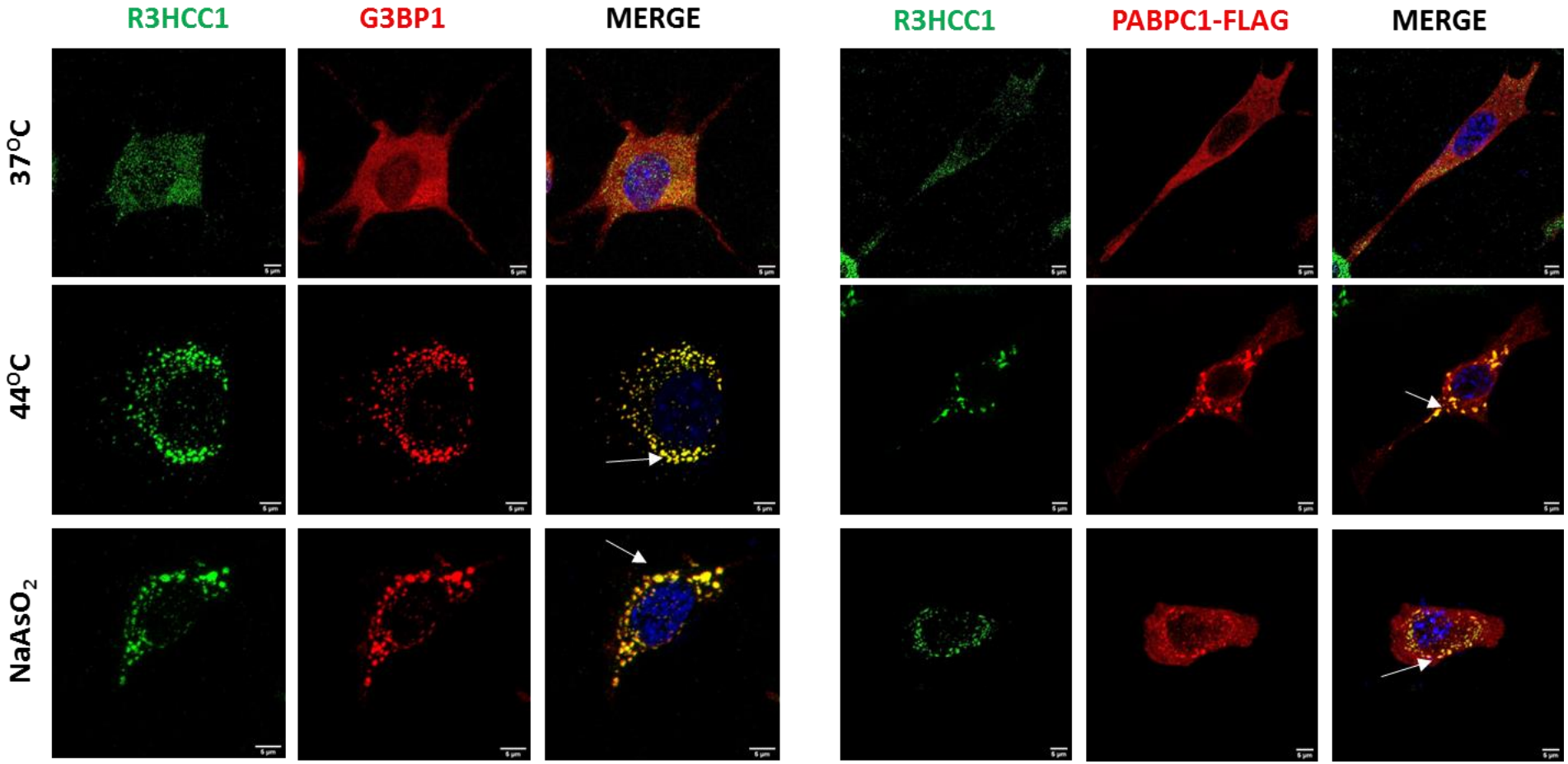


Figure 3.4. Overexpression of R3HCC1 wild-type and variants: (A) Immunofluorescence and (B) Western analysis of cultured SH-SY5Y cells transfected with R3HCC1 wildtype and R3HCC1 variants, p.Val142Met, p.Arg431Gln, and p.Ala433Thr.

3.3.2. R3HCC1 localizes to Stress granules

During the immunofluorescence experiments, it was occasionally observed that R3HCC1 localizes to large cytoplasmic granule structures resembling stress granules (SGs). SGs aggregate in the cytoplasm in response to stress stimuli. These membrane-less organelles consist of the non-translating mRNA, stalled translational preinitiation complex, RNA binding proteins (RBP) involved in mRNA processing, transport, storage and stability, and a subset of non-RBPs; post-translational modification enzymes, ubiquitin-modified proteins, protein or RNA remodeling complex, and signaling molecules (Mazroui et al., 2007; Mahboubi et al., 2013; Protter and Parker, 2016; Mahboubi and Stochaj, 2017). The proposed role of SG is to arrest translational machinery. SGs also facilitate the expression of certain specific genes that are essential for cell viability (Arimoto et al., 2008). SGs size is typically 200-400nm in diameter, but they can be up to 1000nm in size. Upon induction of heat and oxidative stress, R3HCC1 was observed to localize to SGs stained with RasGTPase-activating protein-binding protein 1 (G3BP1) (Figure 3.5A). The localization was also confirmed using Polyadenylate-binding protein 1(PABPC1), another marker for SG (Figure 3.5A). In order to recapitulate the SG finding in neurons, we treated primary rat hippocampal neurons with arsenite. We observed similar localization of R3HCC1 to SGs in these neurons as well.

A.



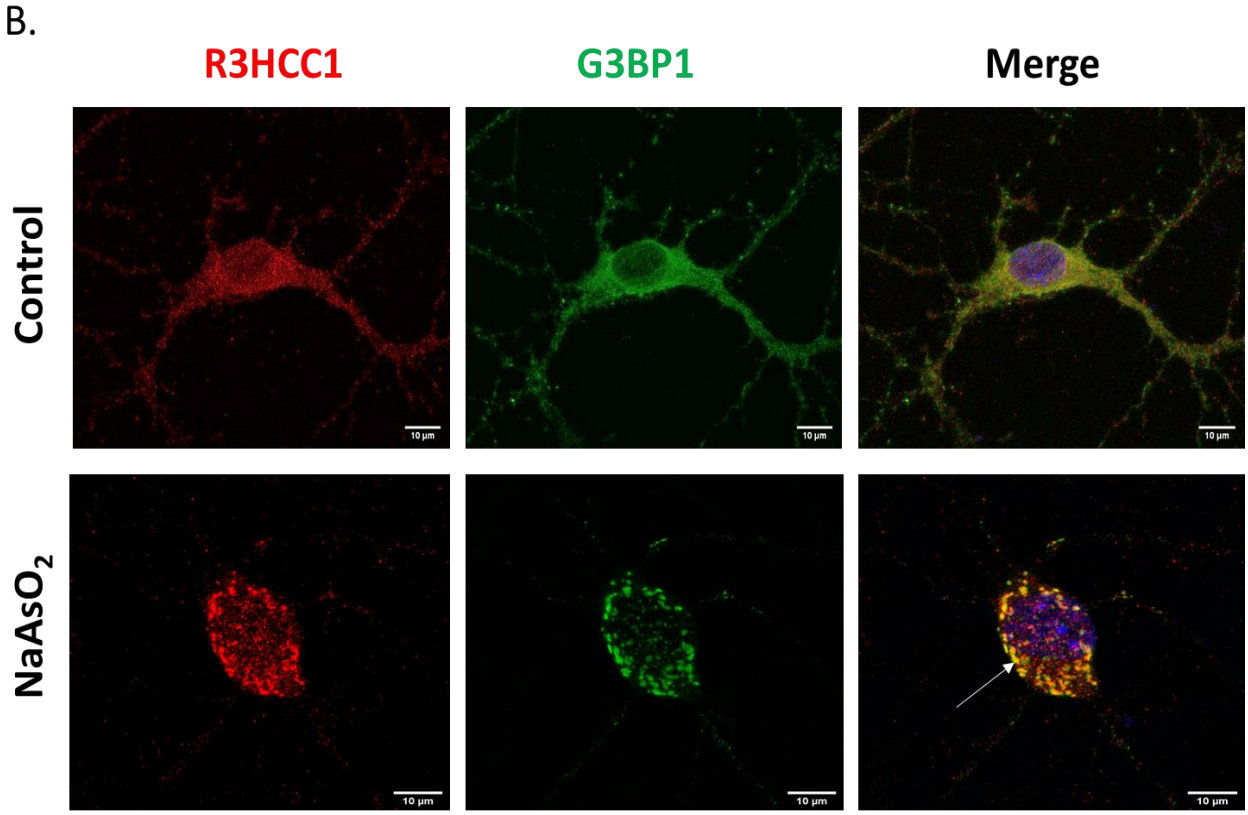


Figure 3.5. R3HCC1 localizes to SGs: (A) Cultured SH-SY5Y cells subjected to heat shock at 44°C for 1 hour and to 600µM $NaAsO_2$ treatment for 45 minutes and stained for endogenous R3HCC1 (green) along with endogenous G3BP1 (right panel) and FLAG-PABPC1 (left panel), (B) Rat primary hippocampal neurons subjected to 600µM of $NaAsO_2$ for 1 hour and stained for endogenous R3HCC1 (red) and G3BP1 (green).

To examine if the patient variants in R3HCC1 impact its localization to SG, formation of SG, and size and number of SG, eGFPC2-R3HCC1 wildtype and patient variants were overexpressed in SH-SY5Y, followed by the heat stress treatment. It was observed that all three variant R3HCC1 localized to SGs, and the SG formation was not disturbed. However, a significant increase in the size of SG after the stress treatment was observed in the case of p.Val142Met and p.Ala433Thr variants, but no significant change in the number of SG under the same conditions (Figure 3.6, A.1.5).

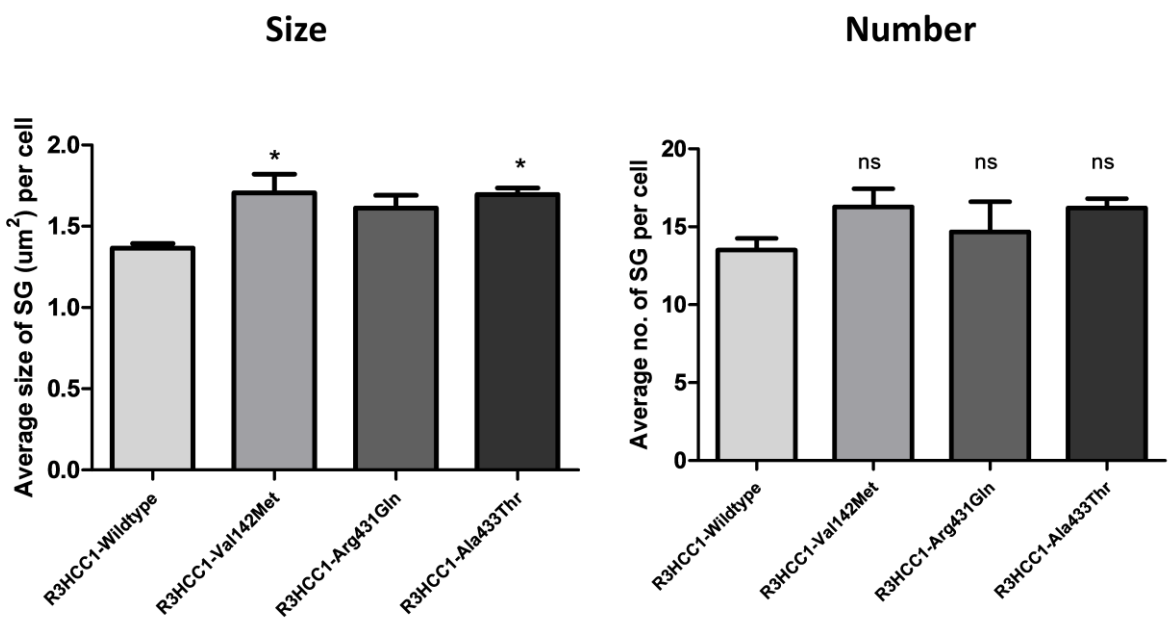
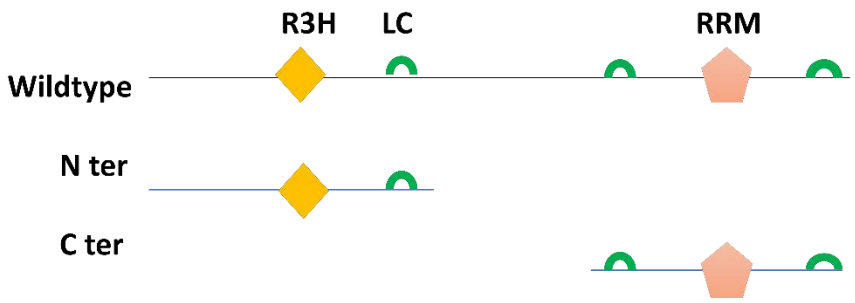


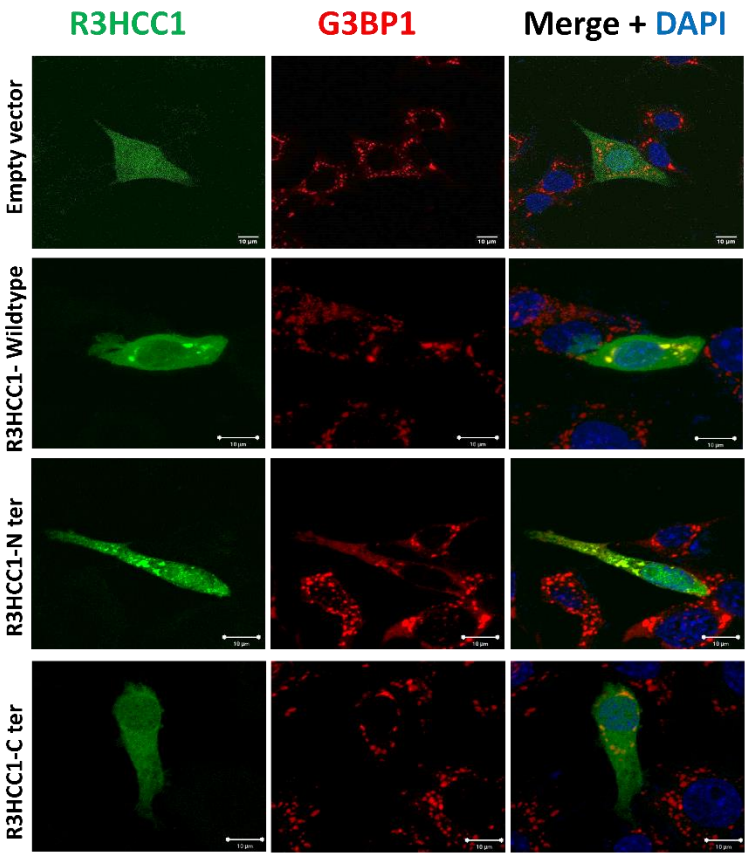
Figure 3.6. Overexpression of R3HCC1 mutants affects SG size: EGFPC2 constructs for R3HCC1-wild-type and variants, p.Val142Met, p.Arg431Gln, p.Ala433Thr were overexpressed in SH-SY5Y and subjected to 44°C for 1 hour following staining for endogenous G3BP1. The size and number of SG were recorded for ten cells for each experiment. Data are represented as mean ± SEM for four independent replicates (N=4). One-way ANOVA followed by Dunnet's multiple comparison test was used to compare wild type to mutant constructs *$p<0.05$.

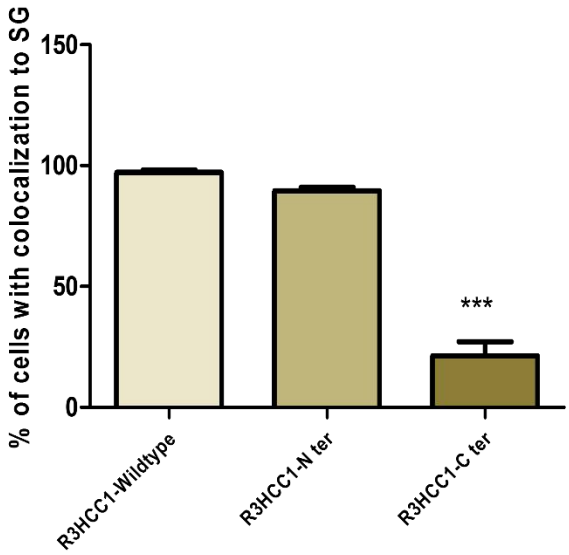
In order to examine which of the predicted RNA-binding domains are crucial in recruiting R3HCC1 to SG, R3HCC1 truncations were overexpressed in SH-SY5Y cells (Figure 3.7A). I observed that the C terminal region (has 238-440 a.a, lacks 1-237 a.a) comprising the RRM domain and two low complexity regions show a drastic reduction in recruitment to SGs upon heat shock treatment. And the N terminal region (has 1-147 a.a, lacks 148-440 a.a.) containing the R3H domain and one low complexity domain is observed to be sufficient and necessary to recruit R3HCC1 to SGs (Figure 3.7B). Moreover, I also observed that over-expression of N terminal (has 1-147 a.a) formed aggregates and led to a significant reduction in the number of cells containing SGs (Figure 3.7C).

A.

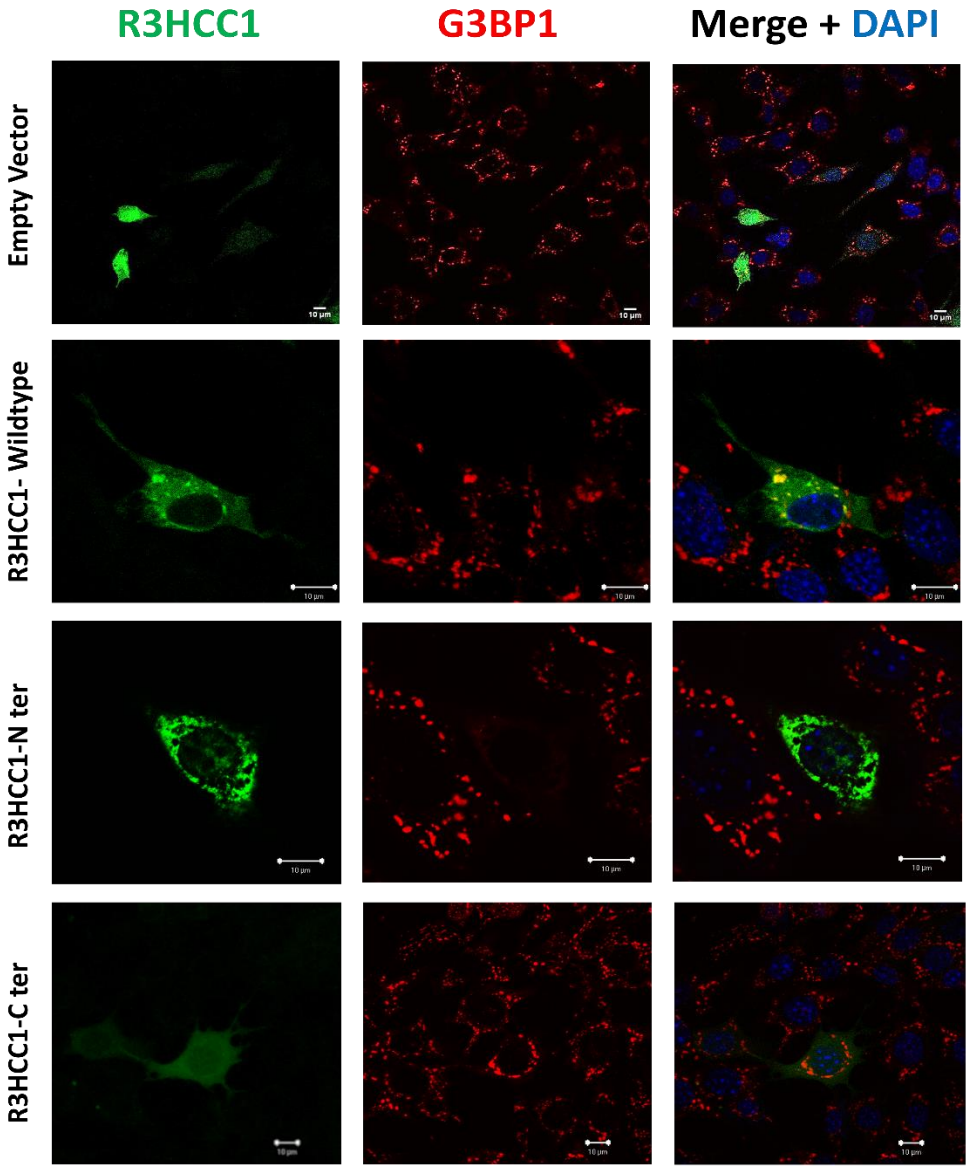

R3H
LC
RRM
Wildtype
N ter
C ter

B.


R3HCC1
G3BP1
Merge + DAPI
Empty vector
R3HCC1- Wildtype
R3HCC1-N ter
R3HCC1-C ter


% of cells with colocalization to SG
150
100
50
0
***
R3HCC1-Wildtype
R3HCC1-N ter
R3HCC1-C ter

C.



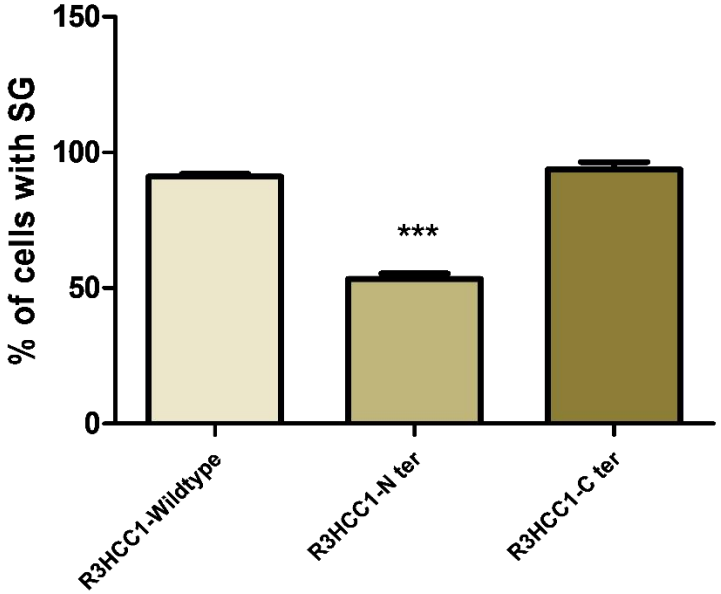
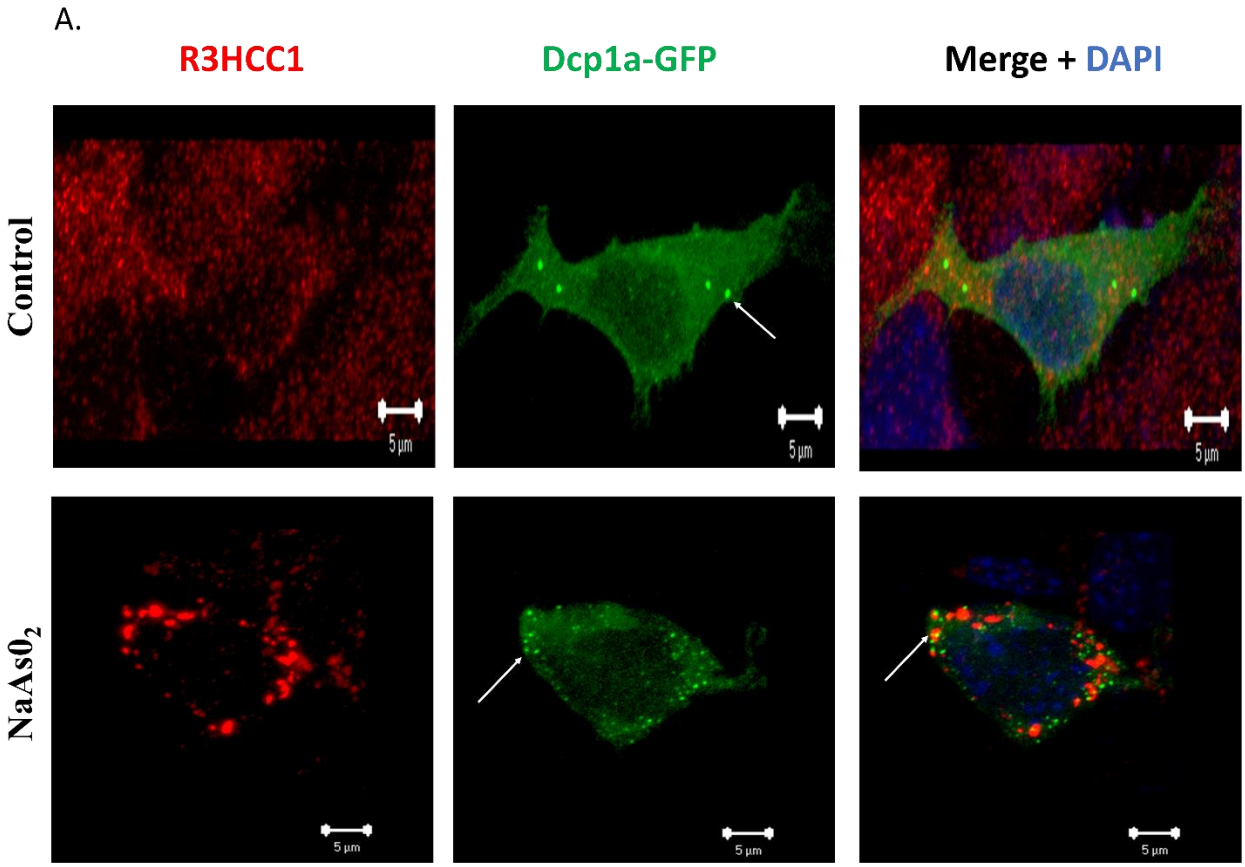


Figure 3.7. Overexpression of R3HCC1 truncations: (A) Schematic representation of R3HCC1-wildtype and -truncations: R3HCC1 1-147-a.a (N ter), 238-440 a.a (C ter). (B) Overexpression of truncated R3HCC1 C ter shows decreased localization to SG in SH-SY5Y. n = 80-90 cells for R3HCC1-WT and R3HCC1-C ter, per experiment. n=50-60 cells for R3HCC1-N ter, per experiment. (C) Overexpression of R3HCC1-N ter expressing cells show a reduction in the number of cells with SG. n = 80-90 cells per experiment for 4 independent experiments. One-way ANOVA followed by Dunnet's multiple comparison test was used to compare WT to mutant constructs ***$p<0.001$. Data are represented as mean ± SEM for four independent replicates.

3.3.3. R3HCC1 does not localize to the P body

While stress granules are the site of translational arrest, P bodies are of mRNA degradation. Both these classes of mRNP granules are known to interact with each other by a dynamic interchange of mRNA and protein complexes (Kedersha et al., 2005). Since both these granules are known to share some proteins (Kedersha et al., 2005), I examined if R3HCC1 also localizes to P bodies under stress conditions. The mRNA Decapping Enzyme 1a (Dcp1a), a marker for the P body, was overexpressed in SH-SY5Y cells, and the cells were subjected to heat- and oxidative- stress. The R3HCC1 localization seems restricted to SGs, and there is no recruitment of the protein to P bodies (Figure 3.8).



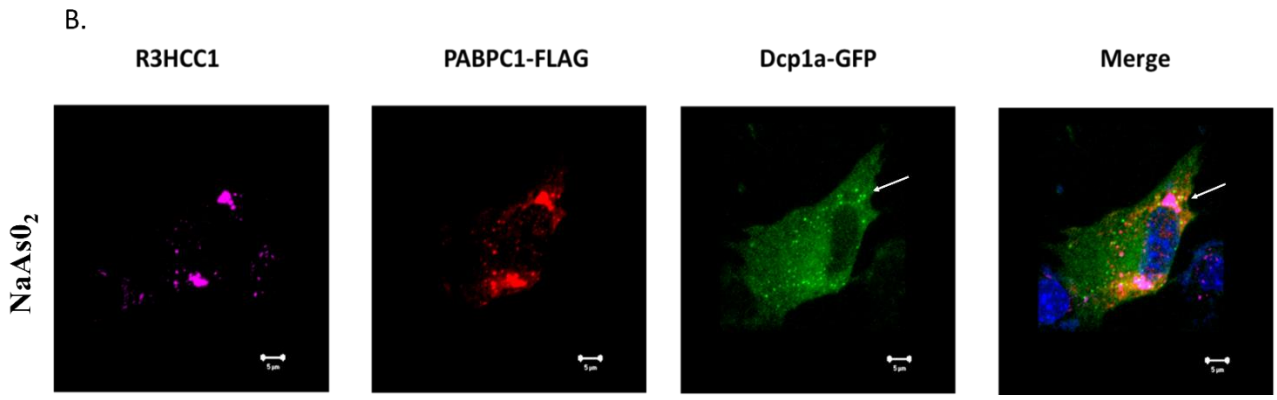


Figure 3.8. R3HCC1 does not colocalize to P bodies: (A) SH-SY5Y cells transfected with P body marker DCP1a-GFP and subjected to heat- and arsenite- stress followed by R3HCC1 (red) staining to examine colocalization to P bodies (marked by arrow), (B) SH-SY5Y cells stained for endogenous R3HCC1 (magenta), overexpressed PABPC1-FLAG (SG marker, red) and DCP1a-GFP (green), upon oxidative stress.

3.3.4. R3HCC1 interaction with the cytoskeleton is RNA-dependent

Earlier studies have shown that certain RNP granules are known to associate with the cytoskeleton. In neurons, this plays a significant role as transport of dendritic mRNA via neuronal granules along microtubules regulates the local translation (Martin and Zukin, 2006), synaptic strength (Kang and Schuman, 1996), and long-term memory formation (Huber et al., 2000; Doyle and Kiebler, 2011). Some of the RBP that are known to have an association with the cytoskeleton to assist mRNA localization are FMRP (Wang et al., 2008), hnRNPA2 (Kosturko et al., 2005), and Staufen (Köhrmann et al., 1999). I performed a subcellular fractionation assay to check if R3HCC1 shows a similar association with the cytoskeleton. It was observed that R3HCC1 is present in mock and paclitaxel stabilized cytoskeleton pellet of SH-SY5Y cell extract (Figure 3.9A). Upon adding RNase A to the cell lysate, a disruption of the interaction between R3HCC1 and cytoskeleton was seen (Figure 3.9B). Also, treating cells with a microtubule-destabilizing agent, nocodazole, led to the reduction of R3HCC1 in the cytoskeletal pellet (Figure 3.9C). These results indicate that the association of R3HCC1 protein with microtubule is sensitive to RNase treatment, suggesting that it might be a component of mRNA transport granule.

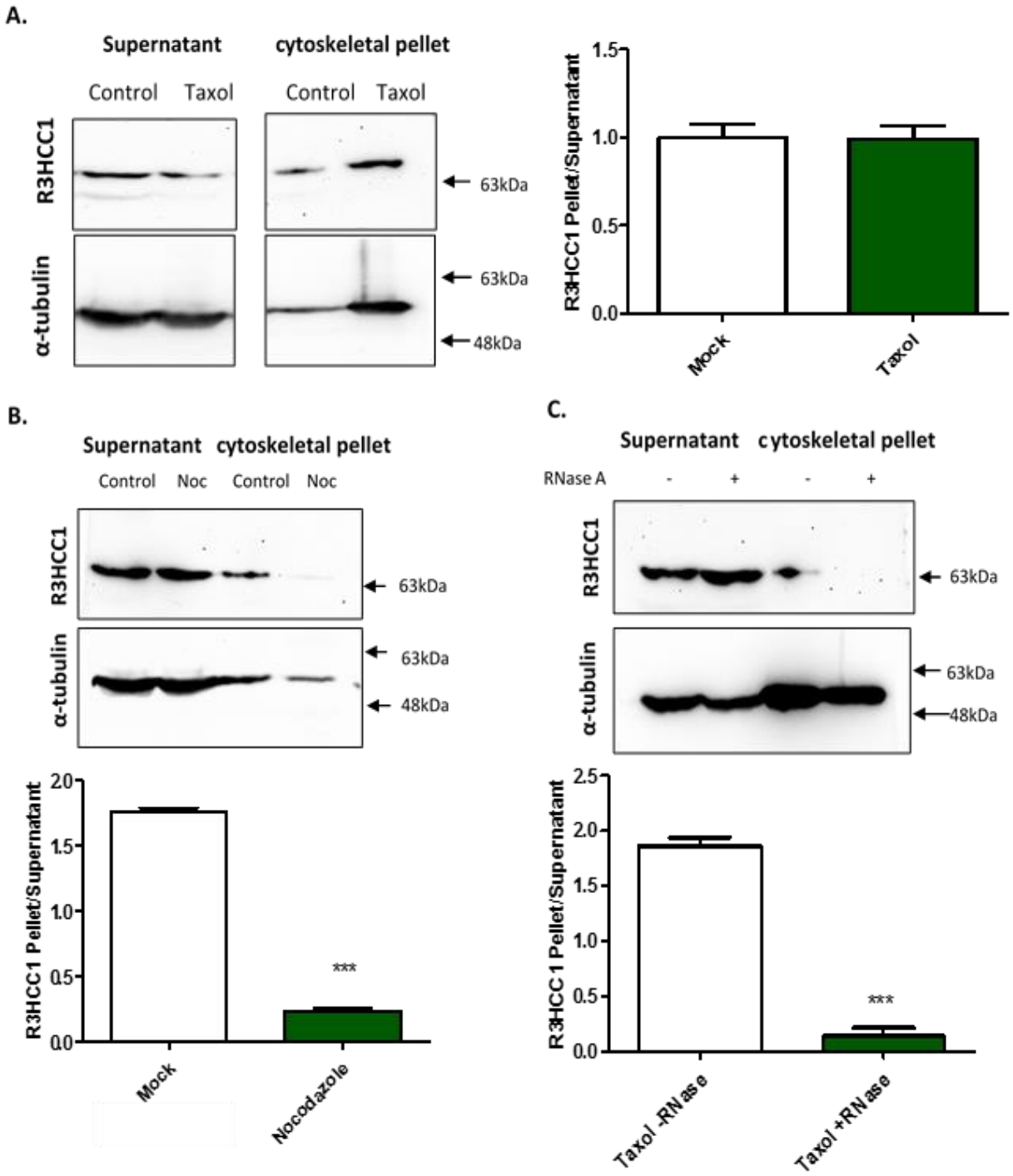


Figure 3.9. R3HCC1 interacts with microtubules: (A) R3HCC1 is associated with cytoskeletal in mock-treated control and paclitaxel-treated SH-SY5Y cells. (B) The association of R3HCC1 with cytoskeletal is abolished upon RNase A treatment. (C) Treatment of cells with nocodazole also reduced R3HCC1 association with microtubules. The ratio of densitometer reading of R3HCC1 in the cytoskeletal pellet and supernatant was calculated, normalized to the mock-treated control, and plotted. Data are represented as mean ± SEM for three independent experiments. Statistical significance was calculated by Student's *t*-test ***p<0.001

3.3.5. R3HCC1 interacts with SG proteins G3BP1 and UPF1

To get further insights into molecular pathways that R3HCC1 might be involved in; its potential protein interacting partners were looked at through co-immunoprecipitation experiments. Since R3HCC1 was observed to localize to SGs, its interaction with G3BP1, a core regulator for SG assembly, was first checked (Tourrière et al., 2003; Matsuki et al., 2013). It was found that the two proteins co-immunoprecipitated under basal conditions as well as when cultured cells were subjected to heat stress. During the stress conditions, the extent of immunoprecipitation of R3HCC1 with G3BP1 was reduced (Figure 3.10A), possibly suggesting changes in the interactome of G3BP1 upon stress and the existence of sub-structures within stress granules (Jain et al., 2016; Cirillo et al., 2020). In addition, both proteins were bound together upon RNase treatment before immunoprecipitation, suggesting an RNA-independent interaction (Figure 3.10B). The domain that is important for mediating this interaction was examined. Due to lower transfection efficiency in SH-SY5Y in our experiments, reverse-immunoprecipitation with eGFPC2-R3HCC1 truncation constructs (Figure 3.10C) was performed in cultured HEK293T cells. It was observed that immunoprecipitation with the N-terminal region (1-209aa), which was also responsible for recruitment to SGs, is essential for its interaction with G3BP1, and in the absence of this region, immunoprecipitation apparently does not occur.

Next, R3HCC1's interaction with Up-Frameshift Suppressor 1 Homolog (UPF1) was examined since UPF1 is also a SG protein and a predicted interactor of R3HCC1 in the STRING database. Our co-immunoprecipitation study with endogenous UPF1 confirmed that it interacts with R3HCC1 in SH-SY5Y (Figure 3.11A). Colocalization of endogenous UPF1 and R3HCC1 was observed in SH-SY5Y cells (Figure 3.11B). Upon overexpression of UPF1-GFP and p3X-FLAG R3HCC1 truncations (Figure 3.11C) followed by immunoprecipitation, it was observed that both the N terminal and C terminal regions were capable of mediating this interaction.

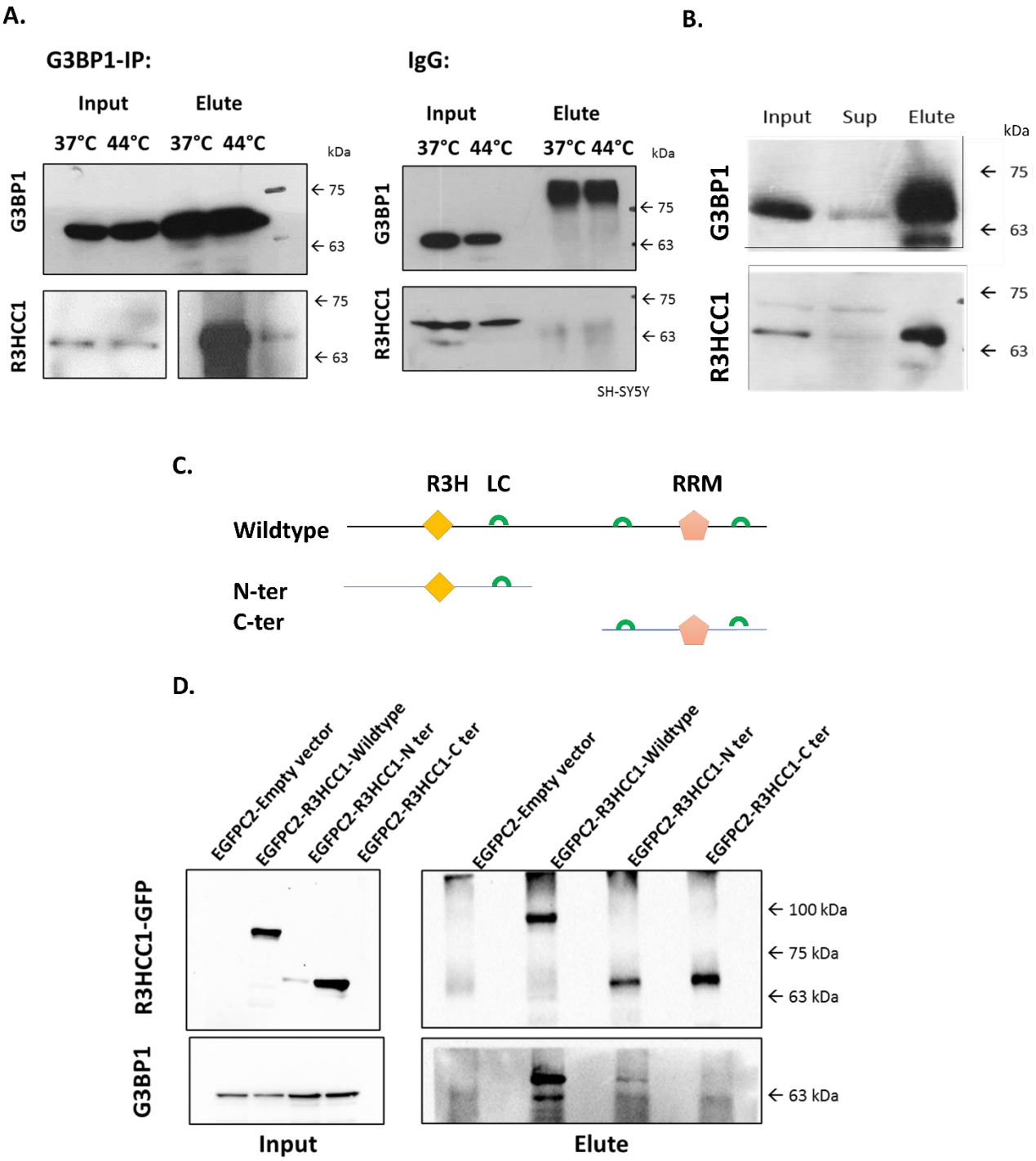


Figure 3.10. Interaction of R3HCC1 with G3BP1: (A) Western blot depicting co-immunoprecipitation of endogenous R3HCC1 with G3BP1 using anti-G3BP1 antibodies at 37°C and 44°C, (B) Co-immunoprecipitation in the presence of RNase A, (C) Representation of R3HCC1 wildtype and truncations; R3HCC1: 1-209 a.a. (N ter) and 238-440 a.a. (C ter) in EGFPC2 and p3X-FLAG constructs and (D) Immunoprecipitation of EGFPC2-empty vector and R3HCC1 wildtype, N ter and C ter in HEK293T; anti-GFP used for pull-down, probed for the endogenous G3BP1.

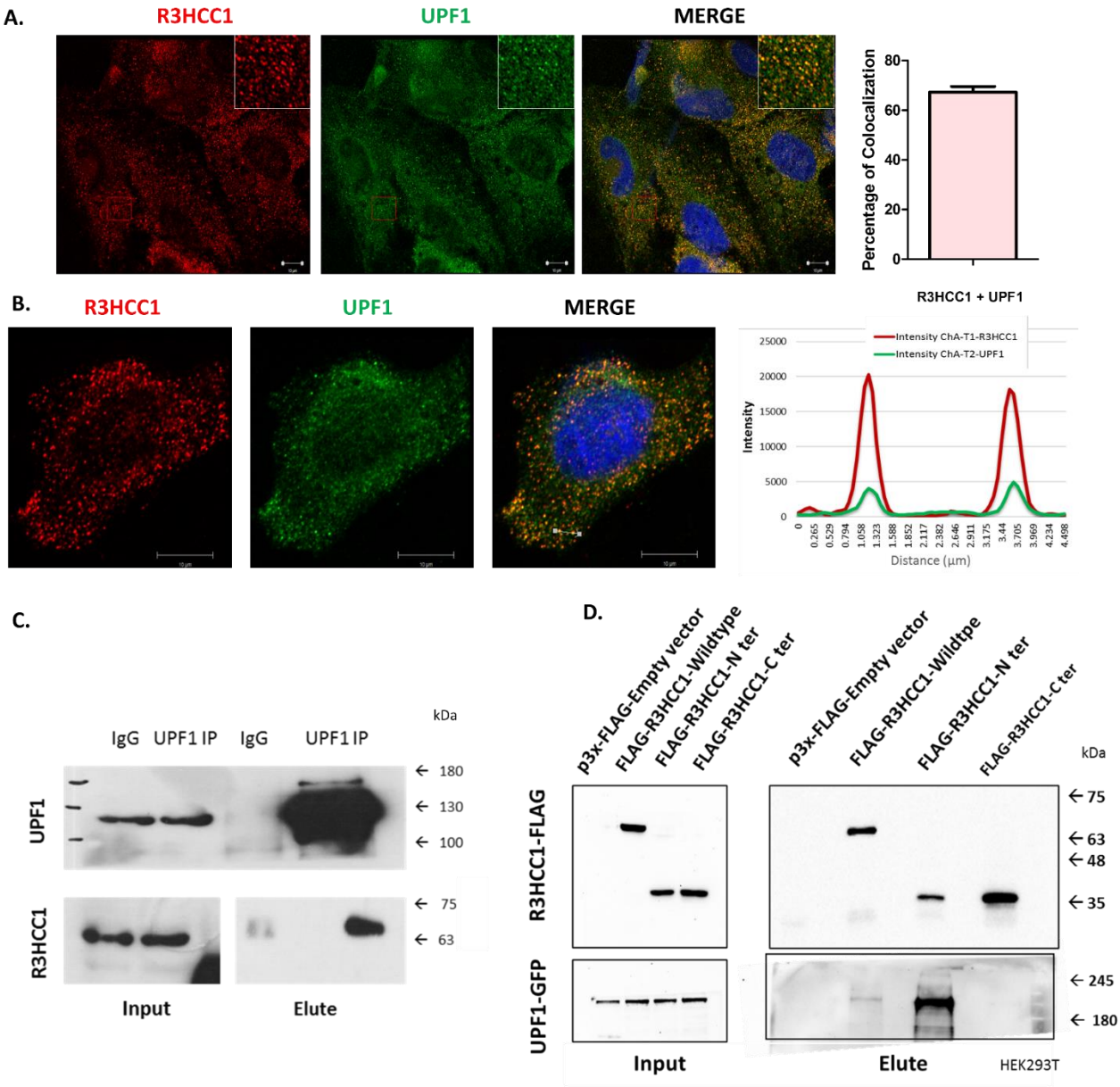


Figure 3.11. Interaction of R3HCC1 with UPF1: (A) Immunofluorescence images to depict colocalization of endogenous R3HCC1 and UPF1. Pearson's correlation coefficient was calculated using Zeiss Zen colocalization software to plot the percentage of colocalization, n=15 cell for three independent experiments, (B) Line profile depicting overlap in the intensity of R3HCC1 (red) and UPF1 (green), (C) Immunoprecipitation of endogenous R3HCC1 with UPF1 using anti-UPF1 antibody and (D) Immunoprecipitation of p3X FLAG empty vector and R3HCC1 wildtype, N ter, and C ter in HEK293T; anti-FLAG was used for pull-down and probed for overexpressed UPF1-GFP. A very faint band for UPF1-GFP with FLAG-R3HCC1-C ter was observed here, which was confirmed on high exposure of the blot (not shown here).

3.3.6. R3HCC1 does not interact with exon-junction complex protein

Exon-junction complex protein is an RNA-binding protein complex consisting of three core proteins: EIF4A3 (DDX48), RBM8A (Y14/Tsunagi), and MAGOH (Mago Nashi). This complex plays a vital role in regulating diverse aspects of mRNA metabolism, which includes nucleocytoplasmic mRNA export, subcellular localization, quality control, translation, and non-sense mediated decay. Kashima and colleagues identified a conserved sequence to be present SMG6 and UPF3b, which would confer direct binding of these proteins to exon-junction complex and called it exon-junction complex binding motif (Kashima et al., 2010). In this study, they identified R3HCC1 and its paralogue R3HCC1L proteins to have the exon-junction complex binding motif, and additionally, R3HCC1L was shown to interact with exon-junction complex proteins, Y14 and MAGOH. We therefore tested for similar interaction with R3HCC1 and found that none of the three exon-junction complex proteins interact with R3HCC1 (Figure 3.12).

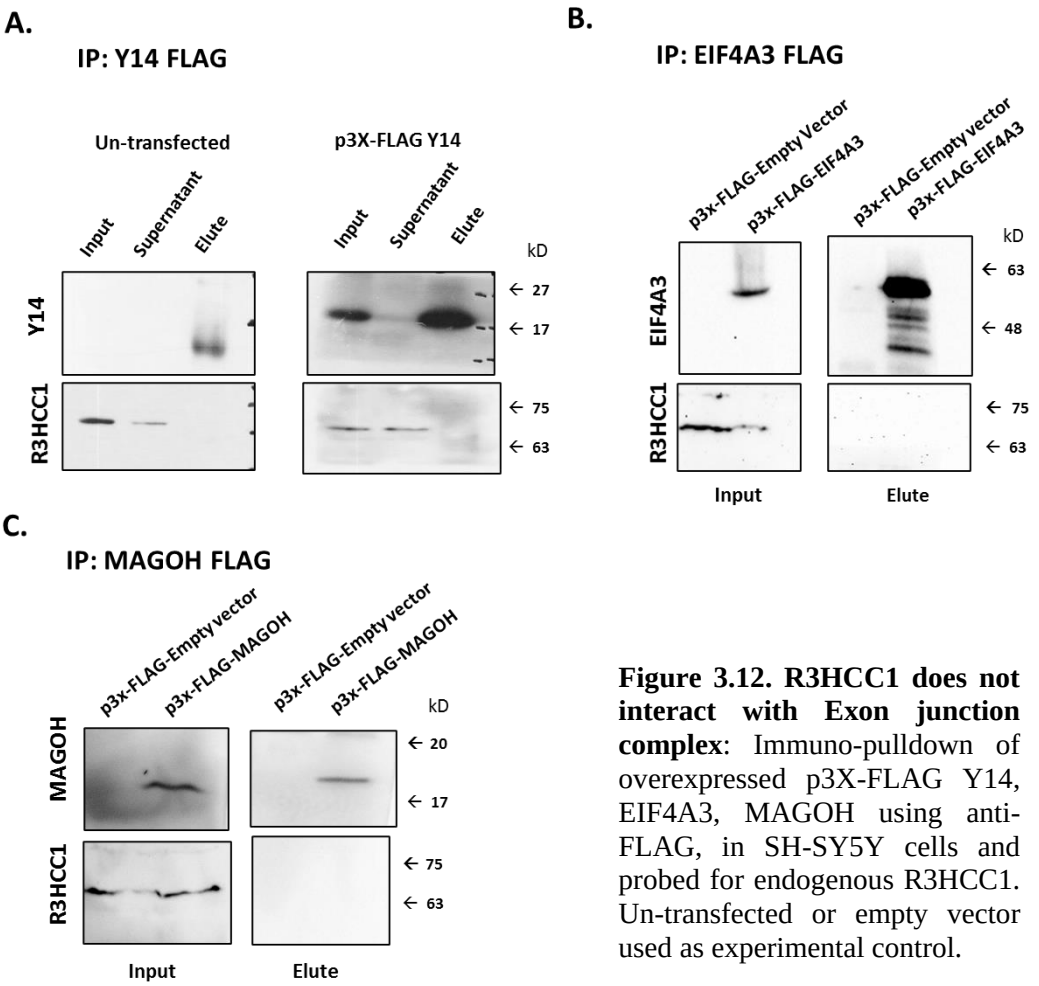


Figure 3.12. R3HCC1 does not interact with Exon junction complex: Immuno-pulldown of overexpressed p3X-FLAG Y14, EIF4A3, MAGOH using anti-FLAG, in SH-SY5Y cells and probed for endogenous R3HCC1. Un-transfected or empty vector used as experimental control.

3.3.7. *CG2162* is a *Drosophila* ortholog of human *R3HCC1*

The predicted ortholog of *R3HCC1* in *Drosophila melanogaster*, named *CG2162* was analyzed for its gene sequence. The longest isoform of *CG2162* codes for a transcript of 2560 bp and protein of 662 amino acids (A.1.7). CG2162 has three protein isoforms, and the longest isoform shows a 35.6% similarity with human R3HCC1. The domains R3H and RRM exhibit 77% and 64% similarity, respectively (A.1.6). The mutation associated with the HWE 244 family is conserved in *Drosophila* at the amino acid position p.Arg639.

I confirmed the transcript expression of *CG2162* in the *Drosophila* adult brain in w^{1118} wildtype control flies (A.1.8). Further, I used an enhancer trap line NP3333 to examine the tissue-specific expression of *CG2162*. NP3333 line contains a P element of 11.3kb in *CG2162* (A.1.10) that consists of Gal4/LacZ under weak promotor sensitive to enhancer regulation, and hence, the presence of an enhancer upstream of the gene of interest will drive the expression of Gal4/Lacz. The NP3333 line was crossed with a UAS-GFP line; herein, the Gal4 binds UAS and promotes GFP expression, which is the read-out for *CG2162* expression. The GFP expression was studied in *Drosophila* 3rd instar larval and adult brain using immunohistochemistry. In the larval brain, abundant expression was seen in the ring gland (A.1.9). The expression was also observed in the innervations from the ring gland extending towards the cerebral lobes and ventral cord (A.1.9). A prominent expression in cell bodies was detected in the cerebral lobes as well (A.1.9). The adult brain had expression in the cell bodies lateral and dorsolateral to antennal lobes, and I observed one or two cells innervating projection into antennal lobes. Expression in a cluster of cells in the anterior-most region of the brain was also observed (A.1.9). Using the enhancer-trap system, we can propose that *CG2162* is expressed in the fly brain; however, this must be confirmed using a more direct staining method with antibodies targeted against CG2162 protein.

3.3.8. NP3333 larvae and adult flies are sensitive to heat stress

Since the NP3333 line consists of a transposable P element inserted in exon 5 of *CG2162* (A.1.10), I checked for the presence of *CG2162* at the transcript level in the line. The complete transcript *CG2162* was absent in the NP3333 line (A.1.11). However, upon using primer pairs within the transcript, P element insertion was confirmed at the transcript level as well, which would cause a change in the translational frame,

subsequently leading to the synthesis of truncated protein (A.1.11). However, this needs to be confirmed by immunoblotting.

We used the NP3333 insertion mutant line to examine certain behavioral aspects. In HWE, the temperature is an essential component along with the touch stimulus to trigger seizures. *Drosophila melanogaster* represents an ideal model for studying mechanisms underlying the neuronal response to acute heat stress due to its small size, large surface-to-volume ratio, and inability to regulate body temperature internally. Hence, NP3333 and control w^{1118} flies were subjected to heat shock treatment at 42ºC, and their response was assessed. It was observed that after 25-35 seconds into the exposure, both the sexes of NP3333 flies fall to the bottom of the vial and exhibit seizure-like activity. The seizure-like activity was described by leg twitching, muscle tremors, and jerks. Most flies become paralyzed over the time interval of 2 minutes (Figure 3.13A). Controls did not display any apparent behavioral phenotype initially, but around 1 minute after the onset of warm temperature, flies settled at the bottom of the vial without displaying any seizure-like activity or paralysis. Vials were removed from the water bath after 2 minutes, and the time taken by NP3333 flies to recover from the paralysis and gain an upright position was also quantified (Figure 3.13B). While recovering, seizure-like activity was observed occasionally in a few NP3333 flies.

Video for heat shock assay can be accessed: Left vial: control w^{1118} flies, right vial: NP3333 flies. If the below link does not open upon clicking, please copy and paste the link as a web URL.

https://drive.google.com/drive/folders/1yDLcIeSoD6uSwHYyt_4AxI_cFIQC-1Q7?usp=sharing

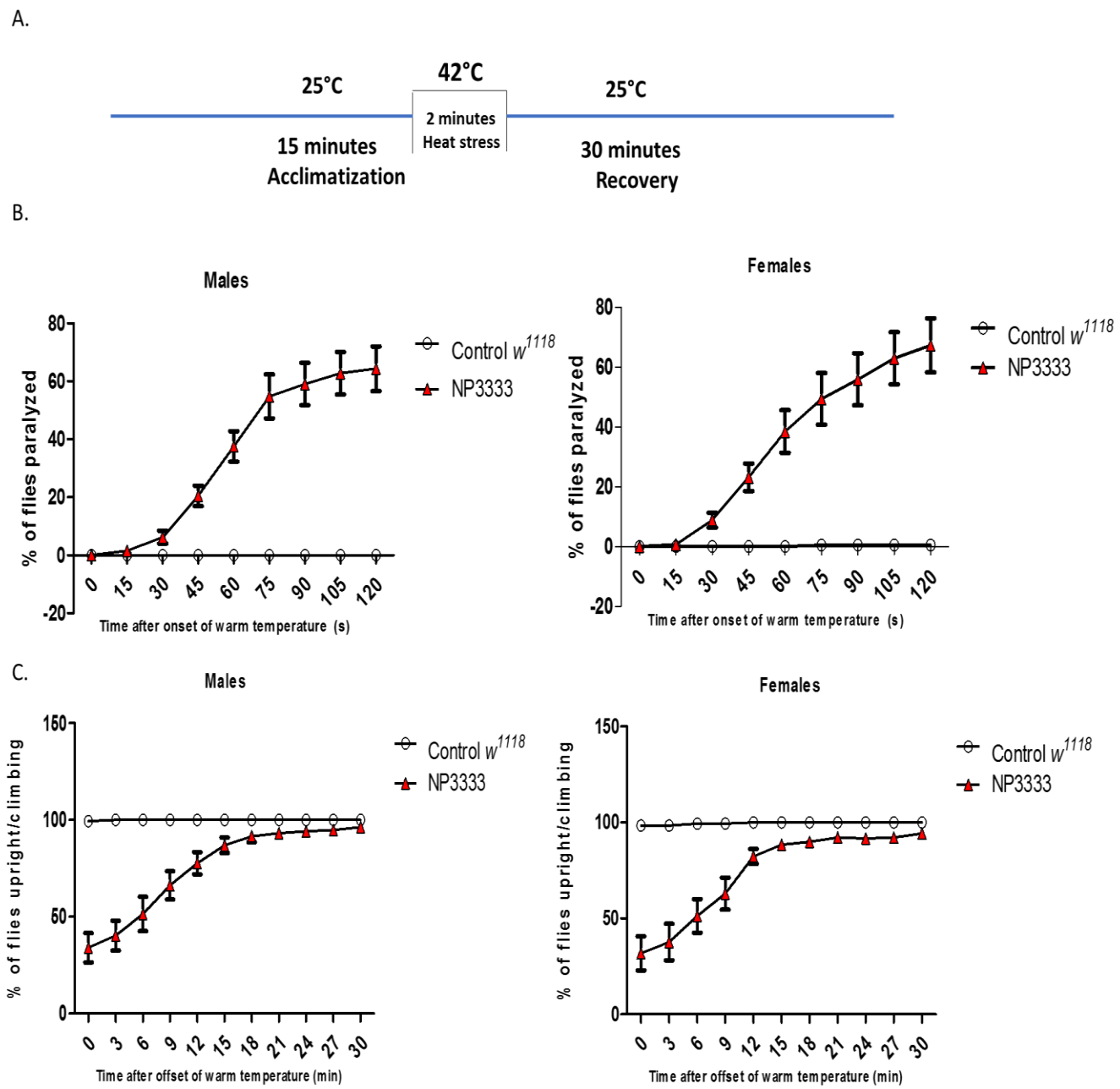


Figure 3.13. NP3333 flies show sensitivity to heat stress: (A) Schematic representation of heat treatment given to the flies, (B) The percentage of flies exhibiting paralysis at an interval of 15 seconds for 2 minutes is plotted for male and female control and NP3333 flies, (C) Percentage of flies that are upright/climbing prior to removal of heat stress is plotted male and female control and NP3333 flies. Data is represented as mean ± SEM across 14 trials.

Since adult NP3333 flies exhibit sensitivity towards warmer temperatures, NP3333 larvae were also assayed for temperature sensitivity. NP3333 and control larvae were incubated at 37ºC for 2 hours, and eclosion over the next few days was analyzed. The number of NP3333 flies that emerged was significantly lower than that of control, indicating heat stress impacts eclosion in NP3333 flies (Figure 3.14). The same experiment was also performed at the optimum temperature, i.e., 25ºC, and no change in fly's emergence was detected, concluding heat sensitivity in NP3333 larvae.

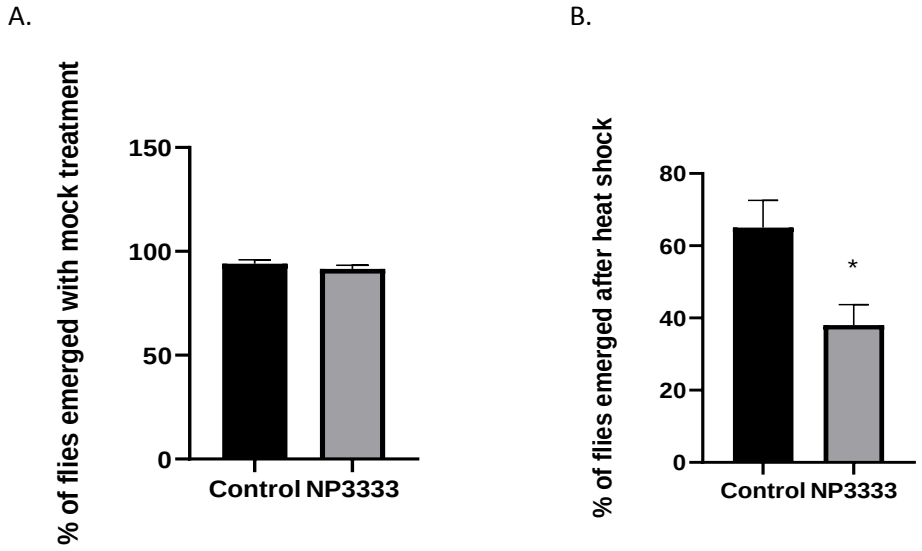


Figure 3.14. NP3333 larvae are heat-sensitive: (A) Percentage of flies eclosed for mock treatment at 25ºC in Control w^{1118} and NP3333 larvae. (B) Percentage of flies eclosed for control w^{1118} and NP3333 larvae after heat shock treatment. The number of larvae per trial= 20. Data are represented as mean ± SEM across ten trials. Statistical analysis was done using a two-tailed unpaired Student's t-test, *$p<0.05$.

3.4. Discussion

Our previous work comprising genome-wide linkage and whole genome sequencing studies suggested the involvement of *R3HCC1* in HWE. *R3HCC1* is expressed in various different tissues and cells, including the brain. Towards understanding its biological functions, I found an interesting lead from R3HCC1 protein localization studies which suggested that R3HCC1 is recruited to stress granules (SGs) under heat and oxidative stress. SGs are mRNP granules formed during cellular stress conditions like heat, oxidative, nutrient deprivation, endoplasmic reticulum stress, and osmotic shock, among others. The stress is sensed by stress-sensitive kinases - HRI, PERK,

PKR, and GCN2 - which phosphorylate translation initiation factor, eIF2a (Anderson and Kedersha, 2002), resulting in stalled translational complexes. These complexes become a site for the recruitment of additional RBPs as well as non-RBPs, which eventually coalesce to form SG (Protter and Parker, 2016). Upon removal of stress, these granules disassemble either via ATPase-driven remodeling events or via autophagy (Mazroui et al., 2007; Buchan et al., 2013; Jain et al., 2016). Mutations in several RBPs, such as FUS, TDP-43, TIA, hnRNPA1, are known to result in aberrant regulation of SG formation or disassembly, leading to neurodegenerative diseases like amyotrophic lateral sclerosis and frontotemporal dementia (reviewed in Li et al., 2013; Ramaswami et al., 2013). In my study, I observe that the overexpression of two R3HCC1 variants, p.Val142Met and p.Ala433Thr, causes an increase in the SG size during heat stress. This increase in size could be either a consequence of additional sequestration of proteins or aberration in the dynamic property of SGs, which allows the exchange of RNA and proteins with surrounding cytoplasm and other cytoplasmic granules. Interestingly, amyotrophic lateral sclerosis-linked FUS (Baron et al., 2013; Vance et al., 2013) and TDP43 (Dewey et al., 2011) mutants also cause a similar morphological change in SG. Further studying the persistence of SGs in cells overexpressing R3HCC1 variants will reflect whether the variants impact assembly or disassembly of these granules. Curiously, I also noticed that overexpression of the N-terminal fragment of R3HCC1 (1-147a.a), which consists of the predicted RNA binding domain R3H and lacks RRM, leads to a reduction in the number of cells with SGs. I also observed that overexpression of R3HCC1 in HEK293T cells leads to formation of G3BP1 stained granules without stress stimuli (data not shown) and implications of this need further studies. These observations suggest that R3HCC1 might play a role in SG dynamics. Induced hyperthermia in mice leads to the formation of TIAR-positive and polyadenylated mRNA-positive granules in the cytoplasm of hippocampal and cortical neurons (Shelkovnikova et al., 2017), suggesting that temperature sensed via different "mechanisms" can induce a stress response in neurons and the proteins involved in SG dynamics during such an event may play a role in maintaining cellular homeostasis.

Another type of RNA granule that I discuss in the study is the neuronal transport granule, which plays an essential role in the localization of RNA to different "compartments" in neurons and contributes to the local translation and synaptic plasticity (Kiebler and Bassell, 2006). I show an RNA-dependent association of R3HCC1 with microtubules,

which are key players in facilitating the transport of RNA. However, it is important to examine if R3HCC1 is a part of these neuronal granules using colocalization studies with neuronal granules proteins- STAU1, SMN1, and PUMILIO2, among others.

The studies conducted also show that R3HCC1 interacts with two RBPs: G3BP1 and UPF1. Apart from its role as a critical regulator of SG assembly, G3BP1 is also present in neuronal granules in the dendrites and axons (Sahoo et al., 2018) and is involved in neuronal plasticity and calcium homeostasis in the hippocampus (Martin et al., 2013). G3BP1 knockout mice exhibit impaired motor coordination and ataxia-like phenotype (Martin et al., 2013). UPF1 is known to play an essential role in the non-sense mediated decay pathway, an mRNA surveillance mechanism that eliminates transcripts with a premature stop codon (Bhattacharya et al., 2000; Chakrabarti et al., 2011; Kurosaki and Maquat, 2016). UPF1 is also known to be present in the transport RNA granules in the dendrites of rat hippocampal neurons, where its interaction with STAU2 plays a role in local translation and mGluR -mediated long term depression (Tyson et al., 2017). A study by Mooney and colleagues suggests the importance of UPF1 in neurological disorders like epilepsy, wherein its levels are high in resected hippocampus from patients with intractable temporal lobe epilepsy (Mooney et al., 2017). This study shows that status epilepticus in mice have enhanced levels of UPF1, leading to increased NMD in hippocampal cells, which in turn, could cause aberrant regulation of transcripts. Also, a recent study has identified that G3BP1 and UPF1 together regulate structure-mediated RNA decay (Fischer et al., 2020). Based on our observations, we propose R3HCC1's involvement in the process of RNA decay pathways as well, including NMD. While the localization of R3HCC1 to P bodies was not observed, several studies have indicated that P bodies are not the only site for NMD, and its depletion does not alter endogenous NMD substrates (Eulalio et al., 2007; Stalder and Mühlemann, 2009). R3HCC1 also contains an exon-junction complex binding motif, which would facilitate its association with EJC protein Y14, MAGOH, and EIF4A3 (Kashima et al., 2010). However, in co-immunoprecipitation studies, I did not observe an interaction between R3HCC1 and exon-junction complex proteins Y14, MAGOH, and EIF4A3.

Based on the findings described above, I propose that R3HCC1 may play a role in mRNA surveillance. RBPs are associated with various neurological disorders, including epilepsy. FMRP is one of the RBP whose loss-of-function causes Fragile X syndrome, characterized by intellectual disability, autism, and seizures (Qiu et al., 2008). Deletion

mutants in the gene RBFOX1, which codes for a neuronal splicing factor, cause a wide range of neurological disorders, including genetic/idiopathic generalized epilepsy, childhood focal epilepsy, and self-limited childhood benign epilepsy and autism (Fogel et al., 2012; Lal et al., 2013). Other examples include Pumilio-2 and CELF4, where mice deficient in these proteins manifest epilepsy (Wagnon et al., 2012; Follwaczny et al., 2017). These RBPs have been shown to regulate genes involved in neuronal excitability, and hence, the misregulation of these proteins leads to abnormal hypersynchronous activity in the brain. In the future, studying R3HCC1's RNA-interacting partners is proposed to gain insights into its role in RNA -processing and its impact on neuronal networks.

Although preliminary in nature, our experiments on *Drosophila* provide an initial insight into the behavioral impact of loss-of-function of R3HCC1's ortholog. *Drosophila* has been extensively used to study human development, behavior, and genetic disorders. Among many neurological disorders examined using *Drosophila*, epilepsies have received considerable attention. Much like the genetic mutants that cause epilepsy in humans, there are seizure-sensitive mutants in *Drosophila*. The seizure phenotype in flies is characterized by leg shaking, abdominal contractions, wing flapping, and scissoring, and proboscis extension (Parker et al., 2011). The seizures often lead to flies becoming paralytic. Bang-sensitivity is the most common phenotype screened for in flies wherein the flies are subjected to mechanical shock via tapping the vial or brief vortexing to induce seizures (Pavlidis and Tanouye, 1995; Parker et al., 2011). Apart from the behavioral phenotypes, seizures are also studied using electrophysiology approaches (Pavlidis and Tanouye, 1995; Lee and Wu, 2002). There are temperature-sensitive paralytic mutants, for example, *shibire(shi^{ts1}), paralytic (para^{TS1}), cacophony(cac^{TS2}), maleness(mle^{napts})* and *seizure* (*sei*) which exhibit seizure phenotype at a higher temperature. These mutants behave normally at the temperature, 23°–25°C, but manifest paralysis at a temperature ranging from 29°–42°C. *shi* codes for Dynamin, which mediates endocytosis and vesicle recycling which are important for chemical synaptic transmission (Van Der Bliek and Meyerowrtz, 1991). *shi^{ts1}* mutant flies at 29°C, exhibit the seizure-like phenotype of behavioral hyperexcitability followed by paralysis and a loss of chemical synaptic transmission (Siddiqi and Benzer, 1976; Salkoff and Kelly, 1978; Koenig and Ikeda, 1989). The gene *cac* codes for a N-type Ca^{2+} channel required for presynaptic neurotransmitter release, and its mutant *cac^{TS2}* also

shows a similar phenotype as *shi*^{ts1} at a temperature of 38°C (Kawasaki et al., 2000; Rieckhof et al., 2003). The *mle* gene encodes an RNA helicase protein, and its mutant *mle*^{napts} at 37°C exhibit loss of action potential, leading to behavioral paralysis and an increased action potential refractory period at room temperature (Wu and Ganetzky, 1980; Lee and Hurwitz, 1993). *Sei* in flies codes for voltage-gated potassium channel, an ortholog of the mammalian ERG channel family. Neuronal-specific knockdown of *Sei* leads to heat-induced seizure and paralysis (Hill et al., 2019). A similar phenotype is manifested by a mutant *para*^{TS1} allele at an elevated temperature of 29°C. The *para* gene encodes a voltage-gated Na^+ channel alpha subunit, and different mutant alleles for the gene exhibit variability in the severity of the phenotype (Siddiqi and Benzer, 1976; Wu and Ganetzky, 1980; Loughney et al., 1989; Ramaswami and Tanouye, 1989; Suzuki et al., 2004a). Mutations in the *SCN1A* sodium channel gene cause a wide spectrum of human epilepsy disorder, including genetic epilepsy with febrile seizure (GEFS+), wherein febrile seizures are observed after six years of age (Scheffer and Berkovic, 1997). In a study by Sun and colleagues, knock-in SCN1A mutation (K170T) in the *Drosophila* sodium channel gene, *para*, led to a semi-dominant temperature-induced seizure phenotype (Sun et al., 2012).

In this part of the work, I show that *R3HCC1* ortholog, named, *CG2162* is expressed in the fly larval and adult brain at the transcript level. It was observed that the *CG2162* mutant flies are sensitive to heat stress (42ºC) and exhibit seizure-like behavior, ultimately leading to a paralysis-like state followed by a slow recovery. The validation of the behavioral phenotype was done using NP3333 line that has been backcrossed for six generations. Looking forward, we propose employing a CRISPR-Cas9 -based knockout system to generate *CG2162* mutations for further studies. It is known that SGs are formed in *Drosophila* in response to heat stress, and they contain homologs of mammalian SG proteins like FMRP, G3BP, and TIA-1 (Farny et al., 2009; Aguilera-Gomez et al., 2017). TORC2 in flies has been shown to participate in SG assembly, and its knockout results in sluggishness and paralysis at high temperatures (Jevtov et al., 2015). On the other hand, dFMRP localizes to SG and neuronal granules, where it promotes the cellular trafficking of its mRNA target. *dFMR1* mutants exhibit delayed motor development, locomotor hyperactivity, defects in courtship, and social interaction (Bolduc et al., 2010; Kashima et al., 2017). The mutants also show increased synapse arborization and branching, increased synaptic bouton numbers, and elevated

neurotransmission at neuromuscular junctions (Zhang et al., 2001; Doll and Broadie, 2016; Drozd et al., 2018). Localization of the CG2162 to the mRNP granules and its role in the formation of SG shall be studied in flies. Since proteins associated with mRNP granules play an essential role in synaptic development, *CG2162* mutants may be examined synaptic abnormalities.

R3HCC1's predicted RNA-binding domains, R3H and RRM, are highly conserved with 77% and 64% similarity in *Drosophila*, respectively. In *Drosophila*, mutations in RNA-binding proteins, *pumilio* and *cpo* have severe phenotypic consequences on nervous system functions. The *pumilio* gene encodes a translational repressor and functions in embryogenesis, germ-line stem cell development, and larval ovary development (Wreden et al., 1997; Schweers et al., 2002). Schweers and colleagues demonstrated its role in maintaining proper neuron excitability. The *pumilio* mutant (*bemuse*), with P element insertion located within the *pum* transcription unit, exhibits sluggishness and increased motor neuron excitability (Schweers et al., 2002). Another such gene is *couch potato* (*cpo*), which also encodes a *Drosophila* RNA-binding protein, which when impaired, causes several neurological abnormalities, including epilepsy. Mutation of *cpo* leads to bang-sensitive paralysis, seizure susceptibility, and synaptic transmission defects (Glasscock and Tanouye, 2005). Interestingly, the first line of experiments done to evaluate the role of R3HCC1 indicated its involvement in regulating the normal functioning of the nervous system. Several available examples of existing genetic mutants in the RNA-binding protein and their phenotypic defects shall form the basis of our future study for understanding the role of R3HCC1 at molecular, cellular, and organismal levels.

Chapter 4

Genetic and cell biological aspects of *ZGRF1*, a potential hot water epilepsy gene

To reinvestigate the genetic involvement of the ZGRF1 gene in a multi-generation, multi-affected family, HWE227, I examined an affected member of the family employing a whole genome sequencing study. A total of 36,392 variants were obtained at Chr4q24-q28 (HWE2, MIM: 613340), a locus previously shown to be linked to the disorder in the family. Among these, four heterozygous rare variants were found in the region of interest, which co-segregated with the disorder in the family. These four variants: c.1805C>T in ZGRF1, c.*2206G>A in EGF, c.*3748G>A in FAM241A, and g.111883909T>C in an open chromatin region, were present with minor allele frequency (MAF)<0.01 in the ethnically matched control individuals. The bioinformatic analysis for the exonic variant c.1805C>T in ZGRF1 and g.111883909T>C in open chromatin region indicated high conservation among mammals. On the other hand, the two 3'UTR variants were poorly conserved; however, both present a predicted micro-RNA binding site. Here, I describe the genetic study carried out further supporting the role of ZGRF1 as a potential candidate gene for HWE. ZGRF1 (Zinc finger GRF-type containing 1) has been reported to be involved in DNA damage repair response. ZGRF1 localizes to the nucleus and centrosome, spindle poles, and mid-body during cell division. In primary rat hippocampal neurons, under basal conditions, ZGRF1 is observed in the nucleus, cytoplasm, and proximal dendrites. Upon overexpression of ZGRF1 patient mutants in cultured HEK293 cells, mitotic defects were observed at different cell cycle stages. ZGRF1 cells carrying a CRISPR/Cas9 -generated, N-term deletion of the DUF2439 domain exhibit an increase in mitotic defects and G2/M phase arrest when compared to the parental wildtype cells. In this chapter, I also discuss the generation of ZGRF1 knockout in mice and the prospective future studies.

4.1. Introduction

A genome-wide linkage study of a four-generation HWE family (HWE277, Figure 4.1) led to the identification of a locus at chromosome 4q24-q28, linked to HWE (Ratnapriya et al., 2009a). The critical region at chromosome 4, situated between the markers

D4S1572 and D4S2277, contains 24 megabases of the genome. To analyze the genetic variants in the region, whole exome sequencing was performed for two affected members of the family, III:4 and IV:3. This helped identify a rare variant, c.1805C>T (p.Thr602Ile) in the gene, *ZGRF1* (Zinc Finger GRF-Type Containing 1), to be co-segregating with clinical phenotype. On further examining the gene in 288 HWE patients, five additional rare variants were found: c.977G>A (p.Arg326Gln), c.1979A>G (p.Glu660Gly), c.5584C>T(p.Arg1862*), c.5818T>C (p.Phe1940Leu) and c.5951A>G (p.Asp1984Gly) (Shalini Roy Chowdhary, Ph.D. Thesis, 2018). To enhance the coverage in other regions of the genome, such as UTRs, enhancers, promoter region, lncRNA, and miRNA, I performed whole genome sequencing for the patient IV:4 in the family. In this chapter, I present variants found in whole genome sequencing analysis and discuss probable variants underlying HWE in the family. Further, certain functional attributes of the candidate gene, *ZGRF1* are discussed.

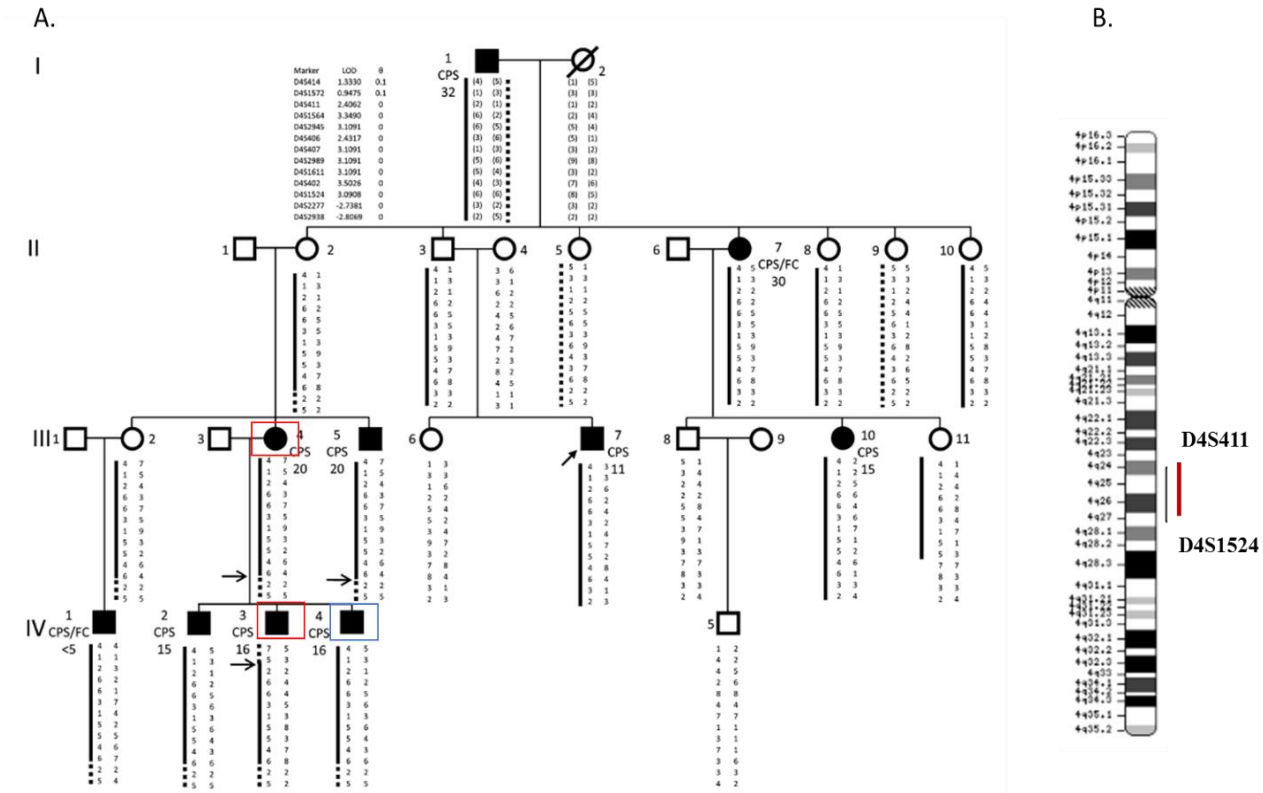


Figure 4.1. A 13-marker haplotype segregation in HWE 227 family: (A) Pedigree of family HWE227: Males are denoted by squares, circles denote females, affected individuals are shown with filled symbols, and empty symbols represent unaffected members. I, II, III, IV indicate generations. Proband is denoted by an arrow pointing towards the symbol. The HWE-linked haplotype is represented as solid lines. The arrow in III:4, III:5, and IV:3 indicates key recombinant events. Red and blue boxes indicate patients whose DNA was used for which whole exome and whole genome

sequencing, respectively. Seizure types and age at onset are indicated beside symbols: CPS, complex partial seizure; FC, febrile convulsions (Ratnapriya et al., 2009a). (B) Representation of chromosome 4 at which the segregating haplotype 4q24-q28 with marker boundaries D4S411 and D4S1524 is present spanning 24Mbp.

4.2. Material and methods

4.2.1. Whole genome sequencing

Whole-genome sequencing for the affected individual, IV:4 was done using NEB NEXT Ultra II DNA library prep kit (New England Biolabs) (for details, see Chapter 2, pages 24-25). Sequencing libraries were prepared with an average insert size of 350bp (A.1.1). The library quality and quantity were checked using the Qubit HS dsDNA kit (Thermo Fisher Scientific) and Agilent Bioanalyzer DNA 1000 kit (Agilent Technologies). The QC-passed libraries were sequenced for 150 bp paired-end reads on an Illumina HiSeqX platform. The quality of the sequence data was checked using FastQC and MultiQC (Ewels et al., 2016) software. The data was also checked for base call quality distribution, percentage of bases above Q20, Q30, percentage of GC content, and sequencing adapter contamination (A.1.2). The sequence data was processed using Trimgalore to remove adapter sequences and low-quality reads. The QC passed reads were mapped to the human reference genome build GRCh38 that was provided in the GATK Resource Bundle using the BWA-MEM algorithm (Li and Durbin, 2009). The alignments were sorted, indexed, and PCR-duplicates were marked and removed using Picard tools. GATK (McKenna et al., 2010) work-flow for short variant discovery (SNPs + INDELs) using GATK v4 was followed for variant calling. The variants were annotated according to GRCh38 assembly using the Variant effect predictor (McLaren et al., 2016). The variants were also annotated separately using the gnomAD v3.1.1 data.

4.2.2. Sequence validation and segregation

Primers were designed to validate the variants obtained from the whole-genome sequencing study (A.2.8). All amplicons were sequenced in the individual IV:4 to check if the variant is a true variant, followed by checking for co-segregation with HWE in the family. The segregating variants were examined in 200 ethnically matched control individuals to estimate their allele frequencies.

4.2.3. Plasmid and antibodies

A 421bp length 3'UTR spanning the *EGF* variant c.*2206G>A, and 397bp spanning the *FAM241A* variant c.*3748G>A were amplified from control-genomic DNA and cloned in the vector pMIR-REPORT™ Luciferase using specific primers with restriction site, Spe1 and Mlu1 (A.2.9). Site-directed mutagenesis using a Quickchange site-directed mutagenesis kit (Agilent Technologies) was done using specific primers to create the desired mutation (A.2.9). All the constructs were sequenced confirmed. The wildtype *ZGRF1* clone in Myc-DDK tagged pCMV6 vector was obtained from Origene, Maryland, USA. The six mutations were incorporated individually using QuikChange II XL Site-Directed Mutagenesis kit (Agilent Technologies) (Shalini Roy Chaudhary, Ph.D. Thesis, 2018).

Antibodies used were: Anti-ZGRF1 antibody (ab122126, Abcam), anti-α tubulin antibody (T5168, Sigma), anti-gamma tubulin antibody (T6557, Sigma) and anti-MAP2 (M9942, Sigma)

4.2.4. Luciferase assay

HEK293T cells were plated in 6-well plates with 1 ml of DMEM growth medium. Transfections were carried at 60% confluency with 1µg of pMIR-REPORT™ Luciferase and 200 ng of β-galactosidase expression vector pCMVβ using LipofectamineTM 2000 (Invitrogen-Life Technologies), in a serum-free medium and replaced with complete DMEM after 4-6 hours of transfection. After 24-30 hours of transfection, cells were washed with 1X PBS and lysed in 200µl of reporter lysis buffer (Promega, Wisconsin, USA) for 1 hour with vortexing the lysate at regular intervals. The lysate was centrifuged at 15,000 rpm for 20 minutes at 4ºC, and the supernatant was collected for the luciferase assay. The lysate was diluted at a 1:50 ratio in lysis buffer, from which 1 µl was mixed with 10 µl of Luciferase Assay Reagent (Promega), and the luciferase activity was measured in the scintillation counter using the C14 window. For measuring β-galactosidase enzyme activity, 4µl of lysate was added to 16µl of lysis buffer following the addition of 20µl of Assay 2× Buffer 2 (Promega) that contains the substrate ONPG. The mix was incubated at 37ºC for 30 minutes, and the reaction was stopped by adding 50 µl of 1M sodium carbonate. The β-galactosidase activity was measured on an ELISA reader using a 420nm filter. To normalize for transfection efficiency, the luciferase activity was presented relative to the β-galactosidase activity. The values of relative

luciferase activity were normalized to wildtype, and differences in luciferase activity were analyzed using the two-tailed Student's *t*-test. All values are plotted as mean ± standard error of the mean (SEM). P values ≤ 0.05 were considered to indicate a statistically significant difference.

4.2.5. Cell culture, cell cycle synchronization, and transfections

Mammalian cells were maintained in DMEM supplemented with 10% FBS, 2 mM L-glutamine, 100 U/ml penicillin, and 0.1 mg/ml streptomycin (Sigma), in a humidified atmosphere of 5% CO_2 and 37°C. For transfection, cells were seeded onto coverslips coated with poly-L lysine in 35mm dishes or 60/90mm dishes (Eppendorf). Upon attaining 60% confluence, cells were subjected to antibiotic-free and serum-free DMEM and transfected with the plasmid of interest using calcium phosphate transfection. A 100µl of the mix was made by adding the plasmid of interest, 12.5mM of $CaCl_2$, and deionized water. This mix was slowly added to 100µl of 2X HEPES Buffer Saline (280mM NaCl, 1.5mM Na_2HPO_4, 50mM HEPES), and the combination was added to the dish. After 5-6 hours of transfection, the medium was replaced by a complete DMEM medium, and 24-hours post-transfection, cells were processed for immunofluorescence and immunoblot assays. To study mitotic defects, cell cycle synchronization was performed to enrich the pool of mitotic cells. Post-transfections, cells were treated with culture media consisting of 2mM thymidine (Sigma) for 14 hours followed by 9 hours in a culture medium containing 24µM deoxycytidine (Sigma). The cells were further processed for immunocytochemistry.

4.2.6. Primary hippocampus neuronal culture

Primary hippocampal neuron cultures were prepared from Sprague Dawley rats at embryonic day 18 (Kaech and Banker, 2006). The cells were plated at the density of 20,000-30,000 cells/cm^2 on poly-L-lysine (Sigma Aldrich) in Minimum Essential Media (Thermo Fisher Scientific) supplemented 10% FBS. The media was changed to a neurobasal medium (Thermo Fisher Scientific) containing B27 (Thermo Fisher Scientific) and Glutamax (Invitrogen-Life technologies) after 3 hours. Neurons were cultured at 37°C in a humidified atmosphere with 5% CO_2. The cells were processed for treatments and staining on Day 14 (DIV14). The animal work was carried out following the procedures approved by the Institutional Animal Ethics Committee and the Institutional Biosafety Committee, inStem, Bangalore, India. The animals were

maintained at 20–22°C temperature, 50–60 relative humidity, 0.3 μm HEPA-filtered air supplied at 15–20 ACPH, and 14-h/10-h light/dark cycle and were freely fed with food and water.

4.2.7. Immunocytochemistry

For immunofluorescence assay, cells were washed thrice in 1X PBS and fixed with 2% PFA for 15 minutes at room temperature. For endogenous staining of ZGRF1 and γ-tubulin in HEK293, cells were fixed with 100% cold methanol for 5 minutes at -20°C. For staining of overexpressed ZGRF1 wildtype and variants and endogenous α-tubulin, the cells were fixed with 2% PFA for 15 minutes. The cells were permeabilized using 0.1% Triton X-100 for 10 minutes, followed by incubation in 5% BSA blocking solution for 4 hours at 4ºC. Next, the cells were treated with the required primary antibody; overnight at 4ºC for ZGRF1 and others, 1 hour at room temperature, followed by incubation with 1:500 secondary antibody-conjugated with Alexa 488/Alexa 568 for 4 hours at 4ºC. The cells were washed in 1X PBS and stained for the nucleus using DAPI (Sigma) for 15 minutes. The coverslips were mounted on a glass slide in PVA-DABCO mounting media (Sigma). For dual antibody staining procedures, antibodies were mixed equally in a 1:1 ratio while treating the cells for both primary and secondary antibodies. For immunostaining of rat hippocampal neurons, the cells were fixed with 2% PFA for 20 minutes at room temperature, then permeabilized using 0.1% Triton X-100 for 10 minutes, followed by incubation in blocking solution (2% FBS, 2% BSA, 0.1% TritonX-100) for 1 hour at room temperature. The cells were treated with the required primary antibody made in blocking solution for 1 hour followed by 1:500 secondary antibody and stained for nucleus using DAPI for 15 minutes. The imaging was done with LSM880 Meta confocal laser scanning microscope (Carl Zeiss) under 63X/1.4 oil immersion objective.

Antibodies used were anti-ZGRF1 rabbit raised (1:100), anti-α tubulin mouse raised (1:3000), anti-γ tubulin mouse raised (1:3000) and anti-MAP2 mouse raised (1:1000). All the antibody solutions were made in 1% BSA (in 1X PBS) except mentioned otherwise.

4.2.8. Generation of *ZGRF1* mutant cell lines

HEK293T cells were transfected with 1μg Cas9D10A-GFP plasmid and 500ng of sgRNA1 and sgRNA2 plasmid (Sigma) (A.2.10) each using lipofectamine transfection.

Forty-eight hours post-transfection, cells were dissociated using trypsin and resuspended in 2% DMEM growth media and filtered using 35µm FACS tube (BD Falcon, New Jersey, USA). The GFP positive cells were single-cell sorted into 96 well plates consisting of 50% fresh media and 50% conditioned media. The sorting was done using BD FACSDIVA 8.0.1. Twenty days after sorting, single colonies were shifted to a 24-well plate. Genomic DNA was extracted (For protocol, refer to A.2.18) and used for PCR amplification and Sanger sequencing. A 574bp of the CRISPR target site was PCR-amplified using ZGRF1-KO-exon 5-FP and ZGRF1-KO-exon-5-RP (A.2.11). The amplified product was purified and subsequently sequenced to identify INDELS present at the site. To identify mutants at the transcript level, RNA extraction and cDNA synthesis were done (For protocol, refer to A.2.20-21). PCR amplification was done using *ZGRF1*-cDNA-P1-FP and *ZGRF1*-cDNA-P1-FP (A.2.11), and Sanger sequencing was done to confirm the deleted region.

4.2.9. Cell cycle analysis with flow cytometry

HEK293T cells were seeded in a 60mm dish, and upon reaching 90% confluence, cells were washed with 1X PBS twice. The cells were fixed by adding 70% ethanol to the cell pellet and vortexed and incubated at 4°C for 30 minutes, followed by washes in 1X PBS twice. The pellet was spun at 2000rpm for one minute each in the washing step. The supernatant was discarded carefully to avoid cell loss. The cells were resuspended in 1X PBS with 200µg/ml of RNase A and incubated at 37°C for 2-4 hours. To this 50µg/ml propidium iodide (PI) was added and incubated at 37°C for 15 minutes. The cells were filtered into FACS tubes with 35µm mesh and used for cell cycle analysis. For analysis, two plots were generated: Forward scatter vs. side scatter plot; to identify cells and PI staining height vs. area plot; to differentiate single cells from doublets. From the single-cell population, cells in G1, S, G2, M phase were plotted for cell count vs. PI plot; using FlowJO software. Using this plot percentage of cells in all cell cycle phases was determined. The number of events counted per sample were 10,000.

4.2.10. Generation of *Zgrf1* knockout mouse and genotyping

The animal work was carried out following the procedures approved by the Institutional Animal Ethics Committee and the Institutional Biosafety Committee, JNCASR, Bangalore. The mice were maintained at 20–22°C temperature, 50–60 relative humidity, 0.3µm HEPA-filtered air supplied at 15–20 ACPH, 14-h/10-h light/dark cycle and were

freely fed with food and water. C57BL/6NCrl carrying *flox* allele in *Zgrf1* was procured (Cyagen US Inc, California, USA). In brief, the protocol used to generate *Zgrf1*$^{flox/+}$mice is as follows: the region harboring exons -5 and -6 of the *Zgrf1* gene was targeted/floxed by homologous recombination using a targeting vector in C57BL/6Ncrl mouse ES cells. The recombinant cells were selected by G418 resistance and confirmed by PCR amplification. The correctly targeted clone (1H5) was injected into C57BL/6NCrl albino embryos, which were then reimplanted into CD-1 pseudo-pregnant females. Founder animals were identified by their coat color, their germline transmission was confirmed by breeding with C57BL/6 females and subsequent genotyping of the offspring. One male and two female heterozygous targeted *Zgrf1*$^{flox/+}$mice were generated from clone 1H5. The mouse line was further bred and expanded at the JNCASR.

Heterozygous recombinant mice (*Zgrf1*$^{flox/+}$) were bred to homozygosity (*Zgrf1*$^{flox/flox}$) and crossed with β-actin-cre mice for ubiquitous deletion of the floxed region in *Zgrf1*. To extract DNA for genotyping, the mouse tail was clipped 0.5cm and transferred to a 1.5ml tube. The cut tailpiece was further chopped into smaller segments and dipped in 180µl of 50mM NaOH, followed by incubation at 90°C with constant rotation at 1200rpm for 15 minutes. The samples were allowed to cool down, and 20µl of 1M Tris pH 8.0 was added to the tubes and centrifuged at 13000rpm for 15 minutes at room temperature. The supernatant consisting of genomic DNA was collected and used for genotyping by PCR at different annealing and cycling conditions as per Cyagen protocol (A.2.12).

4.3. Results

4.3.1. Whole genome sequencing analysis

In the WGS study, we obtained an average read depth coverage of 37.59X and for the 4q24-q28 critical region, coverage was 39.6X (Table 4.1). The critical region spanning from nucleotide position chr4:102848900 to chr4:126353705 had 36,392 variants identified. Further, heterozygous variants were selected since the phenotype in the family segregated in an autosomal dominant manner. These variants were filtered out based on various parameters, such as removal of duplicate reads, variants in predicted transcripts, and exclusion of common heterozygous variants with MAF>0.01. The filtered rare variants were segregated into two categories; variants present in protein-coding genes, and variants in the non-protein coding/ regulatory regions. The variants

in protein-coding genes comprised a total of 411 variants with MAF< 0.01, of which two were exonic variants, ten UTR variants, and one intronic variant within 100bp of intron-exon boundaries (Figure 4.2). These 13 variants were validated by Sanger sequencing in the patient IV:4 DNA and checked for their co-segregation with the clinical phenotype in the family. Among these, four segregating variants were found, which included one non-synonymous variant c.1805C>T in *ZGRF1*, one 5'UTR variation c.-96C>T in *LEF1*, and two 3'UTR variants c.*2206G>A in *EGF* and c.*3748G>A in *FAM241A* (Table 4.2). The c.1805C>T *ZGRF1* variant was previously examined in the control set of 480 ethnically matched individuals, and its frequency was reported to be <0.01 (Shalini Roy Chaudhary, Ph.D. Thesis, 2018). The other three UTR variants were now examined in the same set of control individuals, and variants c.*2206G>A in *EGF* and c.*3748G>A in *FAM241A* were found to occur at low frequencies (Table 4.2).

4.3.2. Analysis of non-coding regulatory variants

The non-coding regulatory variants with MAF<0.01 were taken up for further analysis. The regulatory variants situated in the promoter region, promoter flanking region, CTCF binding site, enhancer region, open chromatin site, and TF binding sites were checked for conservation across species. The regulatory RNA variants present in long non-coding RNA, micro-RNA, small nucleolar RNA were further filtered by selecting variants in the non-coding exonic region and intronic region within 100bp. Out of 559 regulatory variants, I checked 128 variants for conservation status (Figure 4.4). I found 26 highly conserved variants, including 21 cis-regulatory variants and 5 non-coding RNA (Table 4.3). Among these, two were present in the CTCF binding site, three in the enhancer region, two in the open chromatin region, fourteen in the promotor region, and five in lncRNA. FATHMM-XF (Rogers et al., 2018) was used to predict the effects of these variants and found one variant g.111883909T>C in open chromatin region, two variants g.109302443G>A, g.113810007T>C in promoter regions, to be pathogenic (Table 4.3). These three variants g.111883909T>Cg.109302443G>A and g.113810007T>C were checked in genome alignment data and patient's exome data to find whether these are true variants. Out of the three, only one variant in open chromatin region g.111883909T>C was found to co-segregate with the clinical phenotype in the family.

Table 4.1. Whole genome sequencing coverage

Alignment attributes	Whole genome	4q24-q28
Total number of reads	813,696,920	
% Reads aligned to complete genome	99.95%	
Number of mapped reads	813,251,032	6,224,507
Target sequence length	3.2GB	24MB
% of target covered with atleast 10X read depth	92.04%	99.23
% of target covered with atleast 20X read depth	87.18%	98.46
% of target covered with atleast 30X read depth	76.60%	89.63
Average read depth	37.59 X	39.60 X

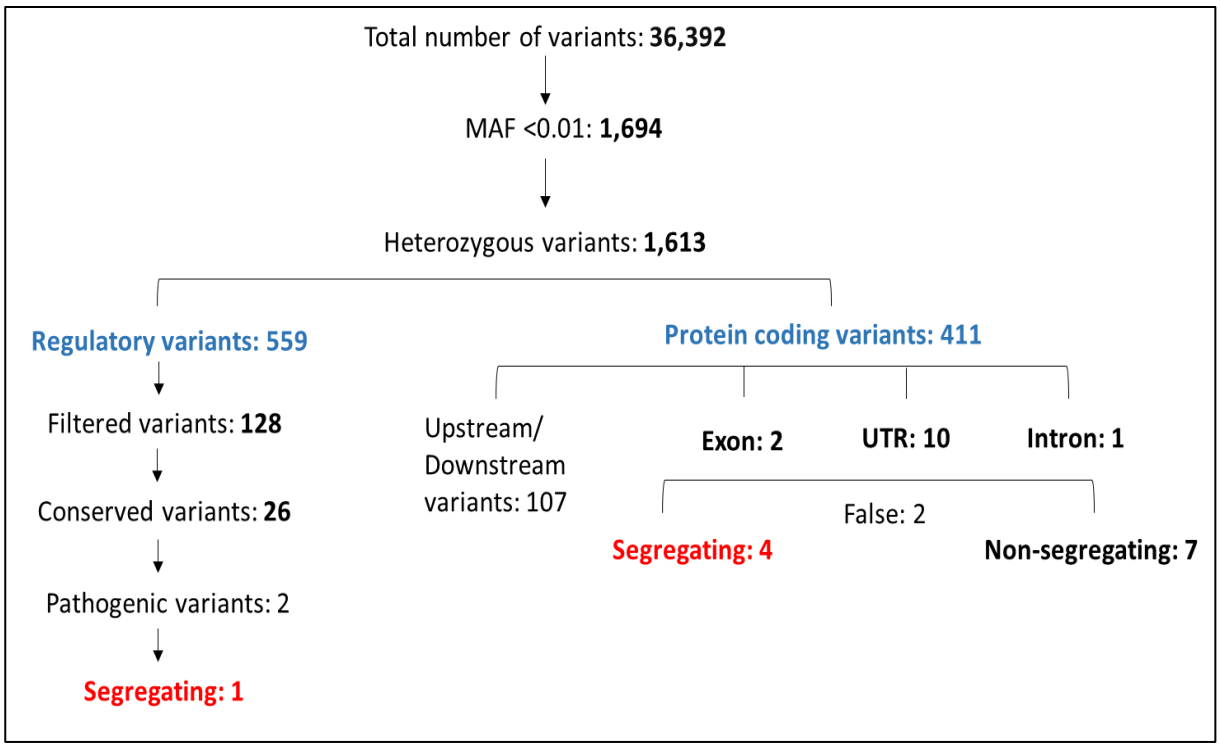


Figure 4.2. Analysis of variants in the 4q24-q28 region: The flow-chart represents the number of variants present at each step during the filtration process. The variants identified were first filtered based on their MAF across databases and their zygosity. Variants are further filtered based on their annotation and their segregation with the disease phenotype in the family.

Table 4.2. Thirteen heterozygous variants identified from WGS analysis with MAF <0.01.

Nucleotide position	Gene	Reference	Amino acid change	Biotype of SNV	True Variant	Family segregation	rs ID	MAF	MAF Inhouse controls
110542868	ENPEP	c.1925C>T	p.Ala642Val	missense_variant	Yes	No	rs778486245	0.000007	-
112618237	ZGRF1	c.1805C>T	p.Thr602Ile	missense_variant	Yes	Yes	rs201904239	0.0008	0.001
105847330	GSTCD	c.*1753G>A	-	3_prime_UTR_variant	Yes	No	rs539474003	0.00001	-
106045657	TBCK	c.*908_*912del	-	3_prime_UTR_variant	No	-	NR	NR	-
107951643	CYP2U1	c.*1220T>G	-	3_prime_UTR_variant	Yes	No	rs562528248	0.00003	-
110013661	EGF	c.*2206G>A	-	3_prime_UTR_variant	Yes	Yes	rs140612901	0.0008	0.002
112190686	FAM241A	c.*3748G>A	-	3_prime_UTR_variant	Yes	Yes	rs1225223011	0.00004	0
121875889	TRPC3	c.*3843_*3846del	-	3_prime_UTR_variant	No	-	rs1337914259	0.01364	-
106319231	AIMP1	c.-2246C>T	-	5_prime_UTR_variant	Yes	No	rs538266361	0.00003	-
106321096	AIMP1	c.-381C>T	-	5_prime_UTR_variant	Yes	No	rs563086421	0.00003	-
108167863	LEF1	c.-96C>T	-	5_prime_UTR_variant	Yes	Yes	rs577527784	0.00005	0.008
109302443	COL25A1	c.-331C>T	-	5_prime_UTR_variant	Yes	No	NR	NR	-
109302443	COL25A1	c.-57+54C>T	-	intron_variant	Yes	No	NR	NR	-

Minor allele frequency (MAF) is cumulative of data in dbSNP 224, Ensembl, TOPMED, ExAC, gnomAD v2.1.1 and v3.1.1
NR: Not reported

Table 4.3: Heterozygous non-coding variants with MAF <0.01 identified in WGS analysis.

Nucleotide position	rs ID	Reference	Change	Symbol	Biotype of SNV	MAF	FATHMM Prediction	Family segregation
124556468	rs796226391	ACT	A	LINC02516	lncRNA	NR	neutral	-
108678401	rs185590924	G	A	AC107071.1	lncRNA	0.0002	benign	-
112443955	rs531616724	A	G	LOC107986304	lncRNA	0.0012	benign	-
117372929	rs186367911	A	C	LINC02263	lncRNA	0.0028	benign	-
123930396	rs755098554	TA	T	LINC01091	lncRNA	NR	neutral	-
110211769	-	G	A	-	CTCF_binding_site	NR	benign	-
110516766	-	T	C	-	CTCF_binding_site	NR	benign	-
106594589	-	A	C	-	enhancer	NR	benign	-
108272046	rs149186900	G	A	-	enhancer	0.0004	benign	-
111042086	rs540127779	C	A	-	enhancer	0.0006	benign	-
111363122	-	TGA	T	-	open_chromatin_region	NR	neutral	-
111883909	rs576757329	T	C	-	open_chromatin_region	0.0002	pathogenic	Yes
104494204	-	C	T	-	promoter	NR	benign	-
106264683	-	G	A	-	promoter_flanking_region	NR	benign	-
107058968	rs551656454	A	G	-	promoter_flanking_region	0.0014	benign	-
107648538	rs531841923	C	A	-	promoter_flanking_region	0.0002	benign	-
107658616	rs571737750	A	C	-	promoter_flanking_region	0.0002	benign	-
107951643	rs562528248	T	G	-	promoter_flanking_region	0.002	neutral	-
108649930	-	AT	A	-	promoter	NR	benign	-
109302443	-	G	A	-	promoter	NR	pathogenic	No
113641157	rs532994355	T	A	-	promoter_flanking_region	0.001	benign	-
120659587	rs574866977	T	A	-	promoter_flanking_region	0.0008	benign	-
121160660	rs563671905	G	C	-	promoter	0.0002	benign	-
122107721	rs542560257	C	A	-	promoter_flanking_region	0.0002	benign	-
110109561	-	T	C	-	promoter_flanking_region	NR	benign	-
113810007	rs556474361	T	C	-	promoter_flanking_region	0.0012	pathogenic	No

Minor allele frequency (MAF) is cumulative of data in dbSNP 224, Ensembl, TOPMED, ExAC, gnomAD v2.1.1 and v3.1.1, NR: Not reported

4.3.3. Segregating rare variant prioritization

Table 4.4. Insilico predictions for rare segregating variants

Gene	Reference	Amino acid change	Biotype of SNV	Insilico Prediction	Conservation Score	MicroRNA binding site
ZGRF1	c.1805C>T	p.Thr602Ile	Missense variant	benign	0.13	-
EGF	c.*2206G>A	-	3'UTR variant	benign	-0.35	Yes
FAM241A	c.*3748G>A	-	3'UTR variant	benign	-2.8	Yes
-	g. 111883909T>C	-	Open chromatin region	pathogenic	1.1	-

Conservation score using USCS genome browser: PhyloP score, microRNA-binding prediction using Target scan, Pathogenicity prediction using FATHMMXF, PolyPhen 2, and SIFT

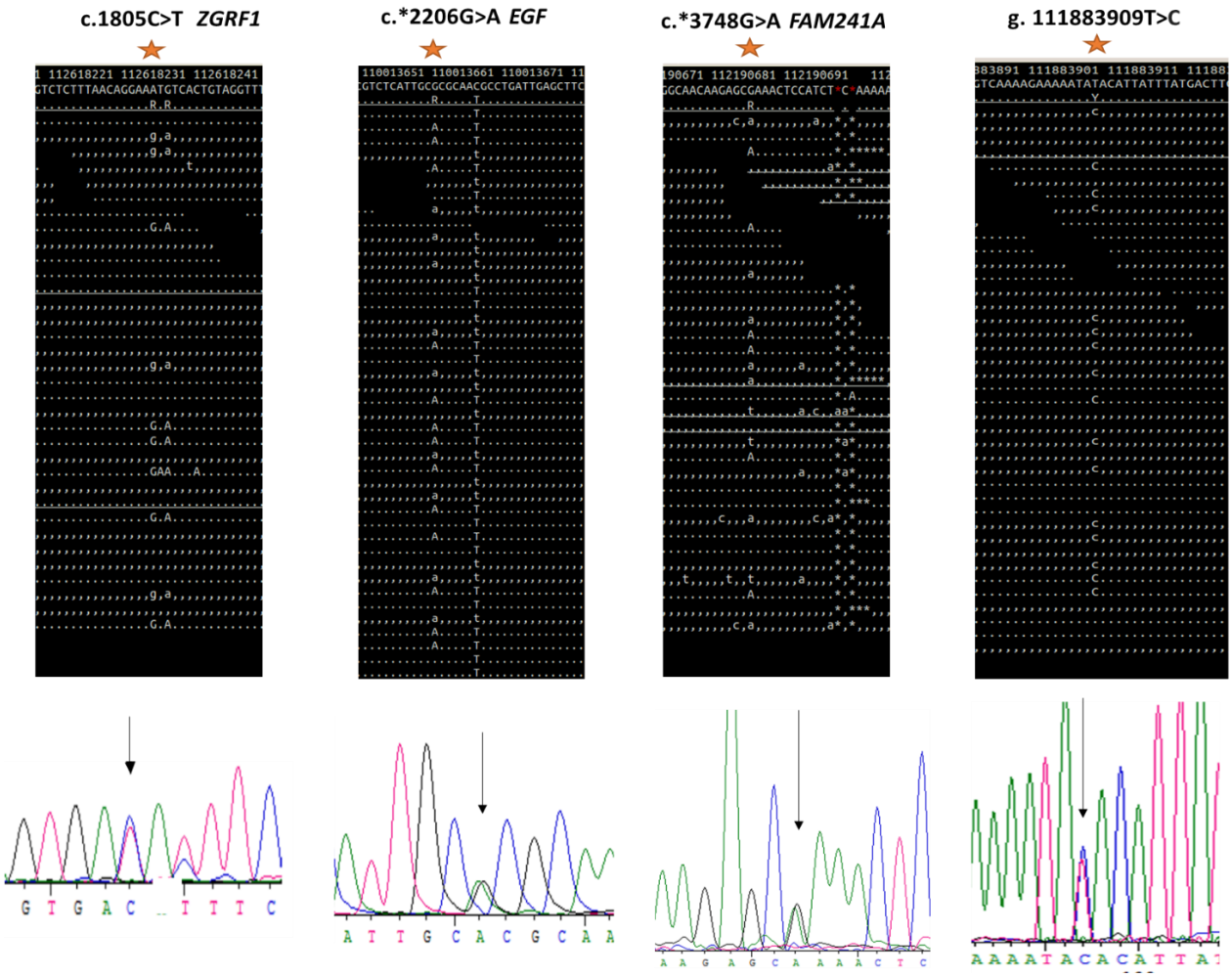


Figure 4.3. Low frequency variants segregating in HWE 227: SAMtools alignment snapshots of the variants: c.1805C>T in *ZGRF1*, c.*2206G>A in *EGF*, c.*3748G>A in FAM241A and g.111883909T>C in Individual IV:4. The variant is marked with an asterisk. The variants are present at high read depth and in both forward (dots) and reverse (commas) strands of the two individuals. Below each alignment is electropherogram representation of the variants in an affected individual of the family.

The insilico predictions suggested benign effects of the variants c.1805C>T *ZGRF1*, c.*2206G *EGF* and c.*3748G *FAM241A*, while g.111883909T>C in open chromatin region was predicted to be pathogenic. The c.1805C>T *ZGRF1* variant, which codes for change p.Thr602Ile, has amino acid threonine conserved in 28 out of 30 mammals examined using the UCSC Genome Browser. The two 3'UTR variants c.*2206G in *EGF* and c.*3748G in *FAM241A*, are conserved in 10 mammals, and in an analysis using PhyloP (http://compgen.cshl.edu/phast/) for 100 vertebrates, the conservation score obtained was negative (Table 4.4). Since 3'UTR are targets of mi-RNA binding, analysis for predicted mi-RNA binding sites was done using Target scan (http://www.targetscan.org/vert_72/) (Agarwal et al., 2015). The region spanning both the positions, c.*2206G in *EGF* and c.*3748G in *FAM241A*, belong to the category of the non-conserved target sites for microRNAs. The predicted micro-RNAs for these sites are miR-4783-5p and miR-6812-3p for *EGF1* and *FAM241A*, respectively.

A reporter-based luciferase assay was carried out to test whether the 3'UTR variants in the genes *EGF* and *FAM241A* cause disruption of mi-RNA regulation at the predicted site. The luciferase gene was fused to short 3'UTRs of wildtype and variant EGF and FAM241A and expressed in HEK293T cells along with β-galactosidase as an internal control for transfection efficiency. The functional relevance of 3'UTR variants, c.*2206G>A in *EGF* and c.*3748G>A in *FAM241A*, was evaluated for their effect on luciferase gene expression. Both the variants did not show any significant change in the reporter luciferase activity as compared to their wildtype allele (Figure 4.5). These results indicate that both 3'UTR variants might not have any functional consequence on the expression of the genes.

The variant, g.111883909T>C is conserved but there are no protein coding functional genes within its flanking 200kb. To study the pathogenic effect of such regulatory variants, multiple experimental approaches need to be taken to first understand the location of the regulatory region through chromatin association and molecular cytogenetic techniques and examine their possible effects on changes in gene expression (Rojano et al., 2019). Chromatin association methods like 3C, 4C, 5C, Hi-C, can help determine the physical connections between different loci, and the effect of mutations on these interactions can be studied. In this study, we prioritize the variant in a protein-coding gene, c.1805C>T *ZGRF1*, since mutations in *ZGRF1* have been previously

reported for other neurological disorders like apraxia of speech (Peter et al., 2016) and schizophrenia (Need et al., 2012).

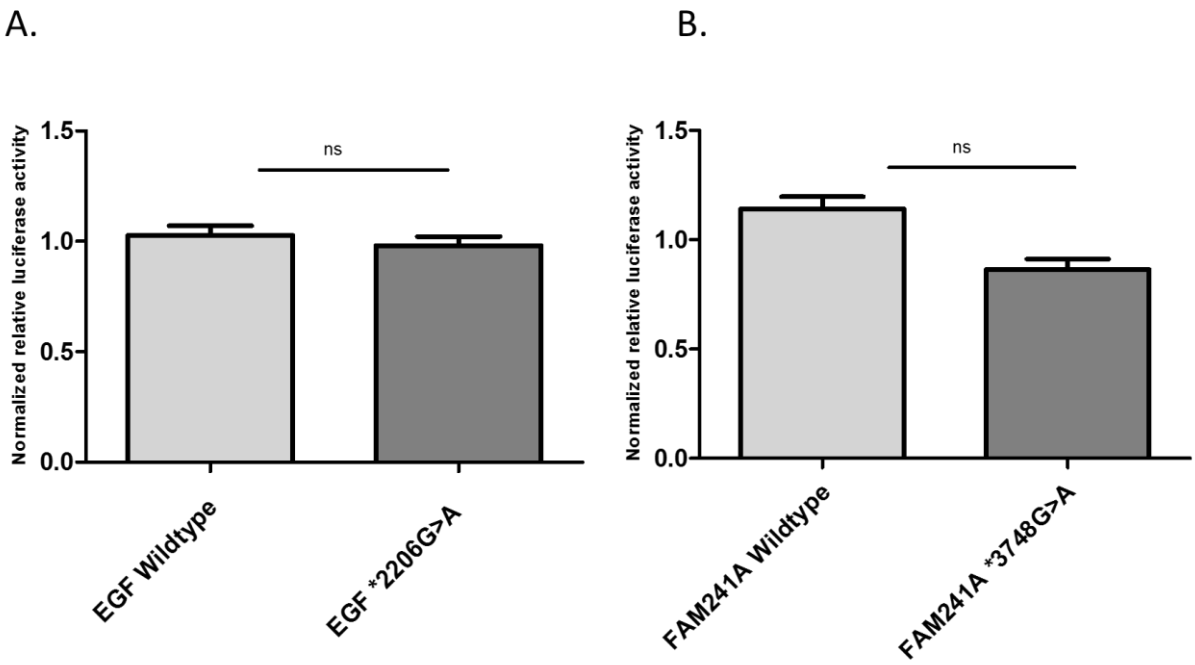


Figure 4.5. Transcriptional regulatory activity of the 3' UTR variants, c.*2206G>A in *EGF* and c.*3748G>A in *FAM241A*: (A) Represents the luciferase activity with the 421 bp of the wildtype 3' UTR, *EGF* compared to the variant, (B) Represents the luciferase activity with the 397 bp of 3' UTR, *FAM241A* compared to the variant. Relative luciferase activity between what was normalized with average activity of what and then to β-galactosidase expression levels and represented as normalized relative luciferase activity (y-axis). Data is represented as mean ± SEM for three independent experiments. Statistical analysis was done using unpaired Student's *t*-test.

4.3.4. *ZGRF1*, a candidate gene for hot water epilepsy

ZGRF1 (zinc finger GRF-type containing 1) codes for a transcript of 6652 nucleotides and protein of 2104 amino acids. The protein consists of five domains: DUF2439, zf-GRF, DNA 2, AAA11, and AAA12 domain (Figure 4.6). DUF2349 (Domain of unknown function) domain is present at the N-terminal, and proteins in this family have been implicated in telomere maintenance in Saccharomyces cerevisiae (Silva et al., 2016) and in meiotic chromosome segregation in Schizosaccharomyces pombe (Raschle et al., 2004; Maria et al., 2013). The zf-GRF (Zinc finger-GRF type) is a zinc-binding domain consisting of three conserved amino acids, glycine, arginine, and phenylalanine and is known to be involved in nucleic-acid binding. The DNA2 domain belongs to family DNA/RNA helicases, and protein consisting of this domain exhibit DNA-

dependent ATPase, ATP-dependent nuclease activity and helicase activity (Kim et al., 2006; Masuda-Sasa et al., 2006). The AAA11 and AAA12 (ATPases Associated with diverse cellular Activities) domains present at the C-terminus are P-loop NTPase domains essential for transferring macromolecules in an energy-dependent process (Iyer et al., 2004; Snider et al., 2008). ZGRF1 is a C-terminal paralogue of UPF1 protein, thus belongs to the class of SF1-UPF1-like RNA helicases. The other proteins in this class of helicase include UPF1 (Up-frameshift 1, RNA helicase and ATPase), AQR (Aquarius intron-binding spliceosome factor), DNA2 (DNA replication helicase/nuclease2), HELZ (Helicase with zinc finger), IGHMBP2 (Immunoglobulin mu binding protein 2), MOV10 (Mov10 RISC complex RNA helicase) and MOV10L1 (Mov10 RISC complex RNA helicase like 1). The domain architecture of ZGRF1 indicates that it has DNA-binding and helicase activity which were validated recently by Brannvoll and colleagues (Brannvoll et al., 2020). In subsequent sections, I present results on ZGRF1's role in cell division pathway.

4.3.5. Localization of ZGRF1 in mammalian cells

Previous cell-localization experiments of ZGRF1 in HEK293 cells were replicated. We observed that ZGRF1 localizes to the centrosome in the prophase stage and the spindle pole and mid-body in a mitotic phase (Figure 4.7A). On modifying the immunostaining protocol from methanol fixation to PFA fixation, I noted protein’s localization to the nucleus also (Figure 4.7B), an observation made in other ZGRF1 studies as well (Brannvoll et al., 2020; Yan et al., 2021). In the rat hippocampal neurons, ZGRF1 localization was observed in the nucleus and somato-dendritic compartments (Figure 4.7C).

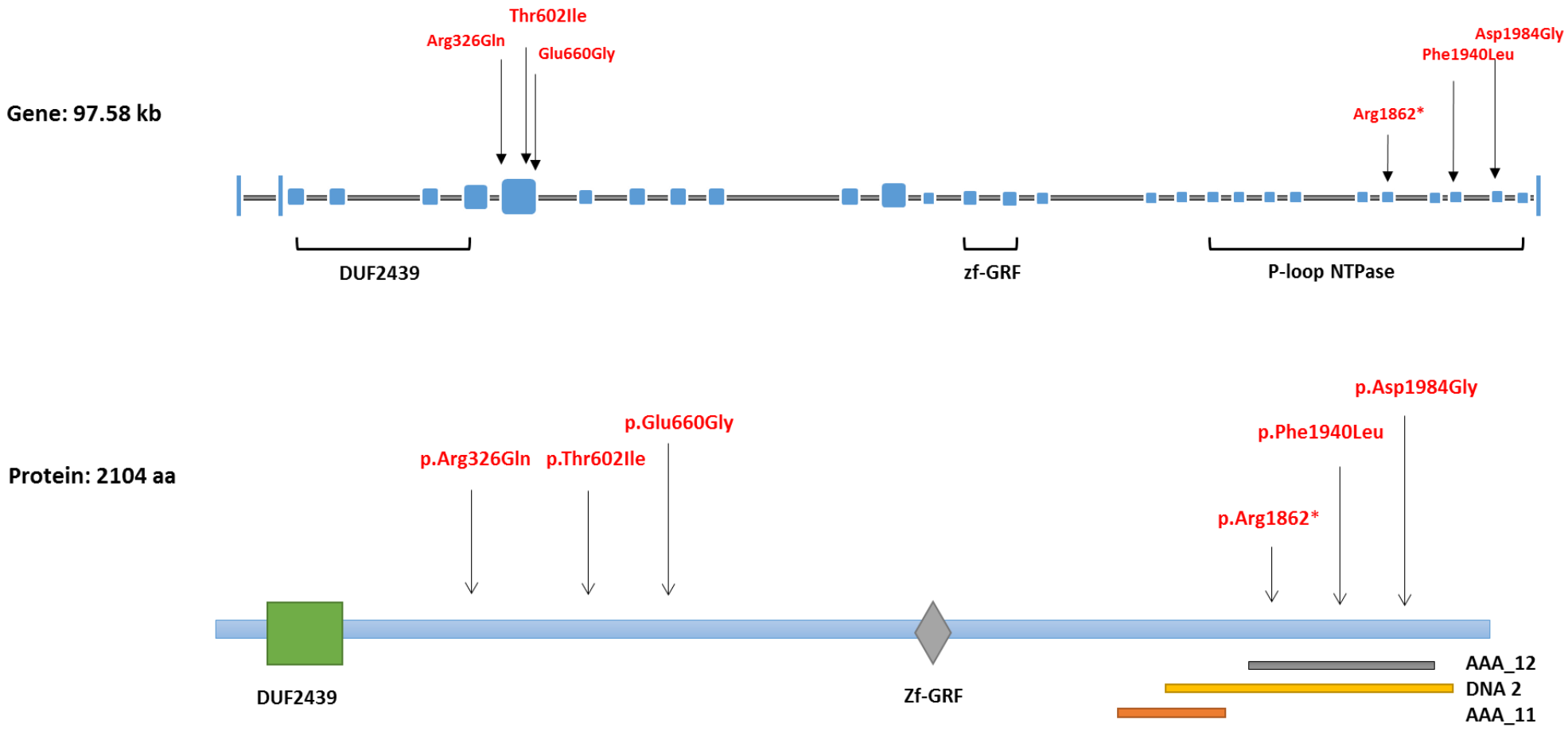
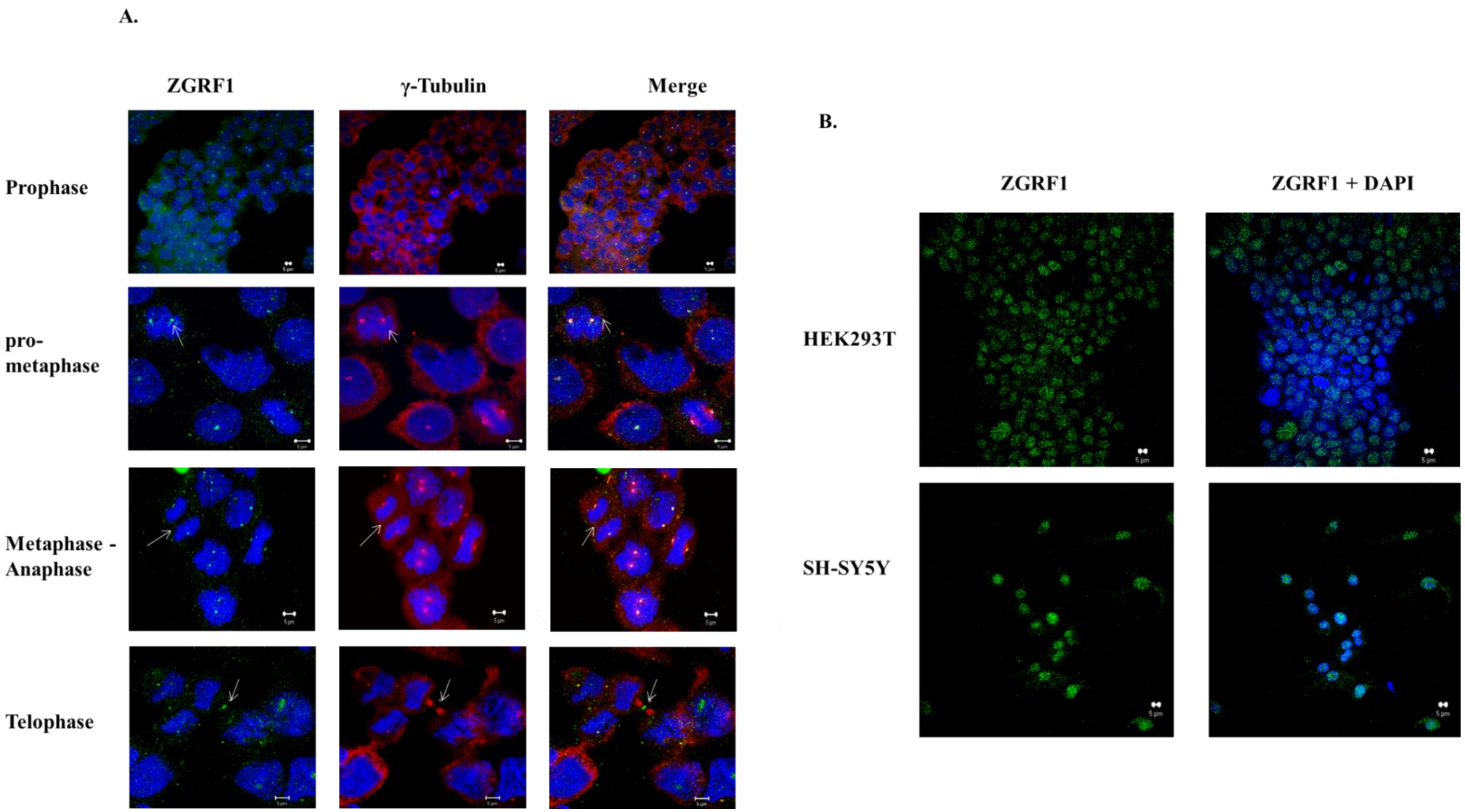


Figure 4.6. Potential functional variants in *ZGRF1*: Depicted gene structure for *ZGRF1* containing 28 exons. Below is the ZGRF1 protein and its domains. Highlighted are the position of patient variants found among the additional HWE patients examined.


A.
ZGRF1
γ-Tubulin
Merge
Prophase
pro-
metaphase
Metaphase -
Anaphase
Telophase
B.
ZGRF1
ZGRF1 + DAPI
HEK293T
SH-SY5Y

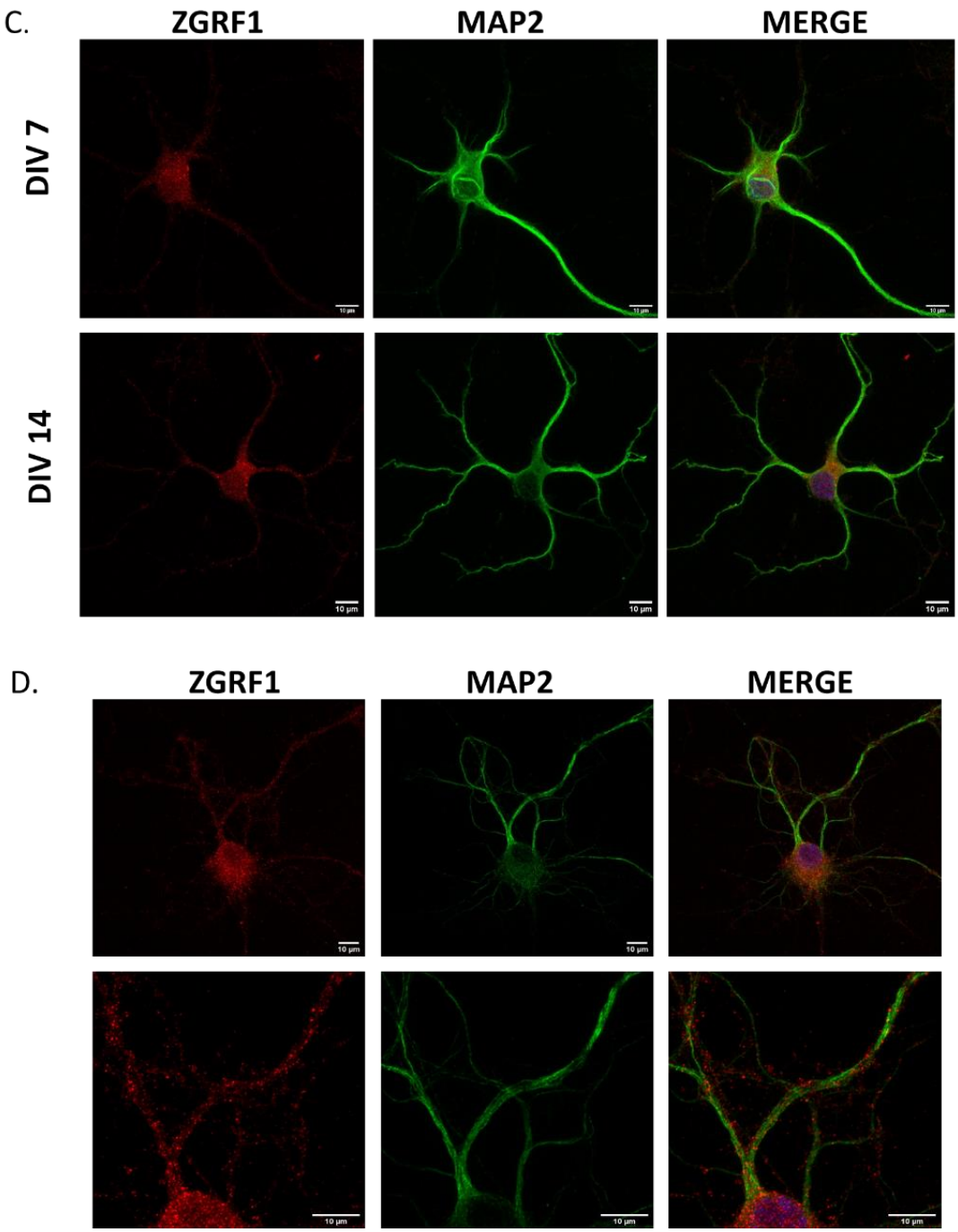
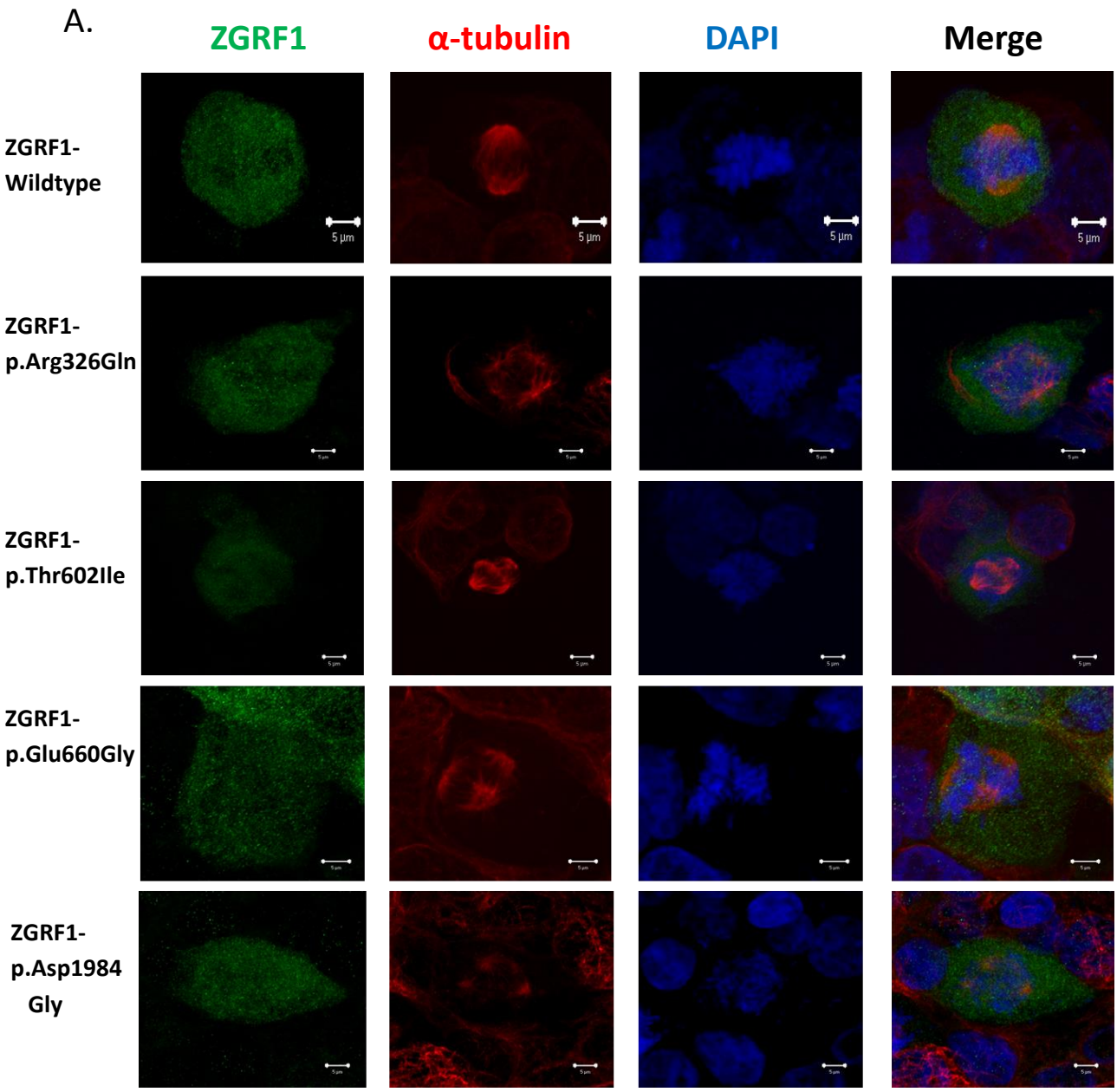


Figure 4.7. Expression of ZGRF1 in mammalian cells: (A) Immunofluorescence image showing ZGRF1 localization on centrosome, spindle poles, and midbody across different stages of the cell cycle. Green indicates the ZGRF1 protein stained using an anti-ZGRF1 antibody. Red indicates centrosomal marker γ-tubulin and blue is DAPI staining. Arrow shows cells at respective mitotic stages. (B) Immunofluorescence image for depicting nuclear localization of ZGRF1. (C, D) Rat primary hippocampal neurons at DIV-7 and -14 stained with anti-ZGRF1 antibody (red) and neuronal marker MAP2 (green) depicting its somato-dendritic localization; the second panel consists of a zoomed-in image of a dendritic shaft of the neuron.

4.3.6. *ZGRF1* variants exhibit mitotic defects

Upon over-expression of ZGRF1 in HEK293 cells, the localization of wildtype and variant ZGRF1 is mostly observed in the nucleus. Previous work has shown that ZGRF1 variants exhibit mitotic defects analyzed using γ-tubulin as a marker. I carried out this experiment for four variants: c.977G>A (p.Arg326Gln), c.1979A>G (p.Glu660Gly), c.1805C>T (p.Thr602Ile) and c.5951A>G (p.Asp1984Gly) using α-tubulin as a marker and found that mitotic cells over-expressing *ZGRF1* variants, p.Arg326Gln, p.Thr602Ile, and p.Asp1984Gly exhibit a significant increase in mitotic defects (Figure 4.8). These defects included abnormal spindle (monopolar and multipolar), chromosomal alignment defects: unaligned, lagging, bridge, and defective cytokinesis.



B.

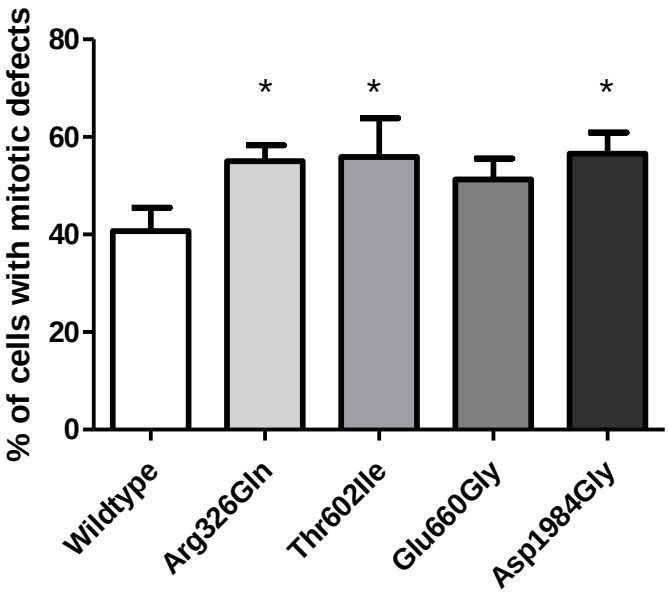


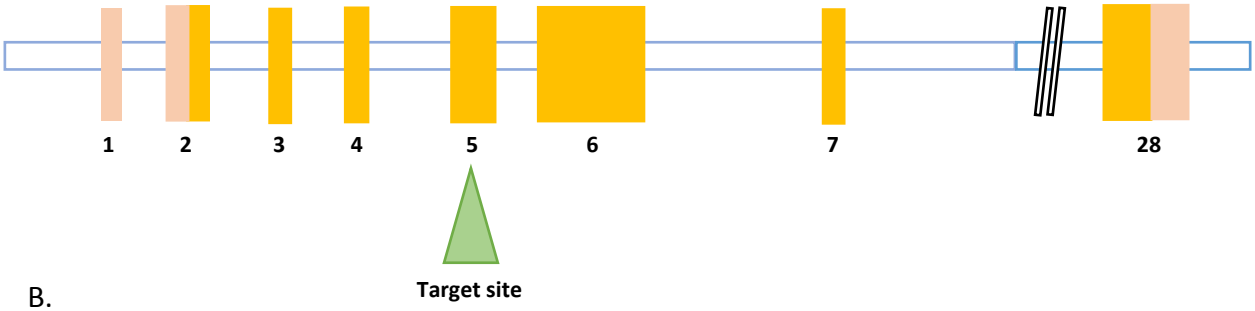
Figure 4.8. Mitotic defects in cells expressing ZGRF1 variants: (A) A representative image showing the mitotic defects at different cell cycle stages in cells with overexpressed ZGRF1 variants. Green indicates the overexpressed ZGRF1 protein stained using an anti-ZGRF1 antibody. Red is α-tubulin, used to analyze defects, and blue indicates DAPI staining, (B) Percentage of cells with mitotic defects in wildtype and the four ZGRF1 variants examined. n=40-60 cells. Data represented as mean ± SEM for three independent experiments. Statistical analysis was done using one-way ANOVA followed by Dunnet's multiple comparison test to compare wildtype and variant constructs * $p < 0.05$.

4.3.7. A *ZGRF1* deletion variant in mammalian cells

To further examine if ZGRF1 plays a role in mitosis, HEK293T cell lines carrying deletion alleles of the gene were generated. In the karyotype of these cells, four copies of chromosome 4 were observed (A.1.12). CRISPR-Cas9 nickase approach was taken for targeting the gene to avoid off-target effects. The fifth exon was chosen as the target (Figure 4.9A), and upon transfection and FACS sorting, 68 clones from ten 96 well plates were obtained. After genotyping and sequencing, I found 23 wildtype clones and 45 clones carrying insertion-deletions (INDELS). Among these 45 clones, certain clones had variants that would, upon translation, induce frameshift and a stop codon downstream in the reading frame. Among these mutant clones, two clones with

homozygous deletion ZGRF1-A$^{\Delta 26bp}$ and ZGRF1-B$^{\Delta 26bp}$ that predicted to generate a stop codon at the 126th codon were examined further (Figure 4.9B).

A.



B.

Cell line ID	INDEL at gDNA	Wildtype	Predicted consequence
ZGRF1-A$^{\Delta 26bp}$ (z52)	26bp del	Absent	Frame shift after 86 amino acids and stop codon after 126.
ZGRF1-B$^{\Delta 26bp}$ (z60)	26bp del	Absent	Frame shift after 86 amino acids and stop codon after 126.

Figure 4.9. CRISPR/Cas9 based targeting *ZGRF1* in HEK293T: (A) Schematic representation of the *ZGRF1* gene structure and target exon against which guide RNA was designed. (B) Homozygous mutant clones obtained and predicted impact of the mutation on the translation of *ZGRF1*.

Since the anti-ZGRF1 antibody gives an unclear outcome in immunoblot analysis, I examined *ZGRF1* transcripts in these cell lines. In the case of ZGRF1-A$^{\Delta 26bp}$ cell line, three transcript forms were present; one with the genomic deletion *ZGRF1*$^{\Delta 26bp}$ which makes a knockout transcript, and the other two with deletions, *ZGRF1*$^{\Delta 189bp}$ and *ZGRF1*$^{\Delta 249bp}$, that are deletions of exon 5, and exon 4-5, respectively. The later deletions arise due to an alternative splicing event (Figure 4.10B). *ZGRF1*$^{\Delta 189bp}$ and *ZGRF1*$^{\Delta 249bp}$ express amino acid deletion mutants ZGRF1$^{\Delta 19aa}$ and ZGRF1$^{\Delta 49aa}$ at the DUF2439 domain, respectively. The other ZGRF1-B$^{\Delta 26bp}$ cell line comprises two forms of transcripts: *ZGRF1*$^{\Delta 26bp}$ and *ZGRF1*$^{\Delta 189bp}$, later producing deletion of 19 amino acids (ZGRF1$^{\Delta 19aa}$) at the DUF2439 domain (Figure 4.10B).

A.

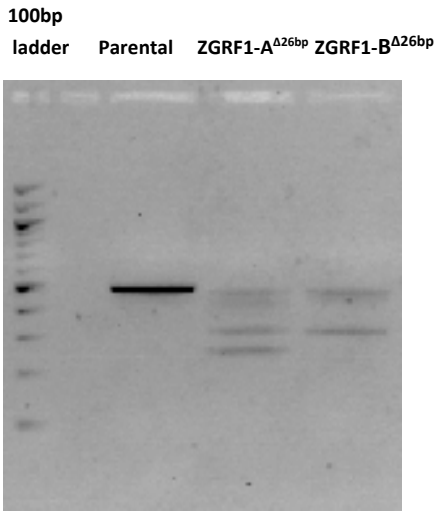
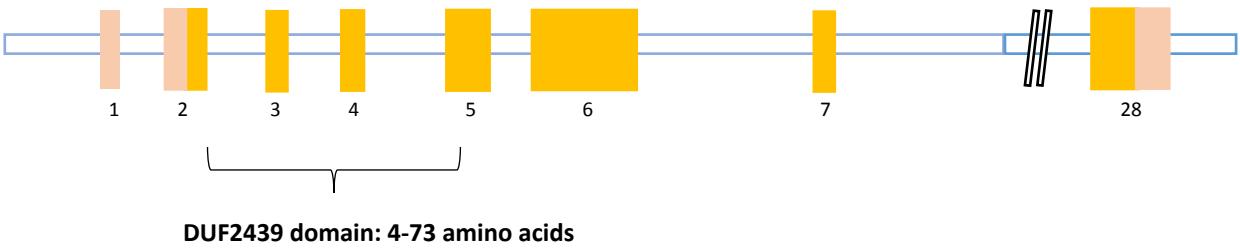


Figure 4.10. *ZGRF1* transcript analysis in the mutant clones: (A) *ZGRF*1-specific, cDNA amplification using primers spanning the exon 5 target site, for each mutant *ZGRF1* line created. The parental wildtype amplicon is 507bp. The shift in the size of the band in comparison with parental was seen in the mutant lines, and the presence of multiple bands indicating multiple transcripts being formed through alternate splicing; (B) Tabulated description of sequence-confirmed, multiple transcripts observed for each mutant line and predicted impact of the mutation on the translation of ZGRF1 protein. (C) Schematic representation of gene structure of ZGRF1 highlighting exons (numbered) coding for DUF2439 domain that is partially deleted in the mutant lines.

B.

Cell line	Mutation at mRNA deletion(del)	Consequence at mRNA level	Predicted consequence at protein level	Effect on DUF2439 domain
ZGRF1-A$^{\Delta 26bp}$	26bp del	26bp del in Ex5	Frame shift after 86 amino acid and stop codon after 126 amino acid	
	189bp del	Ex5 del	63 amino acid del	19 amino acid del
	249bp del	Ex4 and 5 del	83 amino acid del	49 amino acid del
ZGRF1-B$^{\Delta 26bp}$	26bp del	26bp del in Ex5	Frame shift after 86 amino acid and stop codon after 126 amino acid	
	189bp del	Ex 5 del	63 amino acid del	19 amino acid del

C.



4.3.8. Mutant *ZGRF1* lines exhibit mitotic defects and G2/M phase arrest

ZGRF1 mutant lines, ZGRF1-A$^{\Delta 26bp}$ and ZGRF1-B$^{\Delta 26bp}$ were checked for ZGRF1 expression pattern at centrosome and spindle poles (Figure 4.11). The localization observed was similar to that of the parental wildtype line, indicating that deletion of the DUF2439 domain may not be essential for the recruitment of ZGRF1 to the centrosome. Furthermore, I examined *ZGRF1* mutant cell lines for mitotic defects and observed that the mutant lines show a significant increase in the percentage of cells with mitotic defects compared to the parental HEK293T cell line (Figure 4.12A). ZGRF1-A$^{\Delta 26bp}$ and ZGRF1-B$^{\Delta 26bp}$ show 29% and 22% of cells with mitotic defects, respectively. These defects included abnormal spindles (monopolar and multipolar), chromosomal alignment defects: unaligned, lagging, bridge, and defective cytokinesis. A recent report by Brannvoll et al. 2020 reported that the ZGRF1 knockout line displays G2/M phase arrest. On cell cycle analysis of ZGRF1-A$^{\Delta 26bp}$ and ZGRF1-B$^{\Delta 26bp}$, a similar phenotype was observed, wherein there is a significant increase in accumulation of cells in the G2/M phase compared to the parental line (Figure 4.12B).

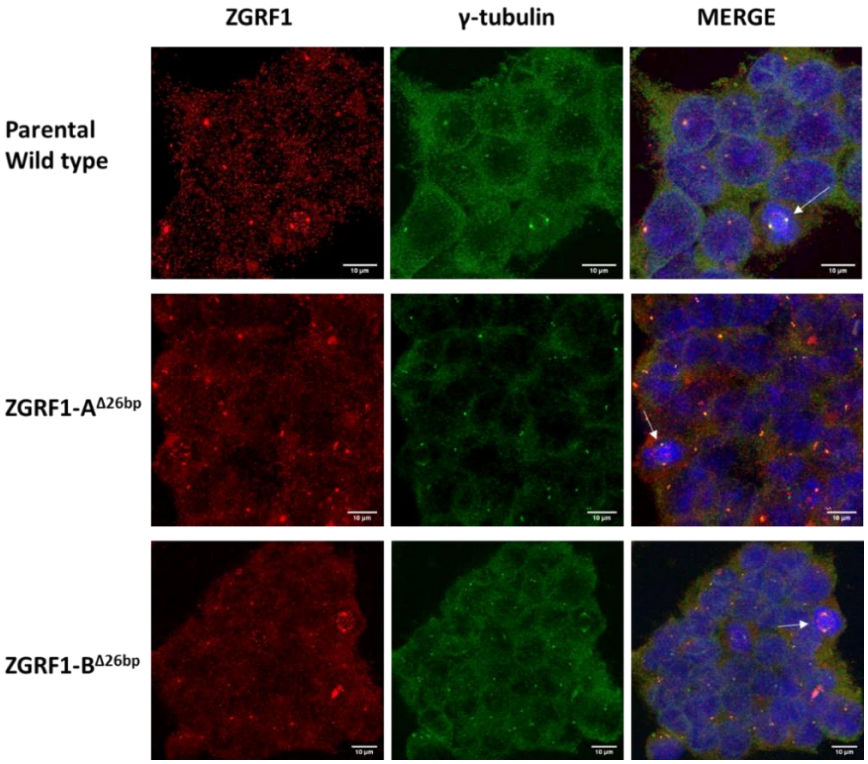


Figure 4.11. ZGRF1 expression: A representative image stained for ZGRF1 in parental wildtype HEK293T and mutant *ZGRF1* cell lines. Red indicates ZGRF1, green indicates the γ-tubulin protein, and blue indicates DAPI staining. The arrow shows cells at mitotic stages, highlighting ZGRF1 staining at spindle poles.

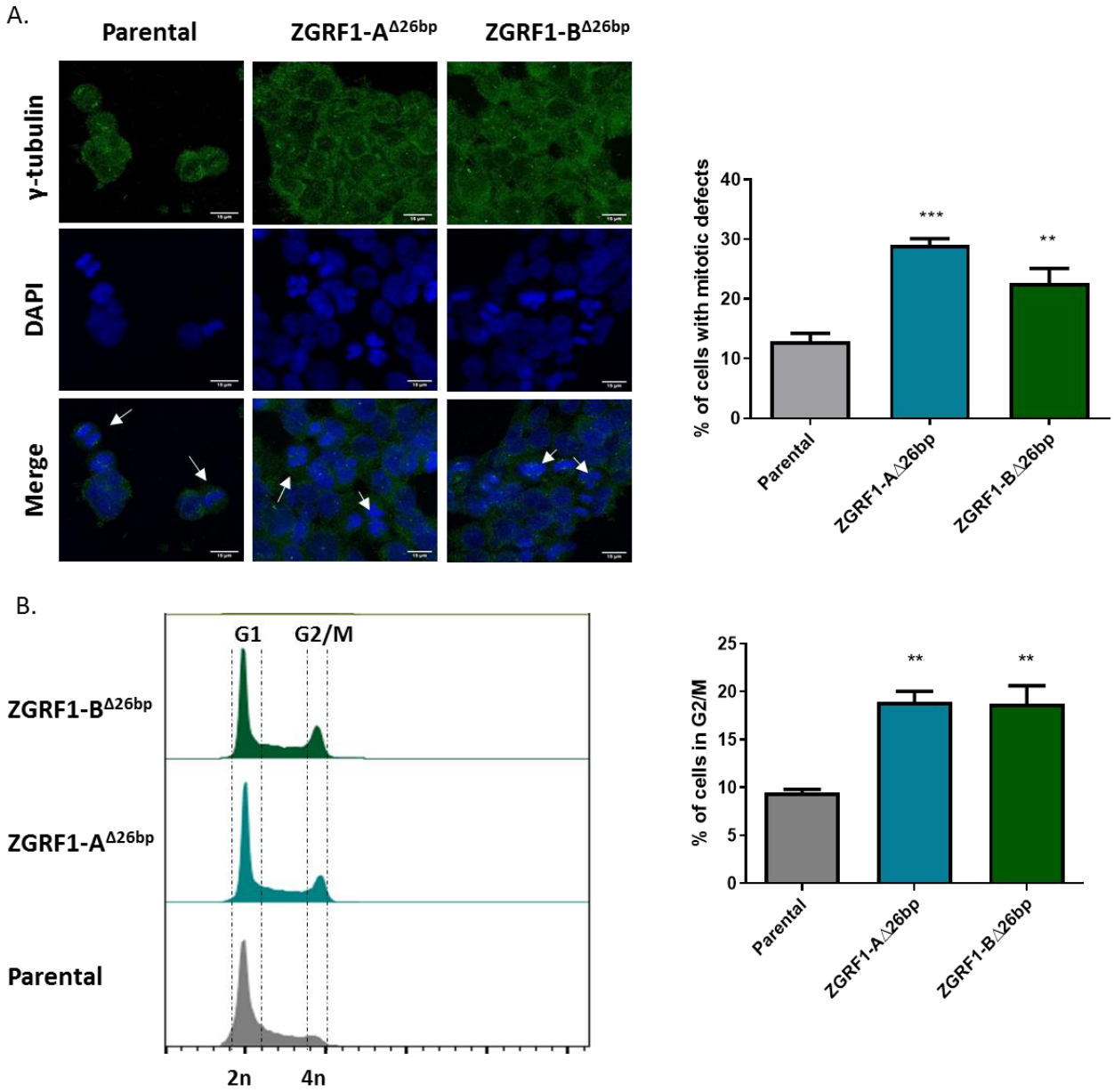


Figure 4.12. Mitotic defects in cells with *ZGRF1* mutants: (A) On the left is a representative image showing the mitotic defects in parental HEK293T and mutant *ZGRF1* cell lines. Green: γ-tubulin protein and blue: DAPI. The arrow shows cells with defective mitosis. On the right, plot for the percentage of cells with mitotic defects in parental and the two *ZGRF1* mutants cell lines for four independent experiments. n=200 cells, ***p<0.001; (B) Representative cell cycle profiles of PI-stained cells for parental and mutant *ZGRF1* cells and quantification of G2/M phase accumulation in parental and mutant *ZGRF1* lines. Percentage of cells in G2/M phase are plotted from three independent experiments. 2n depicts cells with diploid and 4n depicts cells with tetraploidy. Statistical analysis is done using one-way ANOVA followed by Dunnet's multiple comparison test. Data represented as mean ± SEM, ** p < 0.001

4.3.9. *Zgrf1* knockout in mice

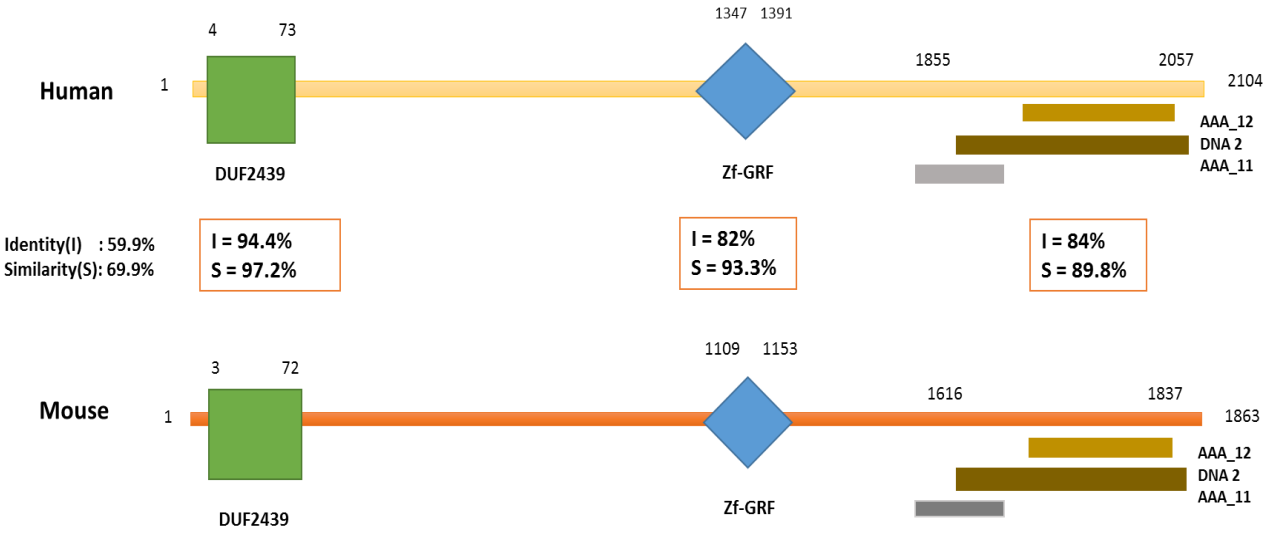


Figure 4.13. Protein conservation between human ZGRF1 and mouse Zgrf1: Schematic depicting full-length protein and domain similarity between ZGRF1 in two species. Conservation analysis is done using pairwise alignment tool in EMBL-EBI, *(https://www.ebi.ac.uk/Tools/psa/)*.

Zgrf1 gene is evolutionary conserved in mouse and is located at chromosome 3 (position 127,553,489-127,618,021), Ensembl ID ENSMUST00000043108.9. The transcript is 7046bp in length, contains 26 exons, and codes for a protein of 1863 amino acids. The mouse protein Zgrf1 shows a 69.9% similarity with human ZGRF1. The domains present in human ZGRF1 are also well conserved in the mouse: DUF2439 with 97.2% similarity, Zf-GRF with 93.3% similarity, and P-loop NTpase domain with 89.8% similarity (Figure 4.13). The transcript expression data available from Allen Brian Atlas for *Zgrf1* in mouse brain using *in-situ* hybridization predicts its expression in the isocortex, hippocampus, striatum, thalamus, midbrain, olfactory area, and cerebellum (https://mouse.brain-map.org/).

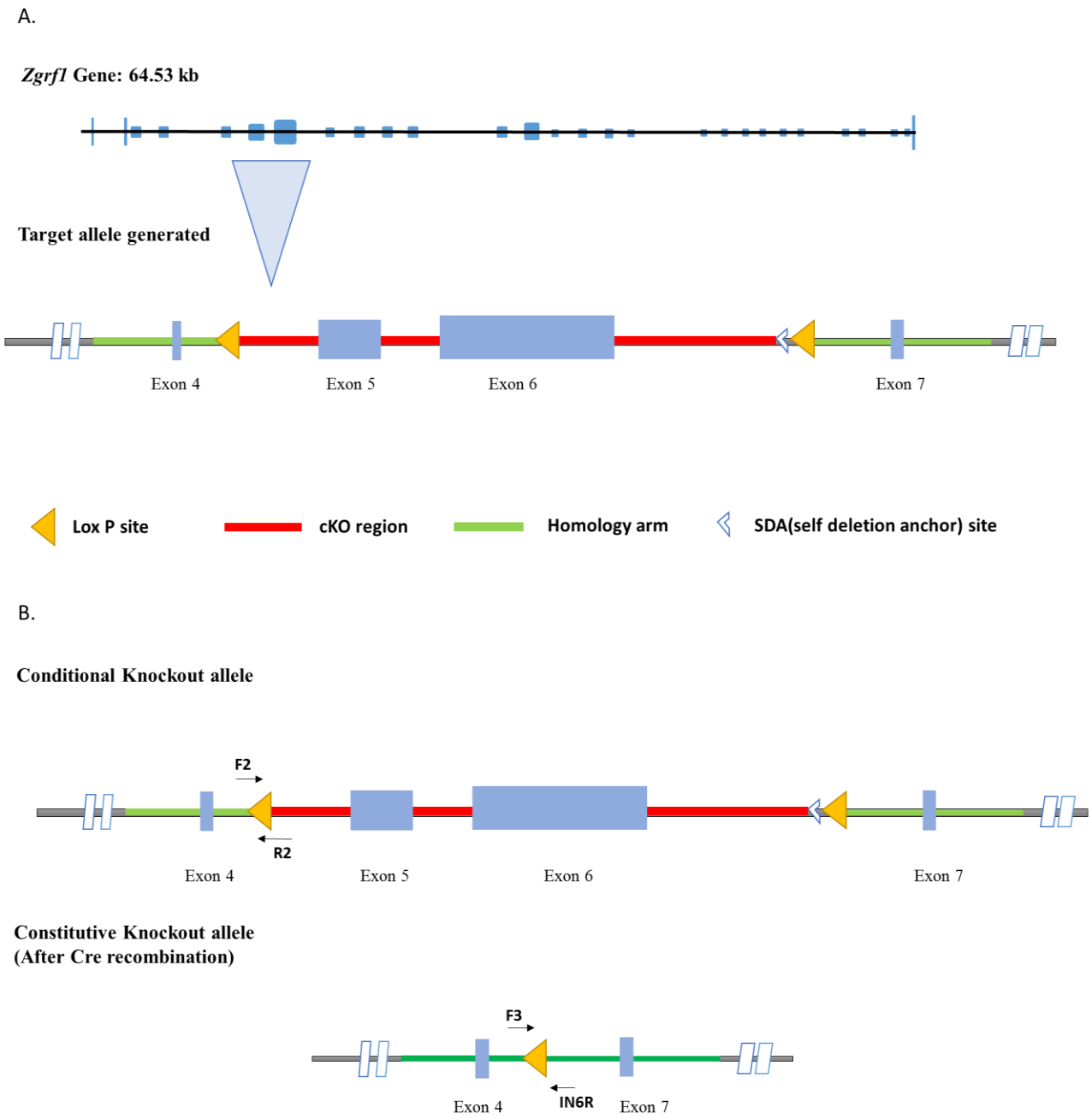


Figure 4.14. Strategy for conditional knockout of *Zgrf1* in M*us musculus*, C57BL/6NCrl: (A) Schematic showing gene architecture of mouse *Zgrf1* and target allele generated for cKO in mice consisting of LoxP sites flanking exon 5 (186 bp) and 6 (1744 bp) in the gene; (B) Representation of cKO allele and constitutive KO allele after cre- recombination and positioning of primers F2-R2, F3-IN6R to confirm *Zgrf1*^{floflox}and *Zgrf1*$^{-/-}$, respectively.

With an objective to explore biological function/s of *Zgrf1*, we undertook generation of its conditional allele in mice (sourced to Cyagen US Inc). The region containing exon 5 (186bp) and exon 6 (1744bp) were targeted to insert flanking *LoxP* sites as their deletion would lead to a frameshift in the coding transcript followed by generation of a stop codon at the 60th position. The targeted allele generated comprises conditional knockout (cKO) region of 5kb, SDA (self- deletion anchor) site, *loxP* sites flanking cKO region, and homology arms (Figure 4.14A). To confirm *Zgrf1* $^{flox/+}$ mice, PCR amplification was done using primer pair F2-R2 (A.2.12) that gives a band of size 251bp for the *floxed* allele and a band of 181bp for the *wildtype* allele (Figure 4.15A). Upon Cre- mediated recombination, to check the presence of the knockout (KO) allele, the primer pairs F2-R2, and F3-IN6R, along with Cre primers, were used. In the heterozygous KO allele, the primer pair F2-R2 gives a band at 181bp and F3-IN6R at 413bp. The heterozygous *Zgrf1* KO allele mice were intercrossed; for *Zgrf1*$^{-/-}$ homozygous mouse, a band of 413bp with F3-IN6R and no band with F2-R2 is expected. From the offspring obtained, mice with *Zgrf1*$^{+/-}$ and *Zgrf1*$^{+/+}$ were seen at ratio 2:1. However, out of 66 progeny examined, no homozygous *Zgrf1*$^{-/-}$ allele was found, suggesting that there is a possibility that these mice are lethal at embryonic stages.

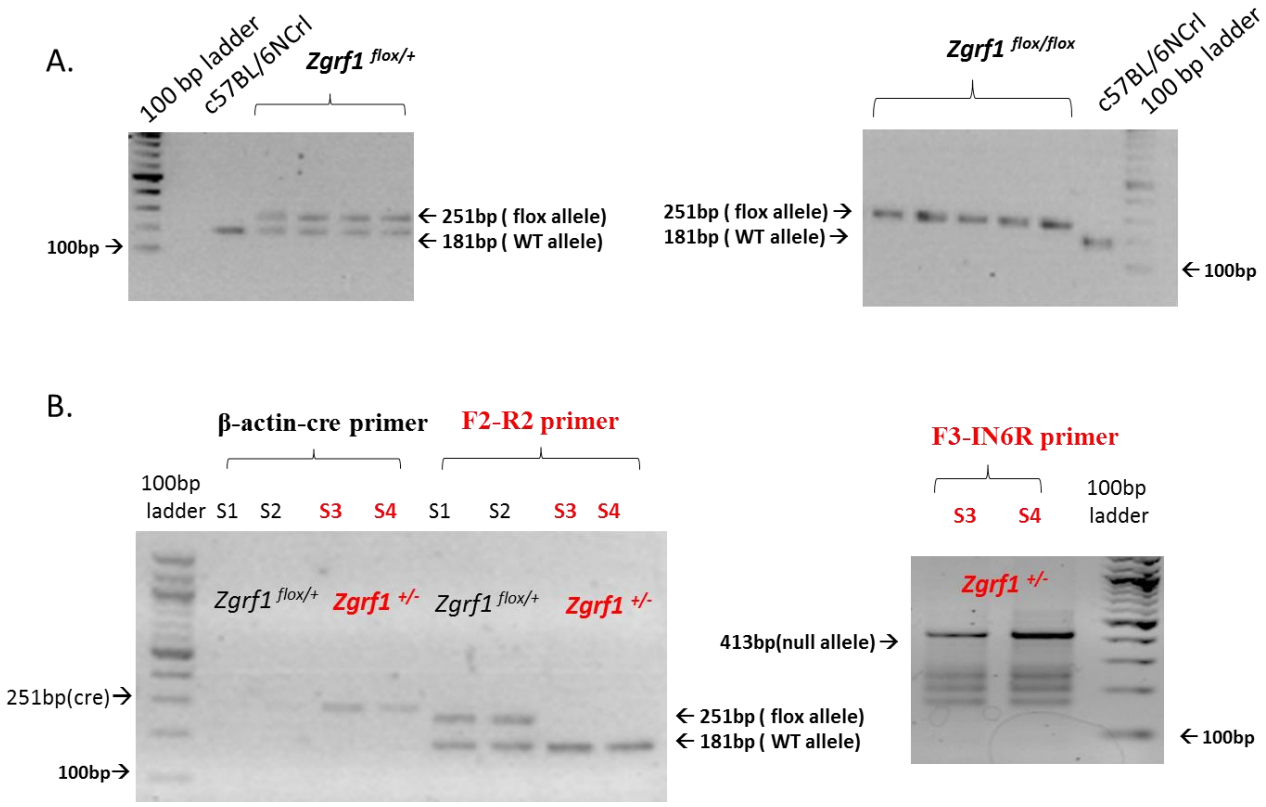


Figure 4.15. Confirmation of *Zgrf1* deletion allele: (A) Genotyping and PCR analysis to confirm the *Zgrf1*$^{flox/+}$and *Zgrf1*$^{flox/flox}$ using F2-R2 primers, size of *flox* band is 251bp; (B) Genotyping and PCR analysis for confirmation of KO allele in *Zgrf1*$^{+/-}$ and *Zgrf1*$^{-/-}$, using two primer pairs; F2-R2, F3-IN6R. In the KO allele, a 413bp band is obtained with F3-IN6R primers. All amplicons were confirmed by sequencing.

4.4. Discussion

Our studies involving a genome-wide linkage (Ratnapriya et al., 2009a) and whole exome sequencing analysis (Roy Choudhury et al., 2019) and whole genome sequencing (this study) in the family HWE227, has helped identify a plausible gene, *ZGRF1,* for this intriguing neurobehavioral disorder. In this chapter, I have reported four rare variants co-segregating with phenotype in the family, of which c.1805C>T in *ZGRF1* was apparently the most important one. Human ZGRF1 biological functions have begun to be explored recently. Until now, a few studies have reported *ZGRF1* mutations to be associated with neurological disorders. Two mutations in *ZGRF1*, rs76187047 (c.4087G>A, p.Glu1363Lys) and rs61745597 (c.142C>A, p.Leu48Met) has been reported in a multi-generation family with childhood apraxia of speech (Peter et al., 2016). In a large-scale exome study, *ZGRF1* variant, c.1504A>G (p.Met502Val), has been identified in three schizophrenic patients (Need et al., 2012). The protein has also been linked to alcohol dependency, psychological and personality disorders (Kalsi et al., 2010). These studies suggest that ZGRF1 plays a role in brain physiology. The cellular function of the protein was first suggested by Elledge and colleagues, wherein they conducted a genome-wide siRNA knockdown screen and implicated ZGRF1 in homologous repair and inter-strand crosslink repair mechanism (Smogorzewska et al., 2010; Adamson et al., 2012). Following these two reports, a relatively recent study by Brannvoll et al., 2020 has reported ZGRF1's role in the stimulation of homologous recombination during the repair of replication-blocking DNA lesions. They show that ZGRF1 has 5' to 3' helicase activity that catalyzes D-loop dissociation and Holliday junction branch migration. Also, ZGRF1 interacts with RAD51, which is a key component in the DNA double-strand break repair mechanism (Brannvoll et al., 2020). In addition, Yan and colleagues studied ZGRF1's role in homologous recombination repair via its association with BRCA1 and EXO1 (Yan et al., 2021). Human ZGRF1 consists of the DUF2439 domain, which shares homology with Mte1 (Mph1-associated telomere maintenance protein 1) in *Saccharomyces cerevisiae*, a protein important for

regulation of crossover recombination, response to replication stress, and telomere maintenance (Silva et al., 2016). In my study, I focused on understanding the role of ZGRF1 in mitotic progression based on a previous observation of ZGRF1's localization to centrosome and spindle poles (Roy Choudhury et al., 2019). Mitotic defects were observed when the ZGRF1 variants were over-expressed in cultured mammalian cells. Mitotic defects were also observed in cell lines carrying deletions of a large portion of the DUF2439 domain in ZGRF1. These lines exhibited G2/M phase arrest as well, also reported in *ZGRF1*$^{-/-}$ by Brannvoll and colleagues. Our observations emphasize the importance of the DUF2439 domain in ZGRF1 and its role in cell cycle progression. The interplay between homologous repair mechanism and mitotic progression has been reported in earlier studies. RAD51 paralogs RAD51B, RAD51C, RAD51D, XRCC2, and XRCC3 are key enzymes involved in double-stranded DNA repair (Masson et al., 2001; Sigurdsson et al., 2001). It has been shown that knockdown of RAD51B and RAD51C induces cell cycle arrest at the G2/M phase (Rodrigue et al., 2013) and cells with defective XRCC3 show increased centrosome numbers, aberrant mitotic spindles, and binucleated cells (Lindh et al., 2006). Another protein that plays an essential role in checkpoint activation and DNA repair is BRCA1 (reviewed in (Roy et al., 2012)). BRCA1 localizes to mitotic centrosomes and BRCA1-deficient cells display a high frequency of disorganized spindles and misaligned chromosomes in metaphase, delayed anaphase onset, relatively high levels of mis-segregated chromosomes in anaphase and chromosomal instability (Hsu and White, 1998; Stolz et al., 2010). Many other DNA damage response proteins have been reported in mitotic cell division processes (reviewed in Petsalaki and Zachos 2020). In view of these findings, we speculate dual function of ZGRF1 in DNA repair and mitotic progression. Though further studies are proposed to evaluate its roles in cell cycle progression by looking at different mitotic stages like centrosome duplication, spindle formation, mitotic checkpoints and cytokinesis.

DNA-damage repair pathways are important in terminally differentiated cells like neurons. The presence of single- and double-stranded breaks in the neuronal genome can have major consequences if not acted upon by repair proteins. Since neurons cannot enter the cell cycle, there is no homologous recombination repair, and non-homologous end joining is the primary mechanism involved in double stranded DNA damage repair. Repair for modified base or single stranded break is carried out by base excision repair and nucleotide excision repair pathways. Several reports show mutations in DNA repair

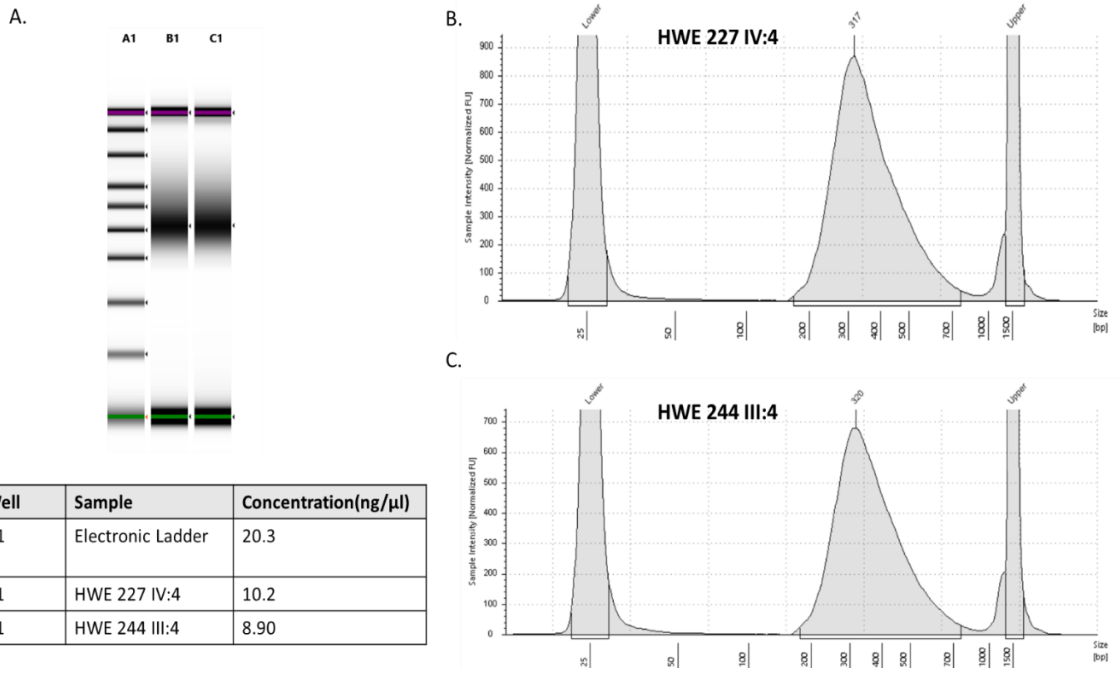
genes cause severe defects in the central nervous system (reviewed in Madabhushi et al., 2014). One such example is a mutation in ATM that leads to ataxia-telangiectasia (A-T), which displays progressive neurodegeneration along with movement and coordination defects, lack of natural eye movements, and slurred speech (Biton et al., 2008). Another example is a mutation in *Ligase IV (LIG4)* that causes LIG4 syndrome characterized by microcephaly, unusual facial features, growth, and development delay (O'Driscoll et al., 2001). Several other disorders with mutations in DNA repair proteins that exhibit neurological defects include Nijmegen Breakage Syndrome, A-T Like Disease, ATR-Seckel Syndrome, and XLF Syndrome (reviewed in (Madabhushi et al., 2014)). In mice, it has been shown that loss of ubiquitin ligase RNF8 leads to neurodegeneration and the presence of reactive astrocytes (Ouyang et al., 2015). Another recent study by Guo and colleagues, shows that RAD6B deficiency leads to increased genome instability and neurodegeneration in mice. These mice exhibit behavioral differences such as reduced learning and memory abilities (Guo et al., 2019).

In our attempt to understand the pathology associated with the knockout of *Zgrf1* in mice, we found embryonic lethality for *Zgrf1*$^{-/-}$ mice. Lethality at embryonic stages in mice is not uncommon among proteins involved in DNA damage repair and mitotic progression. ZGRF1 interactor RAD51's knockout mice fail to undergo early embryonic development (Tsuzuki et al., 1996). Also, disruption of RAD51's paralogue: RAD51B (Shu et al., 1999), RAD51C (Kuznetsov et al., 2009), RAD51D (Pittman and Schimenti, 2000) and XRCC2 (Deans et al., 2000) in mice results in embryonic lethality or early neonatal death. Another important interactor protein of ZGRF1 discussed earlier is BRCA1, whose ablation leads to lethality prior to day 7.5 of embryogenesis due to reduced cell proliferation (Hakem et al., 1996). Also, the absence of BRCA1 in the central nervous system leads to widespread neuroanatomical abnormalities, which includes severe proliferative defects and excessive apoptosis resulting in cellular loss in all laminated structures in the brain, including the neocortex, cerebellum, hippocampus, and olfactory bulb (Pao et al., 2014). It will be important to look at the development stages at which lethality takes place for *Zgrf1*$^{-/-}$. Furthermore, examining heterozygous knockout and tissue-specific knockout mice for neurological phenotypes that are known to be associated with the knockout of DNA damage repair proteins in mice is proposed. Additionally, in the light of HWE, examining whether these mice display sensitivity to hyperthermia or have a lower threshold for seizure induction using hyperthermic

kindling would validate it as a good model. Polynucleotide kinase/phosphatase (PNPK) is so far the only DNA repair protein associated with seizure phenotype (Dumitrache and McKinnon, 2017). Mutations in PNPK cause microcephaly with seizures beginning in infancy, where in addition to microcephaly and epilepsy, there is also developmental delay, intellectual disability, and hyperactivity (Shen et al., 2010). Mutations in proteins involved in cell division are also reported in epilepsy. One such example is EFHC1 (EF-hand domain-containing protein 1) which plays an important role in regulating cell division and neuronal migration during cortical development, and its functional defects lead to juvenile myoclonic epilepsy (Suzuki et al., 2004b; De Nijs et al., 2009; Raju et al., 2017). Another example is mutations in Doublecortin, a microtubule-associated protein, which results in severe epilepsy and mental retardation in humans (Gleeson et al., 1998; Kerjan and Gleeson, 2007). In future, it will be interesting to address questions regarding the role of *ZGRF1* in DNA repair in neurons and its role in mitotic events in neuronal progenitors and non-neuronal cells in the brain. *ZGRF1's* likely role in mitotic progression also opens avenues for studying its role in brain development with the emphasis on neuronal division, migration and maturation.

Appendix 1: Supplementary results

A.1.1. Library quality check: The library quality assessment was done using DNA 5000 Screen Tape in an Agilent 4150 Tape-Station system. 1µl of the library was mixed with 5µl of DNA sample buffer, vortexed for a minute, and spun down to collect the sample to the bottom of the strip. The strip was then loaded onto the Agilent 4150 Tape-Station instrument. The library qualification criteria was based on the presence of a broad peak in the range of 200 bp to 1000 bp, with an average size of 350 bp in the Tape-station system, and the Qubit concentrations were more than 2 ng/µl / 10 nMol. (A) Gel Electrophoresis Image from Agilent 4150 Tape-Station system (B, C) Represents range of sizes obtained after library preparation in HWE 227 IV:4 and HWE 244 III:4 sample, respectively.



Well	Sample	Concentration(ng/µl)
A1	Electronic Ladder	20.3
B1	HWE 227 IV:4	10.2
C1	HWE 244 III:4	8.90

A.1.2. Whole genome sequencing raw data

A. Whole genome sequencing raw data quality

Sample ID	No. of reads	Data in GB	GC%	Read length	%Q20	%Q30
HWE244 III:4	737659038	110.649	41	150	99.945	96.595
HWE227 IV:4	820968662	123.145	41	150	99.94	96.44

B. Sequence of Adapters

P7 adapter read1 AGATCGGAAGAGCACACGTCTGAACTCCAGTCA

P5 adapter read2 AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT

A.1.3. Variants found in *R3HCC1* in 480 ethnically matched control individuals

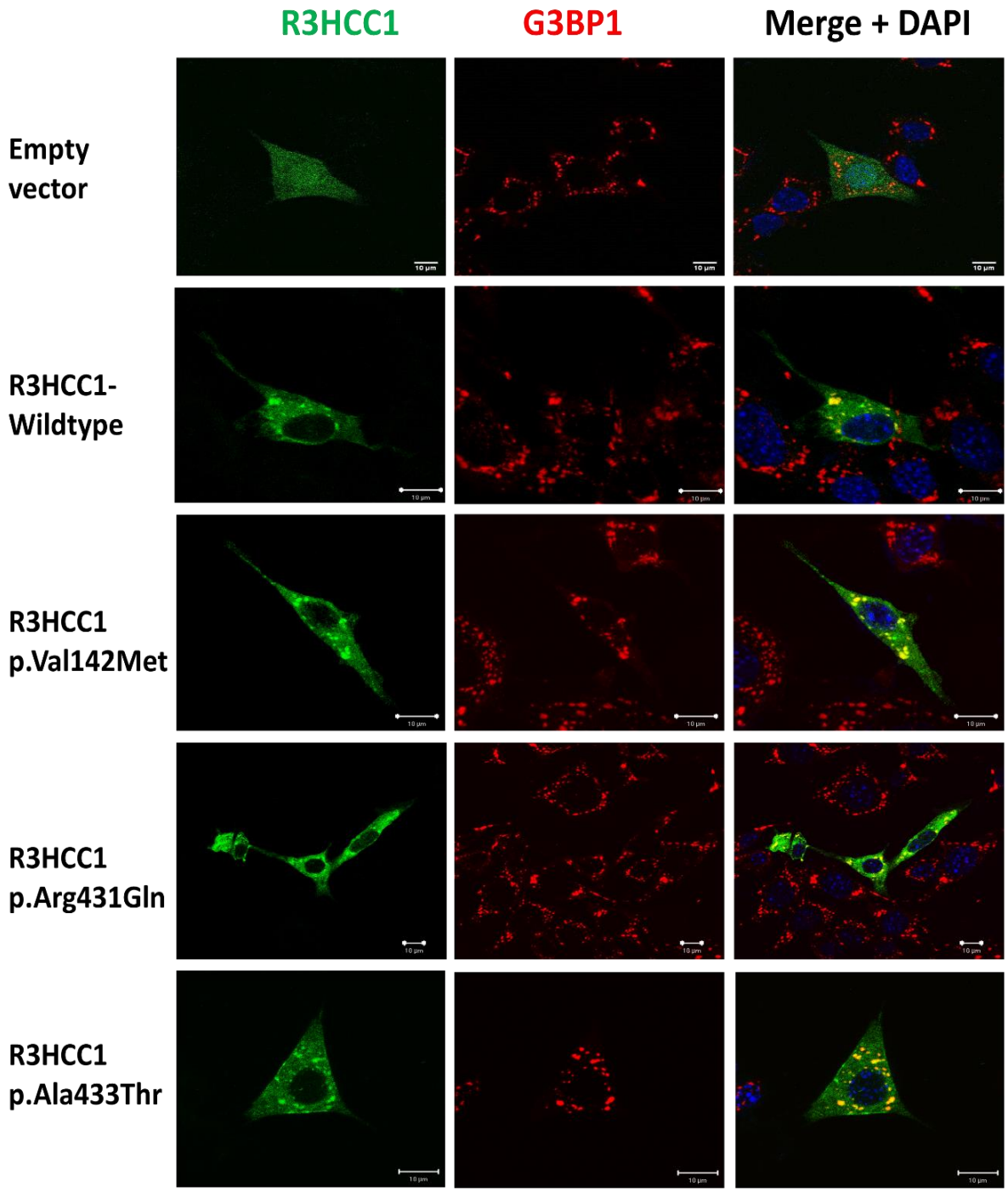
Exon	Nucleotide position	Reference	Het/Hom	Biotype of SNV	Amino acid change	rs ID	MAF	MAF in Inhouse Control
1	23287994	c.-182C>G	Het	5_prime_UTR_variant	-	NR	NR	0.002
1	23288157	c.-19G>C	Het	5_prime_UTR_variant	-	NR	NR	0.001
1	23288256	C>T	Het	intron_variant	-	NR	NR	0.001
2	23288507	c.-17C>G	Het	5_prime_UTR_variant	-	rs1349320161	0.00001315	0.001
2	23288548	c.25G>A	Het	missense_variant	p.V9I	NR	NR	0.001
2	23288582	c.59A>G	Het	missense_variant	p.H20R	rs11546682	0.19	0.1
4	23290051	c.434G>A	Hom/Het	missense_variant	p.R145K	rs3808536	0.53	0.53
4	23290182	c.565A>G	Het	missense_variant	p.T189A	rs761309582	0.000012	0.001
5	23291417	c.909A>G	Hom/Het	synonymous_variant	p.T303T	rs2272762	0.44	0.54
5	23291427	c.919G>A	Hom/Het	missense_variant	p.V307M	rs2272761	0.45	0.55
5	23291465	c.957G>T	Het	synonymous_variant	p.V319V	rs2272760	0.13	0.1
5	23291576	c.1025+43A>G	Hom/Het	intron_variant	-	rs2272759	0.44	0.44
6	23293265	c.1026-38_1026-35del	Hom/het	intron_variant	-	rs375807313	NR	0.41
6	23293364	c.1087C>T	Het	synonymous_variant	p.L363L	rs545985863	0.00003286	0.001
6	23293365	c.1088T>G	Hom/Het	missense_variant	p.L363R	rs13530	0.55	0.52
6	23293487	c.1096+114A>G	Hom/Het	intron	-	rs3860836	0.55	0.4
7	23294821	c.1149A>G	Het	synonymous_variant	p.T383T	rs15946	0.78	0.65
7	23294847	c.1175A>C	Hom	missense_variant	p.K392T	NR	NR	0.001
8	23296056	c.1282A>G	Het	missense_variant	p.K428E	rs574653396	0.00003943	0.001
8	23296091	c.1317G>A	Het	missense_variant	p.P439P	rs527941084	0.0001261	0.001
8	23296145	c.*48G>A	Het	3_prime_UTR_variant	-	rs144495060	0.01	0.01
8	23296169	c.*72A>G	Het	3_prime_UTR_variant	-	rs562679690	0.001072	0.001

Minor allele frequency (MAF) is cumulative of data in dbSNP 224, Ensembl, TOPMED, ExAC, gnomAD v2.1.1 and v3.1.1
NR: Not reported

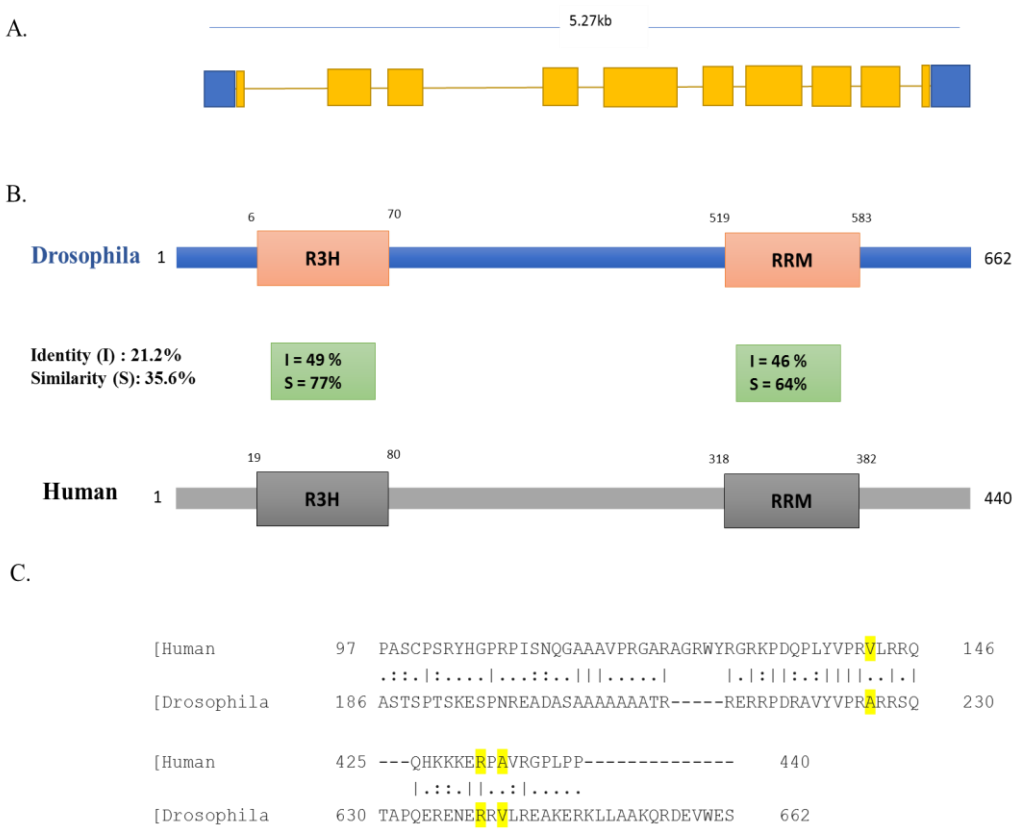
A.1.4. Predicted pathogenic variants exclusive in Bipolar disorder and Schizophrenia patients

Variant ID	Consequence	Biotype of SNV	Insilico Prediction
Bipolar: BiPex			
8-23289050-C-T	p.Arg49Trp	Missense	Damaging
8-23288503-C-A	c.-18-3C>A	Splice site	Damaging
8-23289102-C-G	p.Ser66Cys	Missense	Damaging
8-23289123-G-C	p.Arg73Thr	Missense	Damaging
8-23290002-C-T	p.Arg129Cys	Missense	Damaging
8-23290033-T-A	p.Val139Glu	Missense	Damaging
8-23290033-T-G	p.Val139Gly	Missense	Damaging
8-23291469-G-C	p.Glu321Gln	Missense	Damaging
8-23293343-G-A	p.Ala356Thr	Missense	Damaging
8-23294775-A-G	p.Glu368Gly	Missense	Damaging
8-23294810-C-T	p.Arg380Trp	Missense	Damaging
8-23294811-G-A	p.Arg380Gln	Missense	Damaging
8-23294855-C-G	p.Gln395Glu	Missense	Damaging
8-23289015-G-A	c.111-1G>A	Splice site	Damaging
8-23289903-G-GCTCCTGC	p.Pro101LeufsTer36	Stop gained	Damaging
8-23290438-C-A	p.Ser274Ter	Stop gained	Damaging
8-23293357-T-TCC	p.Cys362ProfsTer10	Stop gained	Damaging
Schizophrenia: SCHEMA			
8-23147825-C-T	p.Thr45Ile	Missense	Damaging
8-23147843-T-C	p.Leu51Pro	Missense	Damaging
8-23147908-G-C	p.Glu73Gln	Missense	Damaging
8-23148890-C-T	p.Thr103Met	Missense	Damaging
8-23152299-C-T	p.Arg185Trp	Missense	Damaging
8-23149048-T-C	c.464+2T>C	Splice site	Damaging
8-23150817-G-T	p.Glu156Ter	Stop gained	Damaging

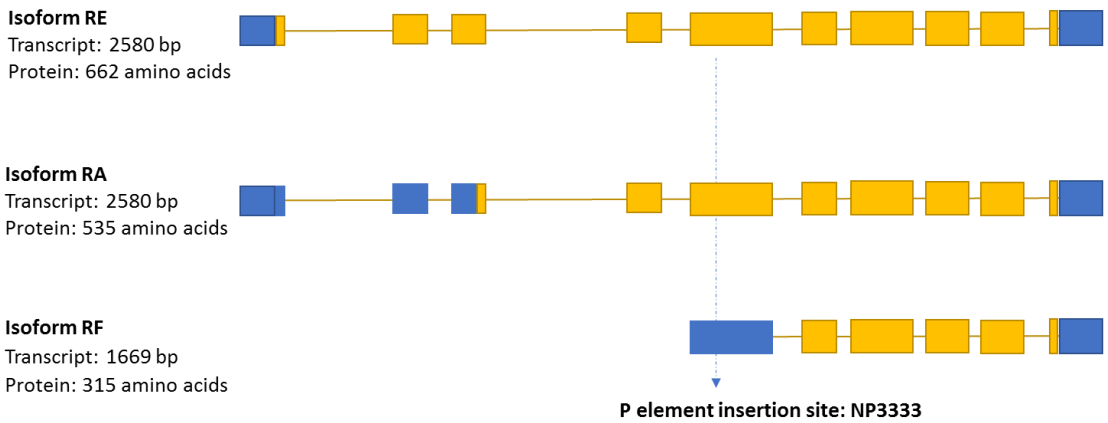
A.1.5. Immunostaining for R3HCC1 mutants in stress granules: EGFPC2 constructs for R3HCC1- wild type and variants, p.Val142Met, p.Arg431Gln, p.Ala433Thr were overexpressed in SH-SY5Y and subjected to 44°C for 1 hour following staining for endogenous G3BP1.



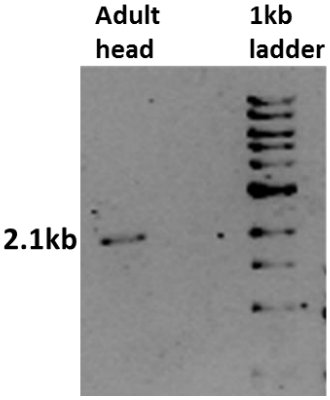
A.1.6. *Drosophila* ortholog of *R3HCC1:* (A) Gene structure of *CG2162* comprising of ten coding exons. Coding exons in yellow and UTR in blue, (B) Schematic depicting full-length protein and domain similarity between CG2162 and R3HCC1 and (C) Amino acid conservation sites in the protein are marked for HWE patient mutations found p.Val142, p.Arg431, p.Ala433 in Humans and *Drosophila*.



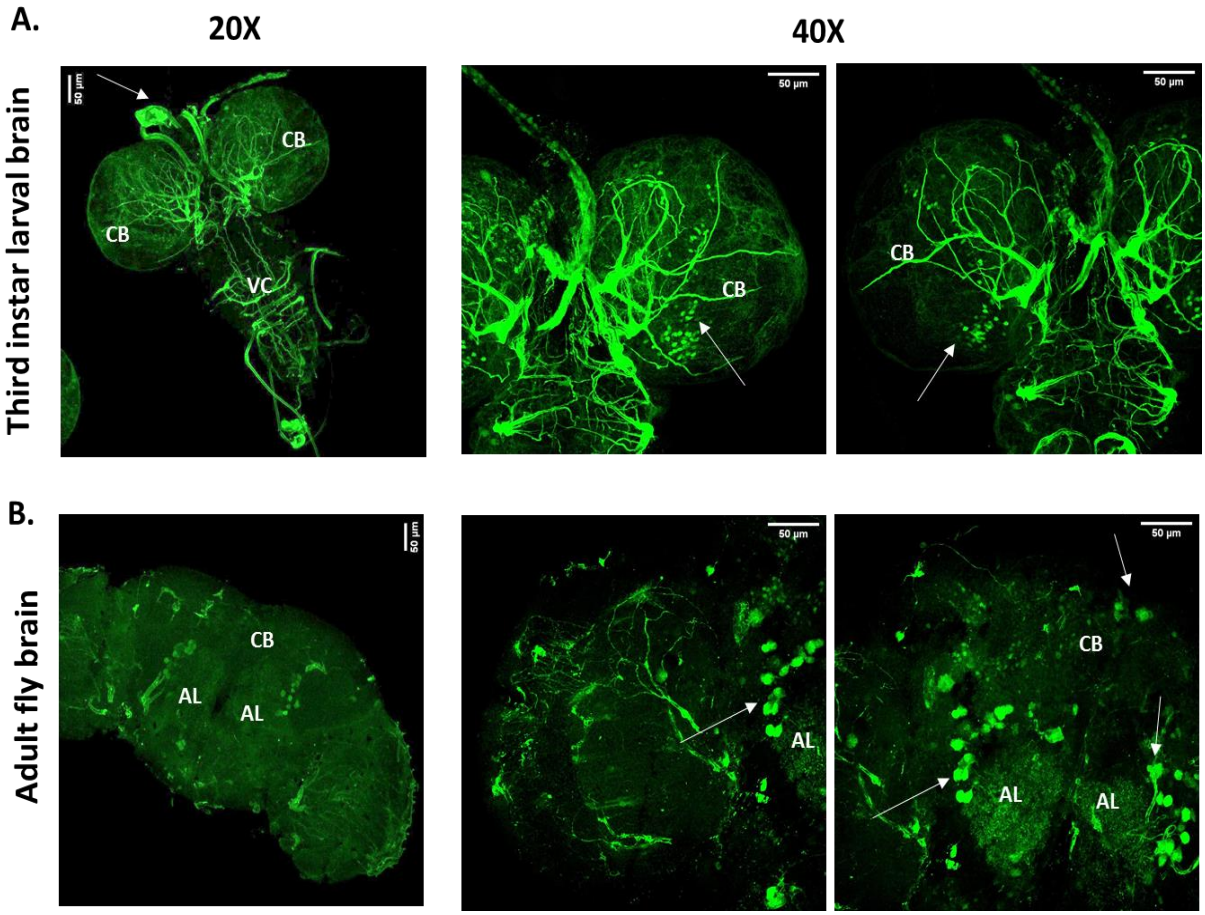
A.1.7. Isoforms of *CG2162*: Schematic depiction of three predicted gene isoforms of *CG2162*; isoform RE and RA comprising 10 exons, isoform RF comprises 6 exons. Blue represents UTR region and yellow represents coding exons. Arrow marks the site of P element insertion in gene *CG2162* in NP3333 line.



A.1.8. ***CG2162* transcript in adult *Drosophila* brain**: PCR-amplified band of size 2.1kb was present in adult wild type w^{1118}*Drosophila* head. The band was sequence confirmed.

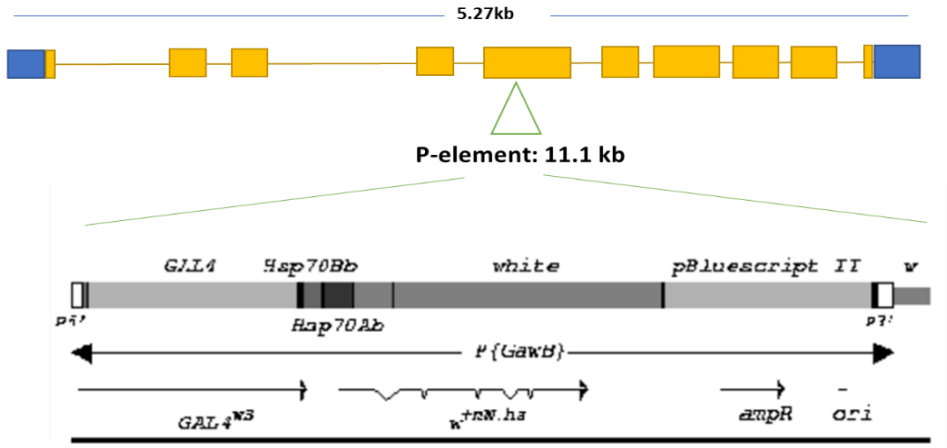


A.1.9. *CG2162* tissue specific expression: (A) NP3333 crossed with UAS-GFP; F1 generation *Drosophila* larval brain stained with anti-GFP, imaged at 20X and 40X. CB: Cerebral lobes, VC: Ventral cord, (B) F1 generation adult D*rosophila* brain stained with anti-GFP, imaged at 20X and 40X. CB: Central brain, AL: Antennal lobes, OL: Olfactory lobe.

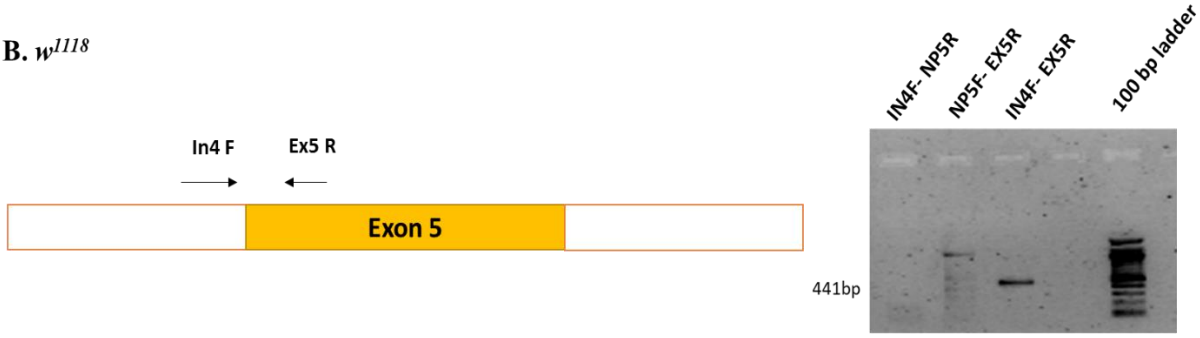


A.1.10. P element insertion in the gene *CG2162*: (A) Schematic depicting 11.1kb P element insertion in exon 5 of *CG2162*. P element comprises of w^{+}mW.hs mini-white visible marker and ScerGAL4 driver/enhancer trap sequences in the P element (source: http://flybase.org/reports/FBti0035338). (B) PCR amplification from the genomic DNA of control w^{1118} flies at 5' region of exon 5, which gives a band of 441bp. A non-specific band is observed for primer pair NP5F-EX5R. (C) Insertion of P element checked using two sets of primer pair amplifying start and end sequence of P element in NP3333 line at genomic DNA. Upon sequencing these PCR amplicons, an additional insertion of 8bp present immediately upstream of the P element was observed, which had not been reported in the database for the line.

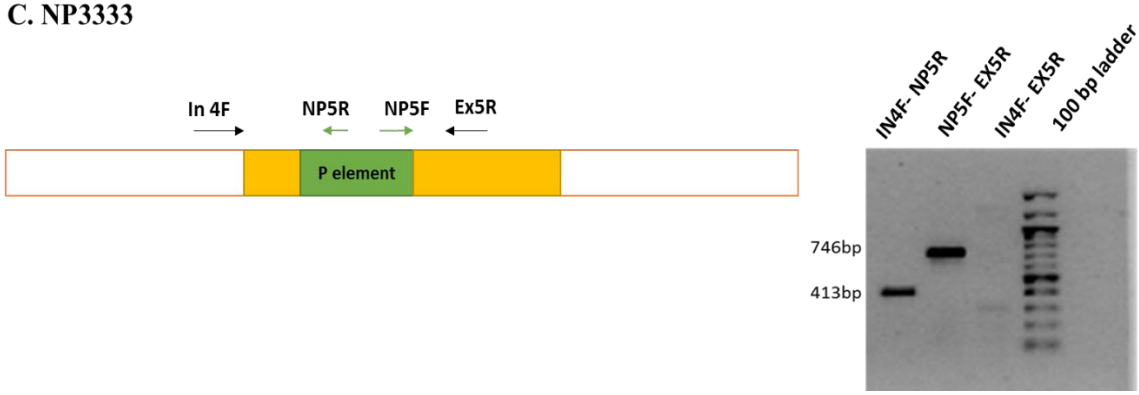
A.



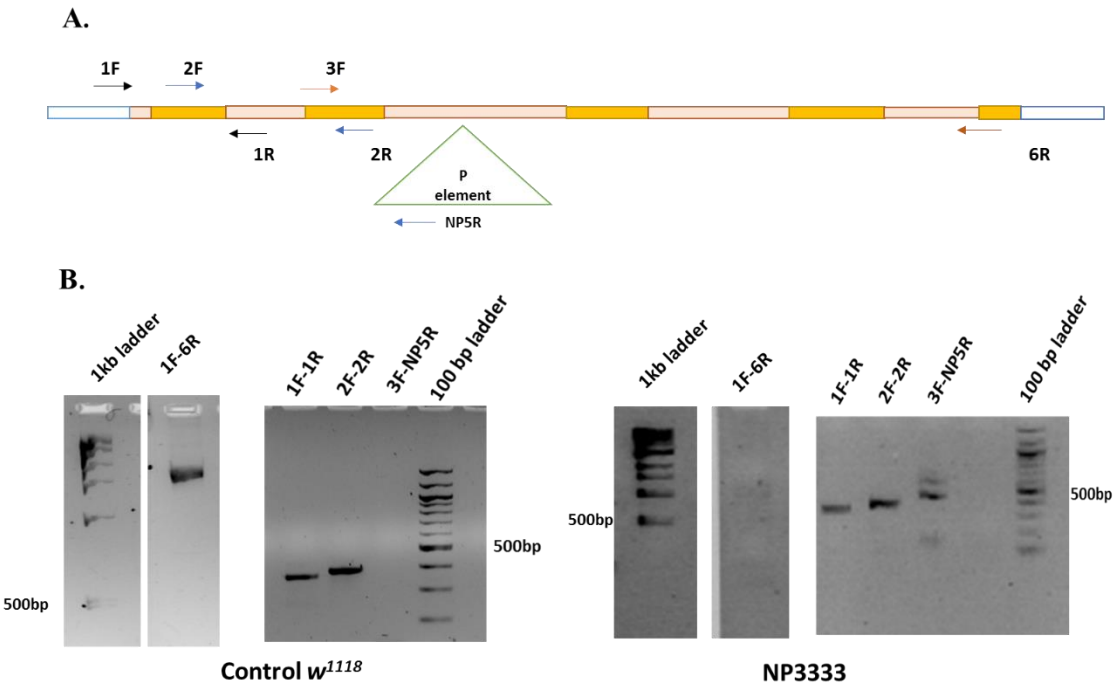
B. w^{1118}



C. NP3333



A.1.11. P element insertion in transcript *CG2162*: (A) Schematic depicting primers used across *CG2162* to check the presence of transcript in NP3333. (B) Right gel indicates the absence of a full-length transcript in NP3333 but the presence of 5' sequence of the transcript in cDNA from NP3333 head, including the P element sequence. The left gel represents amplification of control w^{1118} head cDNA using the same primer pairs.



A.1.12. Karyotype analysis of HEK293T cells: numbers represent chromosome number

Appendix 2: Reagents and Protocols

A.2.1. Primer pairs for screening of uncovered exons in HWE 244 region 8p23-p12 in whole exome sequencing analysis

Primer	Sequence 5' -> 3'	Length (bp)	Tm (°C)	GC (%)
AGPAT5-Ex2-F	TCA TGG ATG GAA ATA TTC AC	20	52.3	35
AGPAT5-Ex-R	ACGGTATTTTGAATGCCCAT	20	54.3	40
DEF104B-Ex2-F	AAC CTC TAC CAC CAT GCA CT	20	58.4	50
DEF104B-Ex2-R	AGG TCC TAC TTC CAG CGA CT	20	60.5	55
DEF106B-Ex1-F	AGT CAT ACA ACC GTA TCT GG	20	56.4	45
DEF106B-Ex1-R	AGA AGC TAG GTT ATG TAT GC	20	54.3	40
DEF105B-Ex2-F	TTA TGC TAT GTC ACA TGT GA	20	52	35
DEF105B-Ex2-R	ATT CTA TTT CTG CCA CTATC	20	52	35
DEF105B-Ex3-F	TCT CCC CAA CTC CCT ACT	20	58	56
DEF105B-Ex3-R	TCC AGG TCT GCT TCT AAG GC	20	60.5	55
DEF105B-Ex1-F	CTG CAC TTC CTT CAT CAG CT	20	58.4	50
DEF105B-Ex1-R	CAG TGA CTG GTT TGC CTT GG	20	60.5	55
PRR23D1-Ex1-F	CTT GCT GGA AGA GAG CCT TA	20	58.4	50
PRR23D1-Ex1-R	CCA CTG GAT GTC CCT TGA AC	20	60	55
ZNF705B-Ex1-F	AAG TGT CTG CAA GTG GAA AT	20	54.3	40
ZNF705B-Ex1-R	ATT TGC CAG CAA TAT GGG TA	20	54.3	40
ZNF705B-Ex2-F	GGA CAT TGA AAG GGA TTT CT	20	54.3	40
ZNF705B-Ex2-R	TGA ATG AGC GAA TGT GTG TC	20	56.4	45
MSRA-EX3-F	GGA GAA ATA GGC TCA TTG ACA CA	23	61	43
MSRA-EX3-R	TGC ACA CAA TTT GGG TTT TC	20	59	40
SOX7-Ex2a-F	AGG GAG GTG TGT GTG CAT GT	20	60.5	55
SOX7-Ex2a-R	CCG GGT ACA GGG GAC ATC AT	20	62.5	60
SOX7-Ex2b-F	ACC CGT ACT CAC CGG AGT AC	20	62	60
SOX7-Ex2b-R	TGG ACG GCT CCT CTG CCA CT	20	65	65
FDFT1-Ex2-F	ACC GCC GTG TGT GTT GTC TG	20	60	62.5
FDFT1-Ex2-R	GGT CCC GGA CAA GCT TTC CT	20	62.5	60
FDFT1-Ex1-F	AGT CCC ACT CCC ACT CCC ACT	21	65.3	62
FDFT1-Ex1-R	AGG TCC CGG ACA AGC TTT CCT C	22	65.8	59

TUSC30-Ex10-F	AGG CAC CAT CGT CTA TCC CCT	21	63	57
TUSC30-Ex10-R	TGA GGA CTT CTC CCT GTG GC	20	62.5	60
EGR3-Ex2a-F	GCT CAG CGT TCC TCT TAC CC	20	62.5	60
EGR3-Ex2a-R	GTA ATA GGG GGC GGG TTG AC	20	62.5	60
EGR3-Ex2b-F	ATC TCT TCC CCA TGA TTC CTG	21	59.5	48
EGR3-Ex2b-R	AGA TCC CCA CTT TTC CCC T	19	57.5	53
ADAMDEC1-Ex2-F	TCC AGC ATT CAC AAT CTC CTT	21	57.5	43
ADAMDEC1-Ex2-R	AAA AAT GGA GAC ATT GGG TCA	21	55.4	38
PTK2B-Ex2-F	TGA GAG AGC AGA GGC ACT GA	20	60	55
PTK2B-Ex2-R	ATT GTC ATT GGG CCA GTT TC	20	60	45
SCARA3-Ex6-F	CCC CTG AAG TGA GGA ACA AA	20	58.4	50
SCARA3-Ex6-F	TCC CTT TCT CCC TTT CCC TA	20	58.4	50
SMIM18-Ex1-F	TCA GCT CCA TTG TGG TTC AA	20	57	45
SMIM18-Ex1-R	GGA GCA AAG CTT TCT GGT TG	20	58.4	50
SMIM18-Ex2-F	GGA GGG CTG AAT TTG TCT GT	20	58.4	50
SMIM18-Ex2-R	GAA AAG GCA GCT GAA TTG GT	20	58.4	50
FDFT1-Ex1-F	TAT CAC GCC AGT CTC CTT CC	20	60	55
FDFT1-Ex1-R	GGG AAG GCT CGA GGA AAG	18	60	60
NAT1-Ex1-F	TCT GCA GAA AAA CAC CAG CA	20	60	45
NAT1-Ex1-R	TCC ACC GTA GGT ATG GAA GG	20	60	55
SMIM18-Ex3-F	TAA GGA CGT TGG CCC AGA TA	20	58.4	50
SMIM18-Ex3-F	GCA ACA GTT GGC CTC AGA AT	20	58.4	50
RBPMS-Ex8-F	AGG CCA CAG AGA AGA CAT CC	20	60	55
RBPMS-Ex8-F	GGT TCA AGC GAT TCT CCT GC	20	60	55
SLC39A14-Ex9-F	TCA AGT GAT CTT CCC GCC TT	20	58.4	50
SLC39A14-Ex9-F	AGC AAA AGA AGC AAG CCA CA	20	56.4	45
NPM2-Ex1-F	TAT CCG CTG AAA CCC CAC TT	20	58.4	50
NPM2-Ex1-F	CAT TCT TTC CTC CTC TGC GC	20	60	55
DPYSL2-Ex1-F	CAG ATC CAA CTT TGC CGC TT	20	58.4	50
DPYSL2-Ex1-F	CTC AAG CTC AAG GAG ACG C	20	59.5	58
ELP3-Ex1-F	AAG ATG CCA CTC CGT TCA GA	20	58.4	50
ELP3-Ex1-R	TCG CGG TAA AAG TTC GGT TC	20	58.4	50

CNOT7-EX7-F	GTCAGGACGTGCGAA AAGC	19	62	57.8
CNOT7-Ex7-R	TTC TAC TTC CCT AAC CCG CC	20	60	55
FZD3-Ex8-F	CTG TCC TTT TGC AGA GAA TGG	21	59.86	47.62
FZD3-Ex8-R	CAA AGC AGT CAC CAC ACA TTT	22	58.69	42.86
ADRA1A-Ex1-F	CTG GGA GGG TGA AAC CAC TA	20	59.96	55
ADRA1A-Ex1-R	TGG TGA AAC CCC GTC TCT AC	20	59.7	55
PRSS55-Ex5-F	CCG GAA AAT TGC TAA GAG AAG A	22	59.86	40.91
PRSS55-Ex5-R	GGC TAC TCC CAT CAC AAA GC	20	59.7	55
PPP2CB-Ex7-F	CCA TTC CAG CCG TAA GTG AT	20	59.96	50
PPP2CB-Ex7-R	ATG CCA AGA CAG ATC CCA AT	20	59.37	45
DEFB107A-EX2-F	ACT TAG CCA GGA CAG CAATT	20	57	45
DEFB107A-Ex2-R	TTT GTT TCT TGC AGC TGG TG	20	60	45
DEFB107B-Ex2-F	GCC TGG GCG ATG TAA TTT T	19	59.91	47.37
DEFB107B-Ex2-R	GCC TGG AAG AAG TTC ACA GC	20	60	55
ZDHHC2-EX1-F	CTG CGG GAT GGG GAG TTAG	19	62	63
ZDHHC2-Ex1-R	CCC CGA GGC TGC TTT TAG	18	60.8	61.11
DEFB105A-Ex1-F	CAC TTC ACG GAT CAA GAG CA	20	50.98	50
DEFB105A-Ex1-R	GCC ATG AGA GCT TGA CCA TT	20	60.23	50
EXTL3-Ex1-F	CTT TGC AGG ACG CGA GTT	18	60.14	55.6
EXTL3-EX1-R	CTG CGG GAT GGG GAG TTA G	19	62	63
DEFA1-Ex2-F	GGA GAA CAG GAG CAT TCG AG	20	59.95	55
DEFA1-Ex2-R	TGA GAC CCA ACT CTG GAA CC	20	60.09	55
DUSP4-Ex1-F	CAC GGC TAG TCC TTC CTT CA	20	60.3	55
DUSP4-Ex1-R	GCT GCT CAG TCC CAA CTT TC	20	60	55
SCARA5-Ex3-F	TGA GGG TTT CTC TGG TCT CG	20	60.38	55
SCARA5-Ex3-R	GCC TCC TTT GAC ACT TGA GC	20	60	55
GSR-Ex1-F	GCT GAG GCA GGA GAA TCA CT	20	59.56	55
GSR-EX1-R	AGA TCT ATG GGT CCC ACT GC	20	60	55
CHMP7-Ex1-F	GGG GCA AGA AGT GAC AAA CA	20	61.08	50
CHMP7-Ex1-R	TAT GTC ATT GGG GAA CAG CA	20	59.92	45
RHOBTB2-Ex1-F	GAA CGC TTT CCA CCA ACT G	20	59.28	52.63
RHOBTB2-Ex1-R	CGC AAG CAG CCA TCT ACT TA	20	59.23	50

EBF2-Ex1-F	TGC CCC TCA TTC TGT GTC AT	20	59	50
EBF2-Ex1-R	CGT GCT GGT ACC CTG TTT TG	20	60	55
WRN-Ex1a-F	GCA CTC ACA TCA TCT CTG CC	20	60	55
WRN-Ex1a-R	TTC CCA TCT TCC CAC CTG TC	20	60	53
WRN-Ex1b-F	GAC TGG ACA GGT GGG AAG AT	20	60	55
WRN-Ex1b-R	CCA CGT GTC TCA GAA GCC TA	20	60	55
SARAF-Ex1-F	AGT GAC GTA GGG TTG GCG	20	60	60
SARAF-Ex1-R	CGG AAA AGG GCG AAC GAA G	19	60	59
ADAM28-Ex14-F	GGG TAA AGC AGG TAG GGA GA	20	60	55
ADAM28-Ex14-R	CGC ACA CAC ACA GAT TTC CA	20	59	50
ENTPD4-Ex13-F	CTT TCC ACC CCT TTC CTT GC	20	60	55
ENTPD4-Ex13-R	TCC CAG GTT CAT GCC ATT CT	20	59	50
MCPH1-Ex8a-F	AGC TCT CCG TTA ATG TTT TCC A	22	59	41
MCPH1-Ex8a-R	GCT TGA TGG CAG CTG AGA TT	20	58	50
MCPH1-Ex8b-F	AAT CTC AGC TGC CAT CAA GC	20	58	50
MCPH1-Ex8b-R	AGT TCC ACA AGA CAG CAA GT	20	57	45
XKR6-Ex3a-F	GGA GCT GGG AAG AGA GTC AA	20	60	55
XKR6-EX3a-R	TGT AGT CCT GCA CGA ACC AG	20	60	55
XKR6-EX3b-F	CCT CGA CTA CTA CCG CAA GG	20	62	60
XKR6-Ex3b-R	ACC CCG CAG TGC TCT TAC	18	59	61
	ACC CCG CAG TGC TCT TAC	18	59	61
PURG-Ex2a-F	CCA TCC CGT GTG TCT GTT TG	20	59	55
PURG-Ex2a-R	AGG ACA CTG TGA GGA TGC TC	20	60	55
PURG-Ex2b-F	AGC ATC CTC ACA GTG TCC T	19	58	53
PURG-Ex2b-R	CTT GTT CTT CAC CAC TGG CC	20	60	55
DEFB4B-Ex2-F	CAG CCA TGA GGG TCT TGT AT	20	58	50
DEFB4B-Ex2-R	TGG AGG GGA ATG AGA GGA GA	20	60	55
FZD3-Ex1-F	AGA GGC AGA GAA CAA CGA GG	20	60	55
FZD3-Ex1-R	CCT CAA ACT AGG CGT CCC C	19	61	63
KIAA1456-Ex1-F	ACG TGG CTA TGT GTT GAG GT	20	58	50
KIAA1456-Ex1-R	AGC AAT GGG GAG CAA AAC AG	20	58	50
FZD3-Ex1-F	CGA GAG GCA GAG AAC AAC GA	20	61.67	55
FZD3-Ex1-R	GAA CGA TCC CTG AGG AGG A	19	60.14	57

PINX1-Ex1-F	CGT GTC ACT CAG CCT TGT CT	20	59.05	55
PINX1-Ex1-R	TCT GGG GTG AAC TCT GCT GT	20	60.86	55
PNMA2-Ex1a-F	AGC AAA GAA GGG GTC TGA CA	20	59.84	50
PNMA2-Ex1a-R	AGT GGT GCA GCT ATG CAC TG	20	59.66	55
PNMA2-Ex1b-F	CCT AGT CTG GGA CCC ATT CA	20	59.53	55
PNMA2-Ex1b-R	TTC CAA CCA GAC CTC AAA GG	20	60.8	50
PNMA2-Ex1c-F	ATT TGT TGG GAC AGG CAA TG	20	60.07	45
PNMA2-Ex1c-R	GGT CTT AGC CCA CTG GCT GT	20	62	60
DEFB104B-Ex1-F	CCC TCC GAA AAC ACC TTC TA	20	60	50
DEFB104B-Ex1-r	CGA CTC TAG GGA CCA GCA CT	20	62	60
NKX3-1-Ex2-F	AGA ACG ACC AGC TGA GCA C	19	57.89	59.14
NKX3-1-Ex2-R	GAG TCC TGC GGT GTC AGA AT	20	60.27	55
PSD3-Ex16-F	TGG CGA TGG AAG ATG GAA	20	54	50
PSD3-Ex16-R	GGC GAC AAA GCA ACA CAA AC	20	59	50
PDLIM2-Ex1a-F	GTT TGG GAT TTG TGG GTG TC	20	60.07	50
PDLIM2-Ex1a-F	GCT GCC GTC TAC CCG ACT	18	61.4	66.67
PDLIM2-Ex1b-F	TCC CCA GAG TCG GGT AGA C	19	60.06	63.16
PDLIM2-Ex1b-R	CTA GCC TCT GTT CCC AGC AC	20	60.01	60
GSR-Ex13-F	CAT GCT TAG TCA CCG TGA GG	20	60	55
GSR-Ex13-R	CCG AAA GAC CGA GAC AAA GA	20	59	50
NEIL2-Ex3-F	GCA AAA GGC AGA GTC CAA AG	20	59.99	50
NEIL2-Ex3-R	ATG TCC CCA CTG CTG AAG TC	20	60.12	55
PRR23D1-Ex1-F	TCA GCA CCA CTA GCC AAT CA	20	60.41	50
PRR23D1-Ex1-F	GCC CTA GTT TGG CCC TTT AG	20	60.09	55
MFHAS1-EX2a-F	ACA GTG GCA ACC TGA AGA CC	20	60	55
MFHAS1-EX2a-R	CCT CAG AGC ACT CAC CAC CT	20	60.46	60
MFHAS1-EX2b-F	CTG GAC GTG AGC CAC AAC	18	58.73	61.11
MFHAS1-EX2b-R	GGT AGC GGA TGC GGT TAT TA	20	59.95	50
MFHAS1-EX2c-F	GAG CTC TAC CTT AGT CGC AA	20	58.4	50
MFHAS1-EX2c-R	GAA AGT GGC GAG GCT CAT AG	20	59.98	55
MFHAS1-EX2d-F	GTT ATG AGG TGA TCC AGC CC	20	60	55

MFHAS1-EX2d-R	AAA TGC AGT TCC TCC AGC AC	20	60.26	50
MFHAS1-EX2e-F	CCA ACT TAC ACA GAG TAC TG	20	56.6	45
MFHAS1-EX2e-R	CTT GCC CTT GGG TTT ATT GA	20	59.93	45
MFHAS1-EX2f-F	TGC TGG AGA AGA TGG GAC TC	20	60	55
MFHAS1-EX2f-R	CCC TGT GTT TCC CTT ACG AA	20	59.96	50
PRR23D1-Ex2-F	AAT GAA GTT TGG GGC TTG TG	20	59.97	45
PRR23D1-Ex2-R	CAA GCA CGA CTT TCC ACT CA	20	60.02	50
TRIM35-Ex1a-F	GTT TGC TGA TTC CCC ATT TC	20	59.37	45
TRIM35-Ex1a-R	GCC ATC CAG TTC TTC CTT GA	20	60.2	50
TRIM35-Ex1b-F	GGG CGA GCT GTC TTT CTA TG	20	59.98	55
TRIM35-Ex1b-R	TAC TCT GCC TTG CCC ATT TC	20	60.21	50
MTUS1-Ex2-F	AGT GTG CTT GTG TCG TGA GG	20	59.94	55
MTUS1-Ex2-R	TAC AGT TTT CAG GCC GCT TT	20	59.88	45
USP17L8-Ex1a-F	TCC CAG CAG AAA AAC CTT GT	20	59.71	45
USP17L8-Ex1a-R	CTT CCT GCT TGC CTC TAT GG	20	59.97	55
USP17L8-Ex1b-F	CCA TAG AGG CAA GCA GGA AG	20	59.97	55
USP17L8-Ex1b-R	TGA TGC TAC AGG CAG TGA CC	20	59.86	55
USP17L8-Ex1c-F	GGT CAC TGC CTG TAG CAT CA	20	59.86	55
USP17L8-Ex1c-R	CGC TTG TGG GTG TAT TTG TG	20	60.03	50
CSMD1-Ex11-F	GCT CCG GAA TGT TGA TCT TG	20	60.6	50
CSMD1-Ex11-r	CTG GTT ATC CGT CCC TGT TG	20	60.37	55
LEPROTL1-Ex4-F	ACC AGA GAG AGC CAG GAA CA	20	59.99	55
LEPROTL1-Ex4-r	GTG TTG CCT ACA TGG TGC AG	20	60.18	55
NUGGC-Ex1-F	AGG AGG AAC ACC TGC TTT GA	20	59.84	50
NUGGC-Ex1-R	TTT CTT GAA CCC AGG AGG TG	20	60.08	50
CTSB-Ex10-F	GCC CCT CTT TTG GAT TCT TC	20	60.02	50
CTSB-Ex10-F	CAG TCC TGG CAA CAG TGA GA	20	60.02	55
FDFT1_45-Ex2-F	CCC GAG GGT TAC ACA TCT GA	20	60.91	55
FDFT1_45-Ex2-R	CAG GCA TGA GGC AGA TAA AA	20	58.87	45
FDFT1_48-Ex1-F	GGA AGG TGA TGC CCA AGA T	19	59.88	52.63
FDFT1_48-Ex1-F	GAG GTG GTG AAA CTG GAG GA	20	60.09	55
FDFT1_49-Ex1-F	CTG CTC CAC TGC GAA TAC AA	20	60.01	50

FDFT1_49-Ex1-R	CGC CCT CTA CTC GAA TAC GC	20	62.15	60
FDFT1_42-Ex1-F	CAA TCA GCA CCC TGT CAA AA	20	59.69	45
FDFT1_42-Ex1-F	GCC TTG TAT TGC CAA ATC CT	20	59.05	45
FDFT1_42-Ex2-F	TGT GCT CGG GTG TGT TAC AG	20	60.78	55
FDFT1_42-Ex2-F	GGA AGG AGA CTG GCG TGA TA	20	60.22	55
MSRA-EX1-F	GTT TGG GCA ACC TCG ATT AC	20	58	50
MSRA-EX1-R	ACA AAG GGA CGT CAG TGG AC	20	60.01	55
MICU3-Ex1-F	AGT GTA GGG GTC GCT CTC G	19	60.42	63
MICU3-Ex1-R	AGC CCT CTC CTC ATC CTC TC	20	60	59.91
TUSC3-Ex1-F	TAG GCC CCA GGT AAA GTG CT	20	60	55
TUSC3-Ex1-R	CCA AGG GGA TCC ATT CTA CC	20	60.52	55
MSR1-Ex1-F	AGT GCC TGA TTC CCC ACA T	19	60.34	52.63
MSR1-Ex1-R	ACA ACA CGG GAA CCA AAG TC	20	59.87	50
SCARA5-Ex1-F	TTC TCG CCC TCC TGT GTA TC	20	60.22	55
SCARA5-Ex1-R	TCGGAGCACATGTTCAAGAG	20	60.02	50
SLC25A37-Ex1-F	GTC CGG TTG CTT GGA AAT G	19	61.43	52.63
SLC25A37-Ex1-R	CTC CTT CTC CGC TCG TTG	18	59.65	61.11
R3HCC1-Ex4-F	TGA TTA CCT CCT CCC ATT CC	20	58.4	50
R3HCC1-Ex4-R	CCA CAC CAT AAC CCT TGG TC	20	60.09	55
TEX15-Ex4-F	CAG TGG ACA GAC AGA GCA CA	20	58.5	55
TEX15-Ex4-F	CCA GGG ACA ACT TCT GAG GA	20	60.23	55
ZNF705B-Ex1-F	TTC TTT CCC TCA CTG CCA CT	20	59.84	50
ZNF705B-Ex1-R	CCC TTC TTT AGC CCC AGT TG	20	60.97	55
ZNF705B-Ex2-F	TGT CCA TGT CTG CTC AAT GC	20	60.85	50
ZNF705B-Ex2-R	AGG GCA TCC ACT TCA TCA TC	20	59.89	50
NAT1-Ex1-F	ATA CGG TTT GTC CCC TGT TG	20	59.71	50
NAT1-Ex1-R	AAC CAA GCA GCC CTC TAC AC	20	59.36	55
SORBS3-Ex1-F	CAG CTG GGA AAC CAG AGA G	19	59	58
SORBS3-Ex1-R	TTC TGG TAT GAG GGC TGG TG	20	60	55
STMN4-Ex1-F	CTT GTG TGA CCT GCC CTT TT	20	58	50
STMN4-Ex1-F	ACA CAG ACC TCA TCC CAC CT	20	59.4	55
ZNF705B-Ex6-F	CAT TTG ACA CAC TTG CCA CA	20	59.14	45

ZNF705B-Ex6-R	CAC TCA GGT GAT TGT GCT TCA	21	59.89	47.62
EGR3-Ex1-F	GAA GGA AGG GGA GAG GAA GA	20	59.75	55
EGR3-Ex1-R	TGG GAG GCT GAG GGA GTA A	19	60.74	57.89

A.2.2. Primer pairs to check segregation in HWE 244 whole genome analysis

Primer	Sequence 5' -> 3'	Length (bp)	GC (%)	Tm (°C)
NEFL-3UTR-F	GCGTAATTCTGTTGCCTCAA	20	45	56.4
NEFL-3UTR-R	ACATGGGGCGTGTATTTGAT	20	45	56.4
MSR1-3UTR-F	TGCAAAATCTCCAATCCACA	20	40	54.3
MSR1-3UTR-R	TCCTGTTAAGTGGCATTTTTGA	22	36	56.4
PTK2B-Intron-F	GTG GAC AGA GGC AAT GCA C	19	58	59.4
PTK2B-Intron-R	GCT CTC ACC TTC ACC ACC AC	20	60	62.5
PDLIM2-5UTR-SF	TAAGAGCCCCAGATGGACAC	20	55	60.07
PDLIM2-5UTR-SR	GCAGGTGACAGCTAAGCACA	20	55	60.21
WRN-3UTR_SF	GCCTCTTCACGGAAAACAAC	20	50	59.72
WRN-3UTR_SR	ATGTAAGGCTGTGCATGGTG	20	50	59.6
BLK-IN8-SF	CGA AGT CTG GAT GGG TGA GT	20	55	60.11
BLK-IN8-SR	CAT TGT CAA AGC CAG CTT CA	20	45	59.99
C8orf58-Intron-F	AAC GTA CCA CTT GCC ATC CT	20	50	58.4
C8orf58-Intron-R	ACC TTG ACC TTG TCC CAG TG	20	55	60.5

A.2.3. Primer pairs for screening *R3HCC1* in HWE patients and control cohort

Primer	Sequence 5' -> 3'	Length (bp)	GC (%)	Tm (°C)	Std Cond.
R3HCC1-5'UTR-F	GAG CCC ATC TTG GAA TCA GA	20	50	60.16	60°C + 10%
R3HCC1-5'UTR-R	AAG TCA GGG GTC AGG GAA G	19	57.89	59.09	
R3HCC1-Ex1-F	GAG CCC ATC TTG GAA TCA GA	20	50	60.16	60°C + 10% DMSO
R3HCC1-Ex1-R	AAG TCA GGG GTC AGG GAA G	19	57.89	59.09	
R3HCC1-Ex2-F	CCT TGG ACC AGA GGG TTT CT	20	55	60.48	60°C
R3HCC1-Ex2-R	CTA AAG GCC AAG GTC AGA GC	20	55	59.08	
R3HCC1-Ex3-F	TTA ACC TGG GAG TGG ACC TG	20	55	59.96	
R3HCC1-Ex3-F	TTC CAA AGG CTG ACC AGC TA	20	50	60.9	
R3HCC1-Ex4-F	TGA TTA CCT CCT CCC ATT CC	20	58.4	50	60°C
R3HCC1-Ex4-F	CCA CAC CAT AAC CCT TGG TC	20	60.09	55	
R3HCC1-Ex5-F	CTG CCT TTC CCA GAC TTT TG	20	50	59.85	60°C
R3HCC1-Ex5-r	ATA CAG AGC CCA GGC AAC AC	20	55	60.14	

R3HCC1-Ex6-F	TGC TGC AAA TGA TTG TCT CC	20	45	59.81	60°C
R3HCC1-Ex6-R	AAG GTC GCC TGA ATC TGA TG	20	50	60.22	
R3HCC1-Ex7-F	TGC CCT GAG CTT TGA ATA GG	20	50	60.34	60°C
R3HCC1-Ex7-F	CCC AAA CCA GAC ACA CAC AC	20	55	59.89	
R3HCC1-Ex8-F	TCTGATGGCTTGTACGGCTG	20	55	61	55°C
R3HCC1-Ex8-F	CGTCCCCTACAGACCACATT	20	55	61	

A.2.4. Primer pairs for cloning of different genes

Primer	Sequence 5' -> 3'	Length (bp)	GC (%)	Tm (°C)	Site
R3HCC1-pCDNA3.1-F	GCTGGCTAGCGCCACCCTGGCCCTTCTCTGCTTGGATGG	39	67	85	NheI
R3HCC1-pCDNA3.1-R	TCTAGACTCGAGTCAGGGCGGCAGCGGACCCCGG	34	71	83	XhoI
R3HCC1-eGFPN1-F	GCTGCTCGAGGCCACCCTGGCCCTTCTCTGCTTGGATGG	39	67	85	Xho I
R3HCC1-eGFPN1-R	TCTAGACCGCGGGGGCGGCAGCGGACCCCGGAC	33	79	86.4	Sac II
R3HCC1-p3XFlag-F	GCTGAAGCTTCTGGCCCTTCTCTGCTTGGATGGTGT	36	56	79.1	HindIII
R3HCC1-p3XFlag-R	TCTAGAGAATTCTCAGGGCGGCAGCGGACCCCGG	34	65	81.4	EcoRI
R3HCC1-eGFPC2-F	GCTGGAGCTCTCCTGGCCCTTCTCTGCTTGGATGGTGT	38	61	72	SacI
R3HCC1-eGFPC2-R	TCTAGAGAATTCTCAGGGCGGCAGCGGACCCCGG	34	65	81.4	EcoRI
Y14-p3XFLAG-F	GCTGAAGCTTATGGCGGACGTGCTAGATCTTCAC	34	53	66.8	HindIII
Y14-p3XFLAG-R	TCTAGAGAATTCTCAGCGACGTCTCCGGACTGG	33	52	65.6	EcoRI
PABPC1-p3XFLAG-F	GCTGGGTACCCATGAACCCCAGTGCCCCCAGCTAC	34	65	71	Kpn1
PABPC1-p3XFLAG-R	TCTAGAGGATCCTTAAACAGTTGGAACACCGGT	33	45	63	BamH1
DCP1A-eGFPN1-F	GCTGCTCGAGGCCACCATGGAGGCGCTGAGTCGAGC	36	69	74	Xho1
DCP1A-eGFPN1-F	TCTAGACCGCGGTAGGTTGTGGTTGTCTTTGTTCTTG	37	49	66	SacII
MAGOH-p3XFlag-F	GCTGAAGCTTATGGAGAGTGACTTTTATCTGCGTTAC	37	43	64	HindIII
MAGOH-p3XFlag-R	TCTAGAGAATTCCTAGATTGGTTTAATCTTGAAGTG	36	33	59	EcoRI
EIF4A3-p3XFlag-F	GCTGGCGGCCGCCATGGCGACCACGGCCACGATGGCG	33	78	78	Not I
EIF4A3-p3XFlag-R	TCTAGATCTAGATCAGATAAGATCAGCAACGTTC	34	38	60	Xba I
R3HCC1-Cterm-p3XFlag-F	GCTGAAGCTTATGCTG GAAAAGGGGAAGGAGAG	33	52	65	HindIII
R3HCC1-Cterm-eGFPC2-F	GCTG GAGCTC TC CTG GAAAAGGGGAAGGAGAG	32	59	68	SacI

A.2.5. Primer pairs used for site-directed mutagenesis to create patient-specific variants and truncations in *R3HCC1*

Primer	Sequence 5' -> 3'	Length (bp)	GC (%)	Tm(°C)
R3HCC1-R431Q-F	CACAAAAAGAAAGAGCAGCCTGCTGTCCGGGG	32	56	80.34
R3HCC1-R431Q-R	CCCCGGACAGCAGGCTGCTCTTTCTTTTTGTG	32	56	80.34
R3HCC1-A433T-F	GAAAGAGCGGCCTACTGTCCGGGGTCC	27	67	80.13
R3HCC1-A433T-R	GGACCCCGGACAGTAGGCCGCTCTTTC	27	67	80.13
R3HCC1-V142M-F	TATGTGCCCCGGATGCTGCGCAGGC	25	68	78.38
R3HCC1-V142M-R	GCCTGCGCAGCATCCGGGGCACATA	25	68	78.38
R3HCC1-SDM-W148*-F	GCAGGCAGGAAGAATGAGGGCTGACCTCT	29	59	65.7
R3HCC1-SDM-W148*-R	AGAGGTCAGCCCTCATTCTTCCTGCCTGC	29	59	65.7
R3HCC1-SDM-G210*-F	GAGTCAGTCAGGGAAGTGAGACATGGTGGAGAT	33	52	65
R3HCC1-SDM-G210*-R	ATCTCCACCATGTCTCACTTCCCTGACTGACTC	33	52	65
R3HCC1-G312*-F	GTGGAGGAGCTGCCTTGAGAGAAGGACCTTG	31	58	80.31
R3HCC1-G312*-R	CAAGGTCCTTCTCTCAAGGCAGCTCCTCCAC	31	58	80.31

A.2.6. Primer pairs spanning cDNA of *R3HCC1*

Primer	Sequence 5' -> 3'	Length (bp)	GC (%)	Tm (°C)
R3HCC1-cDNA-5UTR-F	GCTCTCCCACCTGTCACC	18	67	60.8
R3HCC1-cDNA-1F	TGGATGGTGTCTTCCTCTCC	20	55	60.05
R3HCC1-cDNA-1R	TCTCTTGAGCACCGAGGTAG	20	55	58.19
R3HCC1-cDNA-2F	GTGGTATCGTGGACGCAAG	19	57.89	60.13
R3HCC1-cDNA-2F	CGAAGGAGGATGTGTCCAAA	20	50	61.03
R3HCC1-cDNA-3F	GCTGCAGGAGATCACAGACA	20	55	60.15
R3HCC1-cDNA-3F	AGGCCGCTCTTTCTTTTTGT	20	45	60.37

A.2.7. Primer pairs used for *Drosophila* study: Primers to confirm P element insertion at genomic DNA (gDNA) and cDNA primers for *CG2162*

Primer	Sequence 5'--> 3'	Length (bp)	Tm (°C)	GC (%)
CG2162-gDNA-intron 4F	TCTGTTTTGCTCCGCTCTTT	20	49.7	45
CG2162-gDNA-NP3333-PGawB- R	GCACTTCGGTTTTTCTTTGG	20	59.72	45
CG2162-gDNA-NP3333-PGawB- F	TGTGGAATTGTGAGCGGATA	20	60.07	45
CG2162-gDNA-exon 5R	GTCGGTGTTCTTCGTTGGAT	20	59.97	50
CG2162-cdna-1F	TGGAGGCATTCGCAAATAAT	20	60.42	40
CG2162-cdna-1R	TTCTGCCGTTCGTGTGATAG	20	59.86	50
CG2162-cdna-2F	GCTTTCGAAACCAGCTTCTG	20	60.13	50
CG2162-cdna-2R	CTGTTTGGGGATTCTTTGGA	20	59.9	45
CG2162-cdna-3F	CACCATTAGCTGCAGGAACA		59.86	50
CG2162-cdna-4F	CGTCCAAGTCGAAGAAAAGC	20	59.99	50
CG2162-cdna-4R	CCTTTGTTGCCGATGTCTCT	20	60.25	50
CG2162-cdna-5F	TCAACTCAGATGTGCTGGAGA	21	59.56	47.62
CG2162-cdna-5R	TGGCTTCAACTTGACAAACG	20	59.88	45
CG2162-cdna-6F	CAGTACAGGGGCAGTGGTTT	20	55	60.03
CG2162-cdna-6R	TTCACGTAGTACCCGCCTCT	20	55	59.76
CG2162-ex5-6-R	GCCGCTCATGATCTCCTTAT	20	50	59.25

A.2.8. Primer pairs used for segregation in HWE 277 genome analysis

Primer	Sequence 5'--> 3'	Length (bp)	Tm (°C)	GC content (%)
EGF-3UTR-SF	CAAGGCACCCTGCAGAGATA	20	61.34	55
EGF-3UTR-SR	GCTAGAAAAGAGGCCAATATGC	22	59.4	45.45
FAM241A-3UTR-SF	TGCATGCCTGTAATCTCAGC	20	59.98	50
FAM241A-3UTR-SR	TGTTGGGCAGAGAGACAAGA	20	59.55	50
AIMP1-5UTR-S1F	TCTTGGGATTGAAGCCATTT	20	59.5	40
AIMP1-5UTR-S1R	CGATGCCTCTAATGCAGAATC	21	59.82	47.62
AIMP1-5UTR-S2F	ATTCTCCTGCCTCAGTCTGC	20	59.56	55
AIMP1-5UTR-S2R	TCACTTCCTAGACAGGATGACG	22	59.35	50
LEF1-5UTR-SF	ACCCTCCTCTGCACTTTGG	19	60.24	57.89
LEF1-5UTR-SR	TTGAAGGGGATCATCTCGTC	20	60.01	50

A.2.9. Primer pairs for generating constructs for luciferase assay in HWE 277

Primer	Sequence 5'--> 3'	Length (bp)	Tm (°C)	GC %
EGF-3UTR-F-SPE1	CTAG ACTAGT TGAGAACACGGCTACCCATA	30	61	47
EGF-3UTR-R-MLU1	CTG ACGCGT GCTAGAAAAGAGGCCAATATGC	31	64	52
FAM241A-3UTR-F-SPE1	CTAG ACTAGT TGCATGCCTGTAATCTCAGC	30	61	47
FAM241A-3UTR-R-MLU1	CTG ACGCGT CAAATTGGGGAGCTGTTAGG	29	64	55
EGF-3UTR-SDMG>A-F	CTCATCGTCTCATTGCACGCAACGCCTGATTGA	33	65	52
EGF-3UTR-SDMG>A-R	TCAATCAGGCGTTGCGTGCAATGAGACGATGAG	33	65	52
FAM241A-3UTR-SDMG>A-F	CTCAATCTGGGCAACAAGAGCAAAACTCCATCTCAAAAAAAAA	43	64	37
FAM241A-3UTR-SDMG>A-R	TTTTTTTTTGAGATGGAGTTTTGCTCTTGTTGCCCAGATTGAG	43	64	37

A.2.10. Sequence of CRISPR sgRNA target for *ZGRF1*

Target	Target sequence
HSL-ZGRF1-exon 5	TAGGTATTGTTAAGCAGAA
HSR-ZGRF1-exon 5	AGCACCAGAGTTAAATTCA

A.2.11. Primer pairs to confirm *ZGRF1* knockout at genomic DNA and cDNA in mammalian cells

Primer	Sequence 5'--> 3'	Length (bp)	Tm (°C)	GC (%)
ZGRF1-KO-Exon 5-FP	AGTCTTGCGATTACGGGTGT	20	58.4	50
ZGRF1-KO-Exon 5-RP	AAGTGCCCTTTGAATGACAGA	21	57.5	43
ZGRF1-cDNA-P1-F	GGAAAGCCAAGAATTTATTG	20	52.3	35
ZGRF1-cDNA-P1-R	TACATCTTTCTTGCCAACAG	20	54.3	40

A.2.12. Primer pairs to confirm *Zgrf1* conditional knockout mice at genomic DNA

Primer	Sequence 5'--> 3'	Length (bp)	Tm (°C)	GC (%)	Std condition Ta°C
mzgrf1-gdna-F2	CTTTGTCCATGCTTGTTGTTCTGA	24	62	42	60
mzgrf1-gdna-R2	GAATGCCTACTAAGAAAATGCCTCT	25	62.5	42	
mzgrf1-gdna-F3	ACCTATGTAGCCCATGACCTTTG	23	62.9	48	58
mzgrf1-gdna-IN6R	TTCAATTTCCAGCAACCACA	20	60.9	55	

A.2.13. PCR amplification

1. The reaction mix was assembled into 20µl volume in a 0.2 mL PCR tubes.
2. Reagents were added in the following order: Water, PCR buffer, dNTPs, $MgCl_2$, primers, Taq polymerase, and template DNA according to the concentration mentioned below.

Reagents (NEB Kit)	Working concentration	Working volume µl (20µl reaction)
H_2O		12.4
PCR buffer 10X	1X	2
$MgCl_2$15mM	1.5mM	2
dNTPs 8mM	0.8mM	2
Primer forward 25uM	2.5uM	0.2
Primer forward 25uM	2.5uM	0.2
Taq Polymerase 5U/ul	0.5U	0.2
Template DNA 100ng/ul	100ng	1

3. Gentle mixing was done by tapping the tube followed by centrifugation for 6-7 seconds. The thermocycler for PCR reaction was programmed using the following parameters for 25-30 cycles on GeneAmp® 9700

PCR cycling condition	Temperature	Time
Initial Denaturation	94°C	5 minutes
Denaturation	94°C	30 seconds
Annealing	50-62°C	30 seconds
Extension	72°C	1 minute per kb
Final Extension	72°C	10 minute
Hold	15°C	∞

4. The amplified product was checked on 0.8%-1.2% agarose gel by electrophoresis.

A.2.14. PCR cleanup

1. The PCR amplified products were removed from the refrigerator before setting up the reaction.
2. The PCR cleanup instrument was assembled and to the amplified product (20µl), 80µl of autoclaved MQ water was added.
3. The 100µl mix was transferred to 96-well multiscreen filter plates and the plate was placed in a holder.
4. The instrument was switched on and the suction pressure was allowed to reach 15-20lbps.
5. Once the mix was passed through the filter, the pressure was released and 20µl autoclaved MQ water was added to each column and the plate was incubated at 37°C for 5 minutes with constant shaking.
6. The mix was transferred to 0.6ml tubes.

A.2.15. Cycle sequencing

1. The following components were used to make a reaction mix (Applied Biosystems):
   a) MQ water: 13.5µl
   b) 5X sequencing buffer: 3.5µl
   c) RR mix: 0.5µl
   d) Primer (forward or reverse): 0.03µl
2. The reaction was set for samples in multiple of 16 in 96 well sequencing plate.
3. The sequencing parameters were set on GeneAmp® 9700.

PCR cycling conditions	Temperature	Time
Initial Denaturation	95°C	1 minute
Denaturation	96°C	10 seconds
Annealing	50°C	5 seconds
Extension	60°C	4 minutes
Hold	4°C	∞

A.2.16. Sequencing product cleanup

1. 16µl MQ H_2O + 64µl 95% ethanol was added to the samples in the sequencing plate.
2. The samples were invert mixed and incubated for 30 minutes at room temperature.
3. The samples were centrifuged at 2500g for 30 minutes at room temperature.
4. The ethanol mix was decanted, and a short invert spin was given, followed by addition of 150µl of 70% ethanol to each well. Samples were invert mixed.
5. The sample plate was centrifuged at 2000g for 10 minutes at room temperature and the ethanol was decanted.
6. To the wells, 10µl of formamide was added and subjected to denaturation at 94°C for 5 minutes.
7. The reactions were the snap chilled and the sequencing plate was placed in the machine ABI 3730 Genetic Analyzer for sequencing.

A.2.17. Genomic DNA isolation from the blood samples

Reference: Genomic DNA was extracted from blood using the phenol-chloroform extraction method (Sambrook and Russel 2001).

1. 10ml of blood was taken and diluted to 20ml with cold NKM buffer. The sample was vortexed and centrifuged at 6000rpm at 4°C for 20 minutes.
2. About 10ml of the supernatant was retained with the pellet and it was vortexed to dissolve the clumps.
3. Resuspension buffer was added to the dissolved pellet upto a volume of 20ml and centrifuged for 30 minutes at 6000rpm at 4°C.

4. 0.5ml of 10X TEN solution, 0.25ml of 2mg/ml proteinase K and 0.5ml of 10% SDS was added to ~4ml of the retained supernatant. The supernatant was mixed gently and incubated for 4hrs at 50ºC to aid in digestion.
5. Following this, 5ml of equilibrated phenol pH 8.0 was added and vortexed and centrifuged for 20 minutes at room temperature.
6. The aqueous layer was collected in a fresh 15ml tube and to it an equal volume of chloroform: isoamyl alcohol in the ratio of 24:1 was added. This was centrifuged at 2000rpm for 20 minutes and step 6 was repeated.
7. Genomic DNA was precipitated by adding an equal volume of ice-cold isopropyl alcohol and 5M NaCl (0.1M NaCl as final concentration). The mix was centrifuged.
8. Pellet was transferred to 1.5ml tubes washed twice with 70% ethanol and air-dried.
9. DNA pellet was resuspended in 200µl TE buffer and stored in -20ºC.

NKM buffer	**Proteinase K, 2mg/ml (50ml)**	**Resuspension buffer**	**TEN solutions, 10X**
0.14M NaCl	10mM CaCl2	0.10M NaCl	0.21M Tris-Cl, pH7.5
1.5M MgCl2	10mM Tris-Cl, pH7.5	10mM Tris-Cl, pH7.5	0.29M EDTA, pH8.0
30mM KCl	2mg/ml proteinase K	1.5mM MgCl2	1.07M NaCl

TE buffer
10mM Tris-Cl pH8.0
1mM EDTA pH8.0

A.2.18. Genomic DNA isolation from mammalian cells

Reference: Chery M. Koh 2013 Isolation of Genomic DNA from mammalian cells *Methods in Enzymology*, Volume 529

1. For adherent cells, trypsinization was done and cells were collected in a 1.5ml tube. The cells were centrifuged at 500g for 5 minutes and the supernatant was discarded. This was followed by washing the cells twice in ice-cold PBS, following the centrifugation step.
2. The cells were resuspended in a suitable amount of Lysis buffer. (Used 200µl for one well of 24 well dish or 1ml buffer per 10^8 cells)
3. The samples were vortex and incubated at 50ºC for 12-18 hours.
4. To the sample, an equal volume of phenol extraction buffer was added and vortexed.
5. This was followed by centrifugation at 2000g for 5 minutes at room temperature.
6. The upper aqueous layer was transferred to a new centrifuge tube and 20µg/ml RNase A was added and the sample was incubated at 37ºC for 20 minutes.
7. Following this 0.5 volumes of 7.5M ammonium acetate and two volumes of 100 % ethanol was added, and samples were vortexed.
8. Samples were centrifuged at 200g for 10 minutes at room temperature.

9. The supernatant was removed, and the pellet was washed with 70% ethanol and centrifuged at 1700g for 5 minutes.
10. The supernatant was discarded, and the pellet was air-dried.
11. The DNA was resuspended in TE buffer.

Lysis buffer

100mM NaCl
10mM Tris-HCl, pH 8.0
25mM EDTA, pH 8.0
0.5% SDS
0.1 mg/ml Proteinase K

Phenol extraction buffer

Buffered phenol 25ml
Chloroform 24ml
Isoamyl-alcohol 1ml

A.2.19. Genomic DNA isolation from *Drosophila melanogaster*

1. Approximately 40-50 flies were crushed using chilled mortar pestle in 400µl of homogenizing buffer at 4ºC.
2. The homogenate was incubated at 70ºC for 30 minutes.
3. To the sample, 56µl of 8M potassium acetate solution was added, followed by incubation for 30 minutes on ice.
4. The sample was centrifuged at 13,000rpm for 10 minutes at 4ºC.
5. The upper layer is transferred to a fresh tube and an equal volume of phenol: chloroform (1:1) was added, followed by invert mixing and centrifugation at 13,000 rpm for 5 minutes. This step was repeated twice.
6. An equal amount of isopropanol was added to the supernatant to precipitate DNA.
7. Centrifugation at 13,000 rpm for 10 minutes was done at room temperature to pellet DNA.
8. The pellet was washed with 70% ethanol following centrifugation at 13,000rpm for 10 minutes.
9. Pellet was air-dried and resuspended in TE buffer.

Homogenization buffer:
0.2M Sucrose
0.1M Tris-HCl pH 8.0
0.05M EDTA pH8.0
0.5% SDS

A.2.20. RNA isolation

1. For adherent cells, trypsinization was done and cells were collected in a 1.5ml tube. The cells were centrifuged at 500g for 5 minutes and the supernatant was discarded. This was followed by washing the cells twice in ice-cold PBS and resuspension in 1ml Trizol per 10mm dish.
2. For tissues (Drosophila whole body/brain), homogenization was done in 1ml of Trizol.
3. The samples were incubated at room temperature for 5 minutes.

4. 0.2ml of chloroform was added per ml of Trizol reagent added, and samples were vortexed for 10 seconds. These were incubated at room temperature for 10 minutes.
5. The samples were centrifuged at 12000g for 15 minutes at 4ºC.
6. The upper aqueous layer was transferred to a new tube and 0.5 ml is isopropanol was added per ml of Trizol. Mix was incubated at -20 ºC for 2-4 hours and centrifuged at 12000g for 15 minutes at 4ºC.
7. The pellet was washed with 75% ethanol and centrifuged at 7500g for 5 minutes at 4ºC.
8. The pellet was air-dried and resuspended in nuclease-free water.

A.2.21. cDNA synthesis

Reference: Superscript III First strand synthesis kit for RT-PCR (Invitrogen Catalog no 18080-051)

1. The following mix was made: 1ug RNA + 1µl oligo dT primer(50µM) +10mM dNTP mix, volume made to 10µl with DEPC treated water.
2. The sample was incubated at 65ºC for 5 minutes, then on ice for 1 minute.
3. Each PCR contained:

Reagents	Volume
10X RT buffer	2 µl
25mM MgCl2	4 µl
0.1M DTT	2 µl
RNase OUT (40U/ µl)	1 µl
SuperScript III RT (200 U/ul)	1 µl

4. The 10µl of cDNA synthesis mix was added to the above PCR reaction. The PCR reaction was set as 50 minutes at 50ºC following 5 minutes at 85ºC. The samples were chilled on ice.
5. 1µl of RNase H was added to the tube and incubated for 20 minutes at 37ºC.
6. cDNA mix was stored at -20ºC.

References

Adamson B, Smogorzewska A, Sigoillot FD, King RW, Elledge SJ. 2012. A genome-wide homologous recombination screen identifies the RNA-binding protein RBMX as a component of the DNA-damage response. Nat Cell Biol 14:318–328.

Afshari P, Myles-Worsley M, Cohen OS, Tiobech J, Faraone S V., Byerley W, Middleton FA. 2015. Characterization of a Novel Mutation in SLC1A1 Associated with Schizophrenia. Mol Neuropsychiatry 1:125.

Agarwal V, Bell GW, Nam JW, Bartel DP. 2015. Predicting effective microRNA target sites in mammalian mRNAs. Elife 4:e05005.

Aguilera-Gomez A, Zacharogianni M, Oorschot MM van, Genau H, Grond R, Veenendaal T, Sinsimer KS, Gavis ER, Behrends C, Rabouille C. 2017. Phospho-Rasputin Stabilization by Sec16 Is Required for Stress Granule Formation upon Amino Acid Starvation. Cell Rep 20:935.

Alajouanine T, Gastaut H. 1995. Observations [Synkinesis-startle and epilepsy startle triggered by unexpected sensory and sensitive factors. I. Anatomical and clinical data on 15 cases]. Rev Neurol 93:29–41.

Allen AS, Berkovic SF, Cossette P, Delanty N, Dlugos D, Eichler EE, Epstein MP, Glauser T, Goldstein DB, Han Y, Heinzen EL, Hitomi Y, et al. 2013. De novo mutations in epileptic encephalopathies. Nature 501:217–221.

Allen I. 1945. Observation on cases of reflex epilepsy. N Z Med J 44:135–142.

Ames FR. 1971. “self-induction” in photosensitive epilepsy. Brain 94:781–798.

Andermann K, Oaks G, Berman S, Cooke PM, Dickson J, Gastaut H, Kennedy A, Margerison J, Pond DA, Tizard JPM, Walsh EG, Sherwood SL. 1962. Self-Induced Epilepsy: A Collection of Self-Induced Epilepsy Cases Compared with Some Other Photoconvulsive Cases. Arch Neurol 6:49–65.

Anderson P, Kedersha N. 2002. Stressful initiations. J Cell Sci 115:3227–3234.

Aoyama K, Sang WS, Hamby AM, Liu J, Wai YC, Chen Y, Swanson RA. 2006. Neuronal glutathione deficiency and age-dependent neurodegeneration in the EAAC1 deficient mouse. Nat Neurosci 9:119–126.

Arimoto K, Fukuda H, Imajoh-Ohmi S, Saito H, Takekawa M. 2008. Formation of stress granules inhibits apoptosis by suppressing stress-responsive MAPK pathways. Nat Cell Biol 10:1324–1332.

Aulas A, Velde C Vande. 2015. Alterations in stress granule dynamics driven by TDP-43 and FUS: A link to pathological inclusions in ALS? Front Cell Neurosci 9:423.

Auvin S, Lamblin MD, Pandit F, Bastos M, Derambure P, Vallée L. 2006. Hot water epilepsy occurring at temperature below the core temperature. Brain Dev 28:265–268.

Bac P, Maurois P, Dupont C, Pages N, Stables JP, Gressens P, Evrard P, Vamecq J. 1998. Magnesium deficiency-dependent audiogenic seizures (MDDASs) in adult mice: A nutritional model for discriminatory screening of anticonvulsant drugs and original

assessment of neuroprotection properties. J Neurosci 18:4363–4373.

Baron DM, Kaushansky LJ, Ward CL, Sama RRK, Chian R-J, Boggio KJ, Quaresma AJC, Nickerson JA, Bosco DA. 2013. Amyotrophic lateral sclerosis-linked FUS/TLS alters stress granule assembly and dynamics. Mol Neurodegener 8:30.

Bassell GJ, Zhang H, Byrd AL, Femino AM, Singer RH, Taneja KL, Lifshitz LM, Herman IM, Kosik KS. 1998. Sorting of β-actin mRNA and protein to neurites and growth cones in culture. J Neurosci 18:251–265.

Baulac S, Gourfinkel-An I, Couarch P, Depienne C, Kaminska A, Dulac O, Baulac M, LeGuern E, Nabbout R. 2008. A novel locus for generalized epilepsy with febrile seizures plus in French families. Arch Neurol 65:943–951.

Bebek N, Baykan B, Gürses C, Emir Ö, Gökyiğit A. 2006. Self-induction behavior in patients with photosensitive and hot water epilepsy: A comparative study from a tertiary epilepsy center in Turkey. Epilepsy Behav 9:317–326.

Bebek N, Gürses C, Gokyigit A, Baykan B, Ozkara C, Dervent A. 2001. Hot water epilepsy: Clinical and electrophysiologic findings based on 21 cases. Epilepsia 42:1180–1184.

Bettoni L, Bortone E, Ghizzoni P, Juvarra G. 1986. The sunflower syndrome. Funct Neurol 1:175–181.

Bharath RD, Sinha S, Panda R, Raghavendra K, George L, Chaitanya G, Gupta A, Satishchandra P. 2015. Seizure Frequency Can Alter Brain Connectivity: Evidence from Resting-State fMRI. Am J Neuroradiol 36:1890–1898.

Bhattacharya A, Czaplinski K, Trifillis P, He F, Jacobson A, Peltz SW. 2000. Characterization of the biochemical properties of the human Upf1 gene product that is involved in nonsense-mediated mRNA decay. RNA 6:1226–1235.

Bickford RG, Whelan JL, Klass DW, Corbin KB. 1956. Reading epilepsy: clinical and electroencephalographic studies of a new syndrome. Trans Am Neurol Assoc 100–2.

Biton S, Barzilai A, Shiloh Y. 2008. The neurological phenotype of ataxia-telangiectasia: Solving a persistent puzzle. DNA Repair (Amst) 7:1028–1038.

Bjarnadóttir TK, Fredriksson R, Höglund PJ, Gloriam DE, Lagerström MC, Schiöth HB. 2004. The human and mouse repertoire of the adhesion family of G-protein-coupled receptors. Genomics 84:23–33.

Bliek AM Van Der, Meyerowrtz EM. 1991. Dynamin-like protein encoded by the Drosophila shibire gene associated with vesicular traffic. Nature 351:411–414.

Blume WT, Lüders HO, Mizrahi E, Tassinari C, Emde Boas W Van, Engel J. 2001. Glossary of descriptive terminology for ictal semiology: Report of the ILAE Task Force on classification and terminology. Epilepsia 42:1212–1218.

Bogart K, Andrews J. 2006.

Bolduc F V., Valente D, Nguyen AT, Mitra PP, Tully T. 2010. An assay for social interaction in Drosophila fragile X mutants. Fly (Austin) 4:216–225.

Brannvoll A, Xue X, Kwon Y, Kompocholi S, Simonsen AKW, Viswalingam KS, Gonzalez L, Hickson ID, Oestergaard VH, Mankouri HW, Sung P, Lisby M. 2020. The ZGRF1 Helicase Promotes Recombinational Repair of Replication-Blocking DNA Damage in Human Cells. Cell Rep 32:107849.

Brennan TJ, Seeley WW, Kilgard M, Schreiner CE, Tecott LH. 1997. Sound-induced seizures in serotonin 5-HT(2c) receptor mutant mice. Nat Genet 16:387–390.

Buchan JR, Kolaitis RM, Taylor JP, Parker R. 2013. Eukaryotic stress granules are cleared by autophagy and Cdc48/VCP function. Cell 153:1461–1474.

Byrns CN, Pitts MW, Gilman CA, Hashimoto AC, Berry MJ. 2014. Mice lacking selenoprotein P and selenocysteine lyase exhibit severe neurological dysfunction, neurodegeneration, and audiogenic seizures. J Biol Chem 289:9662–9674.

Cao L, Li L, Zuo Z. 2012. N-Acetylcysteine reverses existing cognitive impairment and increased oxidative stress in glutamate transporter type 3 deficient mice. Neuroscience 220:85.

Chabrol E, Navarro V, Provenzano G, Cohen I, Dinocourt C, Rivaud-Péchoux S, Fricker D, Baulac M, Miles R, Leguern E, Baulac S. 2010. Electroclinical characterization of epileptic seizures in leucine-rich, glioma-inactivated 1-deficient mice. Brain 133:2749–2762.

Chakrabarti S, Jayachandran U, Bonneau F, Fiorini F, Basquin C, Domcke S, Hir H Le, Conti E. 2011. Molecular Mechanisms for the RNA-Dependent ATPase Activity of Upf1 and Its Regulation by Upf2. Mol Cell 41:693–703.

Chen L, Toth M. 2001. Fragile X mice develop sensory hyperreactivity to auditory stimuli. Neuroscience 103:1043–1050.

Cirillo L, Cieren A, Barbieri S, Khong A, Schwager F, Parker R, Gotta M. 2020. UBAP2L Forms Distinct Cores that Act in Nucleating Stress Granules Upstream of G3BP1. Curr Biol 30:698-707.e6.

Darby CE, Korte RA, Binnie CD, Wilkins AJ. 1980. The Self-induction of Epileptic Seizures by Eye Closure. Epilepsia 21:31–42.

Deans B, Griffin CS, Maconochie M, Thacker J. 2000. Xrcc2 is required for genetic stability, embryonic neurogenesis and viability in mice. EMBO J 19:6675.

Devinsky O, Vezzani A, J. O'Brien T, Jette N, Scheffer IE, Curtis M de, Perucca P. 2018. Epilepsy. Nat Rev Dis Prim 4:18024.

Dewey CM, Cenik B, Sephton CF, Dries DR, Mayer P, III, Good SK, Johnson BA, Herz J, Yu G. 2011. TDP-43 Is Directed to Stress Granules by Sorbitol, a Novel Physiological Osmotic and Oxidative Stressor. Mol Cell Biol 31:1098.

Dibbens LM, Tarpey PS, Hynes K, Bayly MA, Scheffer IE, Smith R, Bomar J, Sutton E, Vandeleur L, Shoubridge C, Edkins S, Turner SJ, et al. 2008. X-linked protocadherin 19 mutations cause female-limited epilepsy and cognitive impairment. Nat Genet 40:776–781.

Dictenberg JB, Swanger SA, Antar LN, Singer RH, Bassell GJ. 2008. A Direct Role for

FMRP in Activity-Dependent Dendritic mRNA Transport Links Filopodial-Spine Morphogenesis to Fragile X Syndrome. Dev Cell 14:926–939.

Doll CA, Broadie K. 2016. Neuron class-specific requirements for Fragile X Mental Retardation Protein in critical period development of calcium signaling in learning and memory circuitry. Neurobiol Dis 89:76–87.

Doose H, Petersen B, Neubauer BA. 2002. Occipital sharp waves in idiopathic partial epilepsies - Clinical and genetic aspects. Epilepsy Res 48:121–130.

Doyle M, Kiebler MA. 2011. Mechanisms of dendritic mRNA transport and its role in synaptic tagging. EMBO J 30:3540–3552.

Dravet C, Bureau M, Guerrini R, Giraud N, Roger J. 1992. Severe myoclonic epilepsy in infants. In Roger J, Dravet C, Bureau M, Dreifuss FE, Perret A, Wolf P, Editors. Epileptic Syndromes in Infancy, Childhood and Adolescence., p 75–88.

Drozd M, Bardoni B, Capovilla M. 2018. Modeling fragile X syndrome in drosophila. Front Mol Neurosci 11:124.

Dumitrache LC, McKinnon PJ. 2017. Polynucleotide kinase-phosphatase (PNKP) mutations and neurologic disease. Mech Ageing Dev 161:121–129.

Ellis L. 2017. The potential mechanism of musicogenic epilepsy and future research avenues. Biosci Horizons Int J Student Res 10:1–10.

Engel J. 2001. A proposed diagnostic scheme for people with epileptic seizures and with epilepsy: Report of the ILAE task force on classification and terminology. Epilepsia 42:796–803.

Eom T, Zhang C, Wang H, Lay K, Fak J, Noebels JL, Darnell RB. 2013. NOVA-dependent regulation of cryptic NMD exons controls synaptic protein levels after seizure. Elife 2:e00178.

Etholm L, Bahonjic E, Heggelund P. 2013. Sensitive and critical periods in the development of handling induced seizures in mice lacking synapsins: Differences between synapsin I and synapsin II knockouts. Exp Neurol 247:59–65.

Eulalio A, Behm-Ansmant I, Schweizer D, Izaurralde E. 2007. P-Body Formation Is a Consequence, Not the Cause, of RNA-Mediated Gene Silencing. Mol Cell Biol 27:3970–3981.

Ewels P, Ns Magnusson M, Lundin S, Aller MK. 2016. Data and text mining MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32:3047–3048.

Faingold CL, Riaz A. 1995. Ethanol withdrawal induces increased firing in inferior colliculus neurons associated with audiogenic seizure susceptibility. Exp Neurol 132:91–98.

Farny NG, Kedersha NL, Silver PA. 2009. Metazoan stress granule assembly is mediated by P-eIF2α-dependent and -independent mechanisms. RNA 15:1814–1821.

Fassio A, Patry L, Congia S, Onofri F, Piton A, Gauthier J, Pozzi D, Messa M, Defranchi

E, Fadda M, Corradi A, Baldelli P, et al. 2011a. SYN1 loss-of-function mutations in autism and partial epilepsy cause impaired synaptic function. Hum Mol Genet 20:2297–2307.

Fassio A, Patry L, Congia S, Onofri F, Piton A, Gauthier J, Pozzi D, Messa M, Defranchi E, Fadda M, Corradi A, Baldelli P, et al. 2011b. SYN1 loss-of-function mutations in autism and partial epilepsy cause impaired synaptic function. Hum Mol Genet 20:2297–2307.

Fenger-Grøn M, Fillman C, Norrild B, Lykke-Andersen J. 2005. Multiple processing body factors and the ARE binding protein TTP activate mRNA decapping. Mol Cell 20:905–915.

Fenwick PB. 1998. Self-generation of seizures by an action of mind. Adv Neurol 75:87–92.

Ferlazzo E, Adjien CK, Guerrini R, Calarese T, Crespel A, Elia M, Striano P, Gelisse P, Bramanti P, Bella P di, Genton P. 2009. Lennox-Gastaut syndrome with late-onset and prominent reflex seizures in trisomy 21 patients. Epilepsia 50:1587–1595.

Fischer JW, Busa VF, Shao Y, Leung AKL. 2020. Structure-Mediated RNA Decay by UPF1 and G3BP1. Mol Cell 78:70-84.e6.

Fisher RS, Acevedo C, Arzimanoglou A, Bogacz A, Cross JH, Elger CE, Engel J, Forsgren L, French JA, Glynn M, Hesdorffer DC, Lee BI, et al. 2014. ILAE Official Report: A practical clinical definition of epilepsy. Epilepsia 55:475–482.

Fisher RS, Harding G, Erba G, Barkley GL, Wilkins A. 2005. Photic- and pattern-induced seizures: A review for the epilepsy foundation of america working group. Epilepsia 46:1426–1441.

Fogel BL, Wexler E, Wahnich A, Friedrich T, Vijayendran C, Gao F, Parikshak N, Konopka G, Geschwind DH. 2012. RBFOX1 regulates both splicing and transcriptional networks in human neuronal development. Hum Mol Genet 21:4171–4186.

Follwaczny P, Schieweck R, Riedemann T, Demleitner A, Straub T, Klemm AH, Bilban M, Sutor B, Popper B, Kiebler MA. 2017. Pumilio2-deficient mice show a predisposition for epilepsy. DMM Dis Model Mech 10:1333–1342.

Gachon F, Fonjallaz P, Damiola F, Gos P, Kodama T, Zakany J, Duboule D, Petit B, Tafti M, Schibler U. 2004. The loss of circadian PAR bZip transcription factors results in epilepsy. Genes Dev 18:1397–1412.

Galizia EC, Myers CT, Leu C, Kovel CGF De, Afrikanova T, Cordero-Maldonado ML, Martins TG, Jacmin M, Drury S, Chinthapalli VK, Muhle H, Pendziwiat M, et al. 2015. CHD2 variants are a risk factor for photosensitivity in epilepsy. Brain 138:1198–1207.

Garcia-Cairasco N. 2002. A critical review on the participation of inferior colliculus in acoustic-motor and acoustic-limbic networks involved in the expression of acute and kindled audiogenic seizures. Hear Res 168:208–222.

Gavaret M, Guedj E, Koessler L, Fonseca AT Da, Aubert S, Mundler O, Chauvel P, Bartolomei F. 2010. Reading epilepsy from the dominant temporo-occipital region. J Neurol Neurosurg Psychiatry 81:710–715.

Gilks N, Kedersha N, Ayodele M, Shen L, Stoecklin G, Dember LM, Anderson P. 2004. Stress granule assembly is mediated by prion-like aggregation of TIA-1. Mol Biol Cell 15:5383–5398.

Glasscock E, Tanouye MA. 2005. Drosophila couch potato mutants exhibit complex neurological abnormalities including epilepsy phenotypes. Genetics 169:2137–2149.

Glazova M V., Nikitina LS, Hudik KA, Kirillova OD, Dorofeeva NA, Korotkov AA, Chernigovskaya E V. 2015. Inhibition of ERK1/2 signaling prevents epileptiform behavior in rats prone to audiogenic seizures. J Neurochem 132:218–229.

Gleeson JG, Allen KM, Fox JW, Lamperti ED, Berkovic S, Scheffer I, Cooper EC, Dobyns WB, Minnerath SR, Ross ME, Walsh CA. 1998. Doublecortin, a brain-specific gene mutated in human X-linked lissencephaly and double cortex syndrome, encodes a putative signaling protein. Cell 92:63–72.

Grosso S, Farnetani MA, Francione S, Galluzzi P, Vatti G, Cordelli DM, Morgese G, Balestri P. 2004. Hot water epilepsy and focal malformation of the parietal cortex development. Brain Dev 26:490–493.

Guerrini R, Genton P, Bureau M, Dravet C, Roger J. 1990. Reflex Seizures are Frequent in Patients with Down Syndrome and Epilepsy. Epilepsia 31:406–417.

Guerrini R, Genton P, Dravet C, Viallat D, Bureau M, Horton EJ, Roger J. 1992. Compulsive Somatosensory Self-Stimulation Inducing Epileptic Seizures. Epilepsia 33:509–516.

Guo Z, Tian Y, Guo Y, Li B, Liu X, Xie K, Song Y, Wang D. 2019. RAD6B Plays a Critical Role in Neuronal DNA Damage Response to Resist Neurodegeneration. Front Cell Neurosci 13:392.

Hakem R, la Pompa J de, Sirard C, Mo R, Woo M, Hakem A, Wakeham A, Potter J, Reitmair A, Billia F, Firpo E, Hui C, et al. 1996. The tumor suppressor gene Brca1 is required for embryonic cellular proliferation in the mouse. Cell 85:1009–1023.

Harding GFA, Herrick CE, Jeavons PM. 1978. A Controlled Study of the Effect of Sodium Valproate on Photosensitive Epilepsy and Its Prognosis. Epilepsia 19:555–565.

Hauser WA, Hesdorffer DC. 1990. Epilepsy : Frequency, Causes and Consequences. 378 p.

Hedera P, Blair MA, Andermann E, Andermann F, D'Agostino D, Taylor KA, Chahine L, Pandolfo M, Bradford Y, Haines JL, Abou-Khalil B. 2007. Familial mesial temporal lobe epilepsy maps to chromosome 4q13.2-q21.3. Neurology 68:2107–2112.

Hengst U, Jaffrey SR. 2007. Function and translational regulation of mRNA in developing axons. Semin Cell Dev Biol 18:209–215.

Hennig S, Kong G, Mannen T, Sadowska A, Kobelke S, Blythe A, Knott GJ, Iyer SS, Ho D, Newcombe EA, Hosoki K, Goshima N, et al. 2015. Prion-like domains in RNA binding proteins are essential for building subnuclear paraspeckles. J Cell Biol 210:529–539.

Hibar DP, Stein JL, Renteria ME, Arias-Vasquez A, Desrivières S, Jahanshad N, Toro R,

Wittfeld K, Abramovic L, Andersson M, Aribisala BS, Armstrong NJ, et al. 2015. Common genetic variants influence human subcortical brain structures. Nature 520:224.

Hill AS, Jain P, Folan NE, Ben-Shahar Y. 2019. The Drosophila ERG channel seizure plays a role in the neuronal homeostatic stress response. PLOS Genet 15:e1008288.

Hirsch-Reinshagen V, Pottier C, Nicholson AM, Baker M, Hsiung GYR, Krieger C, Sengdy P, Boylan KB, Dickson DW, Mesulam M, Weintraub S, Bigio E, et al. 2017. Clinical and neuropathological features of ALS/FTD with TIA1 mutations. Acta Neuropathol Commun 5:96.

Hom AC, Leppik IE, Rask CA. 1993. Effects of estradiol and progesterone on seizure sensitivity in oophorectomized DBA/2J mice and C57/EL hybrid mice. Neurology 43:198–204.

Hsu LC, White RL. 1998. BRCA1 is associated with the centrosome during mitosis. Proc Natl Acad Sci U S A 95:12983–12988.

Hua Y, Zhou J. 2004. Survival motor neuron protein facilitates assembly of stress granules. FEBS Lett 572:69–74.

Huber KM, Kayser MS, Bear MF. 2000. Role for rapid dendritic protein synthesis in hippocampal mGluR- dependent long-term depression. Science (80-) 288:1254–1256.

Hughes JR. 2008. The photoparoxysmal response: The probable cause of attacks during video games. Clin EEG Neurosci 39:1–7.

Illingworth JL, Ring H. 2013. Conceptual distinctions between reflex and nonreflex precipitated seizures in the epilepsies: A systematic review of definitions employed in the research literature. Epilepsia 54:2036–2047.

Ingelfinger D, Arndt-Jovin DJ, Lührmann R, Achsel T. 2002. The human LSm1-7 proteins colocalize with the mRNA-degrading enzymes Dcp1/2 and Xrn1 in distinct cytoplasmic foci. RNA 8:1489–1501.

Ingvar DH, Nyman GE. 1962. Epilepsia arithmetices: A new psychologic trigger mechanism in a case of epilepsy. Neurology 12:282–287.

Irmen F, Wehner T, Lemieux L. 2015. Do reflex seizures and spontaneous seizures form a continuum? – Triggering factors and possible common mechanisms. Seizure 25:72–79.

Italiano D, Striano P, Russo E, Leo A, Spina E, Zara F, Striano S, Gambardella A, Labate A, Gasparini S, Lamberti M, Sarro G De, et al. 2016. Genetics of reflex seizures and epilepsies in humans and animals. Epilepsy Res 121:47–54.

Iyer LM, Leipe DD, Koonin E V., Aravind L. 2004. Evolutionary history and higher order classification of AAA+ ATPases. J Struct Biol 146:11–31.

Jain S, Parker R. 2013. The Discovery and Analysis of P Bodies. Advances in Experimental Medicine and Biology, Adv Exp Med Biol, p 23–43.

Jain S, Wheeler JR, Walters RW, Agrawal A, Barsic A, Parker R. 2016. ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. Cell 164:487–498.

Jevtov I, Zacharogianni M, Oorschot MM van, Zadelhoff G van, Aguilera-Gomez A, Vuillez I, Braakman I, Hafen E, Stocker H, Rabouille C. 2015. TORC2 mediates the heat stress response in Drosophila by promoting the formation of stress granules. J Cell Sci 128:2497–2508.

Jiang W, Duong TM, Lanerolle NC De. 1999. The neuropathology of hyperthermic seizures in the rat. Epilepsia 40:5–19.

Jin Z, Tietjen I, Bu L, Liu-Yesucevitz L, Gaur SK, Walsh CA, Piao X. 2007. Disease-associated mutations affect GPR56 protein trafficking and cell surface expression. Hum Mol Genet 16:1972–1985.

Jonas S, Izaurralde E. 2013. The role of disordered protein regions in the assembly of decapping complexes and RNP granules. Genes Dev 27:2628–2641.

Kaech S, Banker G. 2006. Culturing hippocampal neurons. Nat Protoc 1:2406–2415.

Kalsi G, Kuo PH, Aliev F, Alexander J, McMichael O, Patterson DG, Walsh D, Zhao Z, Schuckit M, Nurnberger JJ, Edenberg H, Kramer J, et al. 2010. A systematic gene-based screen of chr4q22-q32 identifies association of a novel susceptibility gene, DKK2, with the quantitative trait of alcohol dependence symptom counts. Hum Mol Genet 19:2497–2506.

Kamel JT, Badawy RAB, Cook MJ. 2014. Exercise-induced seizures and lateral asymmetry in patients with temporal lobe epilepsy. Epilepsy Behav Case Reports 2:26–30.

Kanai Y, Bhide PG, DiFiglia M, Hediger MA. 1995. Neuronal high-affinity glutamate transport in the rat central nervous system. Neuroreport 6:2357–2362.

Kandel ER, Dudai Y, Mayford MR. 2014. The molecular and systems biology of memory. Cell 157:163–186.

Kanemoto K, Watanabe Y, Tsuji T, Fukami M, Kawasaki J. 2001. Rub epilepsy: A somatosensory evoked reflex epilepsy induced by prolonged cutaneous stimulation. J Neurol Neurosurg Psychiatry 70:541–543.

Kang H, Schuman EM. 1996. A requirement for local protein synthesis in neurotrophin-induced hippocampal synaptic plasticity. Science (80-) 273:1402–1406.

Karan KR, Satishchandra P, Sinha S, Anand A. 2017. Rare SLC1A1 variants in hot water epilepsy. Hum Genet 136:693–703.

Karan KR, Satishchandra P, Sinha S, Anand A. 2018. A genetic locus for sensory epilepsy precipitated by contact with hot water maps to chromosome 9p24.3-p23. J Genet 97:391–398.

Kashima I, Jonas S, Jayachandran U, Buchwald G, Conti E, Lupas AN, Izaurralde E. 2010. SMG6 interacts with the exon junction complex via two conserved EJC-binding motifs (EBMs) required for nonsense-mediated mRNA decay. Genes Dev 24:2440–2450.

Kashima R, Redmond PL, Ghatpande P, Roy S, Kornberg TB, Hanke T, Knapp S, Lagna G, Hata A. 2017. Hyperactive locomotion in a Drosophila model is a functional readout

for the synaptic abnormalities underlying fragile X syndrome. Sci Signal 10:8133.

Kawai H, Allende ML, Wada R, Kono M, Sango K, Deng C, Miyakawa T, Crawley JN, Werth N, Bierfreund U, Sandhoff K, Proia RL. 2001. Mice Expressing Only Monosialoganglioside GM3 Exhibit Lethal Audiogenic Seizures. J Biol Chem 276:6885–6888.

Kawasaki F, Felling R, Ordway RW. 2000. A temperature-sensitive paralytic mutant defines a primary synaptic calcium channel in Drosophila. J Neurosci 20:4885–4889.

Kedersha N, Anderson P. 2002. Stress granules: Sites of mRNA triage that regulate mRNA stability and translatability. Biochem Soc Trans 30:963–969.

Kedersha N, Anderson P. 2007. Mammalian Stress Granules and Processing Bodies. Methods Enzymol 431:61–81.

Kedersha N, Stoecklin G, Ayodele M, Yacono P, Lykke-Andersen J, Fitzler MJ, Scheuner D, Kaufman RJ, Golan DE, Anderson P. 2005. Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. J Cell Biol 169:871–884.

Keipert JA. 1969. Epilepsy precipitated by bathing: Water-immersion epilepsy. J Paediatr Child Health 5:244–7.

Kenan DJ, Query CC, Keene JD. 1991. RNA recognition: towards identifying determinants of specificity. Trends Biochem Sci 16:214–220.

Kerjan G, Gleeson JG. 2007. Genetic mechanisms underlying abnormal neuronal migration in classical lissencephaly. Trends Genet 23:623–630.

Khaikin Y, Mercimek-Mahmutoglu S. 1993. STXBP1 Encephalopathy with Epilepsy. GeneReviews®, University of Washington, Seattle,.

Khalfallah Y, Kuta R, Grasmuck C, Prat A, Durham HD, Velde C Vande. 2018. TDP-43 regulation of stress granule dynamics in neurodegenerative disease-relevant cell types. Sci Rep 8:1–13.

Kiebler MA, Bassell GJ. 2006. Neuronal RNA Granules: Movers and Makers. Neuron 51:685–690.

Kim JH, Kim HD, Ryu GH, Kim DH, Hurwitz J, Seo YS. 2006. Isolation of human Dna2 endonuclease and characterization of its enzymatic properties. Nucleic Acids Res 34:1854–1864.

Klauenberg BJ, Sparber SB. 1984. A Kindling-like Effect Induced by Repeated Exposure to Heated Water in Rats. 25:292–301.

Koenig JH, Ikeda K. 1989. Disappearance and reformation of synaptic vesicle membrane upon transmitter release observed under reversible blockage of membrane retrieval. J Neurosci 9:3844–3860.

Koepp MJ, Caciagli L, Pressler RM, Lehnertz K, Beniczky S. 2016. Reflex seizures, traits, and epilepsies: From physiology to pathology. Lancet Neurol 15:92–105.

Köhrmann M, Luo M, Kaether C, DesGroseillers L, Dotti CG, Kiebler MA. 1999.

Microtubule-dependent recruitment of Staufen-green fluorescent protein into large RNA-containing granules and subsequent dendritic transport in living hippocampal neurons. Mol Biol Cell 10:2945–2953.

Kosturko LD, Maggipinto MJ, D'Sa C, Carson JH, Barbarese E. 2005. The microtubule-associated protein tumor overexpressed gene binds to the RNA trafficking protein heterogeneous nuclear ribonucleoprotein A2. Mol Biol Cell 16:1938–1947.

Kovel CGF de, Pinto D, Tauer U, Lorenz S, Muhle H, Leu C, Neubauer BA, Hempelmann A, Callenbach PMC, Scheffer IE, Berkovic SF, Rudolf G, et al. 2010a. Whole-genome linkage scan for epilepsy-related photosensitivity: A mega-analysis. Epilepsy Res 89:286–294.

Kovel CGF de, Pinto D, Tauer U, Lorenz S, Muhle H, Leu C, Neubauer BA, Hempelmann A, Callenbach PMC, Scheffer IE, Berkovic SF, Rudolf G, et al. 2010b. Whole-genome linkage scan for epilepsy-related photosensitivity: A mega-analysis. Epilepsy Res 89:286–294.

Krushinsky L V., Molodkina LN, Fless DA, Dobrokhotova LP, Steshenko AP, Semiokhina AF, Zorina ZA, Romanova LG. 1970. The Functional State of the Brain during Sonic Stimulation. Physiological Effects of Noise, Springer US, p 159–183.

Kurata S. 1979. Epilepsy Precipitated by Bathing A Follow-up Study. No To Hattatsu 11:400–405.

Kurosaki T, Maquat LE. 2016. Nonsense-mediated mRNA decay in humans at a glance. J Cell Sci 129:461–467.

Kuznetsov SG, Haines DC, Martin BK, Sharan SK. 2009. Loss of Rad51c leads to embryonic lethality and modulation of Trp53-dependent tumorigenesis in mice. Cancer Res 69:863.

Lal D, Trucks H, Møller RS, Hjalgrim H, Koeleman BPC, Kovel CGF de, Visscher F, Weber YG, Lerche H, Becker F, Schankin CJ, Neubauer BA, et al. 2013. Rare exonic deletions of the RBFOX1 gene increase risk of idiopathic generalized epilepsy. Epilepsia 54:265–271.

Lee C, Hurwitz J. 1993. Human RNA helicase A is homologous to the maleless protein of Drosophila. J Biol Chem 268:16822–30.

Lee J, Wu CF. 2002. Electroconvulsive seizure behavior in Drosophila: Analysis of the physiological repertoire underlying a stereotyped action pattern in bang-sensitive mutants. J Neurosci 22:11065–11079.

Lee S, Liu B, Lee S, Huang SX, Shen B, Qian SB. 2012a. Global mapping of translation initiation sites in mammalian cells at single-nucleotide resolution. Proc Natl Acad Sci U S A 109:E2424–E2432.

Lee S, Park S-H, Zuo Z. 2012b. Effects of isoflurane on learning and memory functions of wild-type and glutamate transporter type 3 knockout mice. J Pharm Pharmacol 64:302–307.

Lee VLL, Choo BKM, Chung YS, Kundap UP, Kumari Y, Shaikh MF. 2018. Treatment, therapy and management of metabolic epilepsy: A systematic review. Int J Mol Sci

19:871.

Lee YC, Yen DJ, Liring JF, Yiu CH. 2000. Epileptic seizures in a patient by immersing his right hand into hot water. Seizure 9:605–607.

Lenoir P, Ramet J, Meirleir L De, D'Allest AM, Desprechins B, Loeb H. 1989. Bathing-induced seizures. Pediatr Neurol 5:124–125.

Leviton A, Cowan LD. 1982. Epidemiology of Seizure Disorders in Children; pp. 62–83. Neuroepidemiology 1:62–83.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25:1754–1760.

Li W, Su X, Chen T, Li Z, Yang Y, Zhang L, Liu Q, Shao M, Zhang Y, Ding M, Lu Y, Yu H, et al. 2020. Solute Carrier Family 1 (SLC1A1) Contributes to Susceptibility and Psychopathology Symptoms of Schizophrenia in the Han Chinese Population. Front Psychiatry 0:989.

Li YR, King OD, Shorter J, Gitler AD. 2013. Stress granules as crucibles of ALS pathogenesis. J Cell Biol 201:361–372.

Liepinsh E, Leonchiks A, Sharipo A, Guignard L, Otting G. 2003. Solution structure of the R3H domain from human Sμbp-2. J Mol Biol 326:217–223.

Lindh AR, Rafii S, Schultz N, Cox A, Helleday T. 2006. Mitotic defects in XRCC3 variants T241M and D213N and their relation to cancer susceptibility. Hum Mol Genet 15:1217–1224.

Liu-Yesucevitz L, Bilgutay A, Zhang Y-J, Vanderwyde T, Citro A, Mehta T, Zaarur N, McKee A, Bowser R, Sherman M, Petrucelli L, Wolozin B. 2010. Tar DNA Binding Protein-43 (TDP-43) Associates with Stress Granules: Analysis of Cultured Cells and Pathological Brain Tissue. PLoS One 5:e13250.

Lopes Da Silva F, Blanes W, Kalitzin SN, Parra J, Suffczynski P, Velis DN. 2003. Epilepsies as Dynamical Diseases of Brain Systems: Basic Models of the Transition between Normal and Epileptic Activity. Epilepsia 44:72–83.

Loreto V, Nocerino C, Striano P, D'Aulos F, Boccella P, Striano S. 2000. Eating epilepsy. Heterogeneity of ictal semiology: The role of video-EEG monitoring. Epileptic Disord 2:93–98.

Loughney K, Kreber R, Ganetzky B. 1989. Molecular analysis of the para locus, a sodium channel gene in Drosophila. Cell 58:1143–1154.

Luca G De, Giorgio RM Di, Macaione S, Calpona PR, Costantino S, Paola ED Di, Sarro A De, Ciliberto G, Sarro G De. 2004. Susceptibility to audiogenic seizure and neurotransmitter amino acid levels in different brain areas of IL-6-deficient mice. Pharmacol Biochem Behav 78:75–81.

Luo R, Yang HM, Jin Z, Halley DJJ, Chang BS, MacPherson L, Brueton L, Piao X. 2011. A novel GPR56 mutation causes bilateral frontoparietal polymicrogyria. Pediatr Neurol 45:49–53.

Madabhushi R, Pan L, Tsai LH. 2014. DNA damage and its links to neurodegeneration. Neuron 83:266–282.

Mahboubi H, Seganathy E, Kong D, Stochaj U. 2013. Identification of Novel Stress Granule Components That Are Involved in Nuclear Transport. PLoS One 8:e68356.

Mahboubi H, Stochaj U. 2017. Cytoplasmic stress granules: Dynamic modulators of cell signaling and disease. Biochim Biophys Acta - Mol Basis Dis 1863:884–895.

Mani K, Mani A, Kamesh C, Ahuja G. 1972. Hot-water epi lepsy. clinical and electroencephalographic features--study of 60 cases. Neurol India (Suppl II) 237–240.

Mani K, Mani A, Ramesh C. 1974a. Hot-water epilepsy-a peculiar type of reflex epilepsy: clinical and EEG features in 108 cases. Trans Am Neurol Assoc 24–226.

Mani KS, Gopalakrishnan PN, Vyas JN, Pillai MS. 1968. “Hot-water epilepsy”--a peculiar type of reflex epilepsy. A preliminary report. Neurol India 16:107–10.

Mani KS, Mani AJ, Ramesh CK. 1974b. Hot-water epilepsy--a peculiar type of reflex epilepsy: clinical and EEG features in 108 cases. Trans Am Neurol Assoc 99:224–6.

Maria SRS, Kwon YH, Sung P, Klein HL. 2013. Characterization of the interaction between the saccharomyces cerevisiae Rad51 recombinase and the DNA translocase Rdh54. J Biol Chem 288:21999–22005.

Martin KC, Zukin RS. 2006. RNA trafficking and local protein synthesis in dendrites: an overview. J Neurosci 26:7131–7134.

Martin S, Zekri L, Metz A, Maurice T, Chebli K, Vignes M, Tazi J. 2013. Deficiency of G3BP1, the stress granules assembly factor, results in abnormal synaptic plasticity and calcium homeostasis in neurons. J Neurochem 125:175–184.

Martínez AR, Colmenero MIA, Pereira AG, Vilaplana FXS, Morón JA, Marfa MP. 2011. Reflex seizures in Rett syndrome. Epileptic Disord 13:389–393.

Masson JY, Tarsounas MC, Stasiak AZ, Stasiak A, Shah R, McIlwraith MJ, Benson FE, West SC. 2001. Identification and purification of two distinct complexes containing the five RAD51 paralogs. Genes Dev 15:3296–3307.

Masuda-Sasa T, Imamura O, Campbell JL. 2006. Biochemical analysis of human Dna2. Nucleic Acids Res 34:1865–1875.

Matricardi M, Brinciotti M, Trasatti G, Porro G. 1990. Self-induced pattern-sensitive epilepsy in childhood. Acta Paediatr Scand 79:237–240.

Matsuki H, Takahashi M, Higuchi M, Makokha GN, Oie M, Fujii M. 2013. Both G3BP1 and G3BP2 contribute to stress granule formation. Genes to Cells 18:135–146.

Matthews W, Wright F. 1967. Hereditary primary reading epilepsy. Neurology 17:919–921.

Mazroui R, Hout ME, Tremblay S, Fillion C, Labelle Y, Khandjian EW. 2002. Trapping of messenger RNA by Fragile X Mental Retardation protein into cytoplasmic granules induces translation repression. Hum Mol Genet 11:3007–3017.

Mazroui R, Marco S Di, Kaufman RJ, Gallouzi I-E. 2007. Inhibition of the ubiquitin-proteasome system induces stress granule formation. Mol Biol Cell 18:2603–2618.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. 2010. The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20:1297–1303.

McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, Flicek P, Cunningham F. 2016. The Ensembl Variant Effect Predictor. Genome Biol 17:122.

McNamara JO, Puranam RS. 1998. Epilepsy genetics: An abundance of riches for biologists. Curr Biol 8:R168–R170.

Meekings C, O'Brien L. 2004. Borderline pathology in children and adolescents. Int J Ment Health Nurs 13:152–163.

Meghana A, Sinha S, Sathyaprabha TN, Subbakrishna DK, Satishchandra P. 2012. Hot water epilepsy clinical profile and treatment—A prospective study. Epilepsy Res 102:160–166.

Mervaala E, Andermann F, Quesney LF, Krelina M. 1990. Common dopaminergic mechanism for epileptic photosensitivity in progressive myoclonus epilepsies. Neurology 40:53–56.

Miyao M, Tezuka M, Kuwajima K, Kamoshita S. 1982. Epilepsy induced by hot water immersion. Brain Dev 4:158.

Mofenson HC, Weymuller CA, Greensher J. 1965. Epilepsy Due to Water Immersion: An Unusual Case of Reflex Sensory Epilepsy. JAMA J Am Med Assoc 191:600–601.

Mooney CM, Jimenez-Mateos EM, Engel T, Mooney C, Diviney M, Venø MT, Kjems J, Farrell MA, O'Brien DF, Delanty N, Henshall DC. 2017. RNA sequencing of synaptic and cytoplasmic Upf1-bound transcripts supports contribution of nonsense-mediated decay to epileptogenesis. Sci Rep 7:1–14.

Moran J. 1976. So-called water immersion epilepsy. Ir J Med Sci 145–140.

Morimoto T, Hayakawa T, Sugie H, Awaya Y, Fukuyama Y. 1985. Epileptic Seizures Precipitated by Constant Light, Movement in Daily Life, and Hot Water Immersion. Epilepsia 26:237–242.

Motelow JE, Povysil G, Dhindsa RS, Stanley KE, Allen AS, Feng YCA, Howrigan DP, Abbott LE, Tashman K, Cerrato F, Cusick C, Singh T, et al. 2021. Sub-genic intolerance, ClinVar, and the epilepsies: A whole-exome sequencing study of 29,165 individuals. Am J Hum Genet 108:965–982.

Myers KA, Johnstone DL, Dyment DA. 2019. Epilepsy genetics: Current knowledge, applications, and future directions. Clin Genet 95:95–111.

Myles-Worsley M, Tiobech J, Browning S, Korn J, Goodman S, Gentile K, Melhem N, Byerley W, Faraone S, Middleton F. 2013. Deletion at the SLC1A1 glutamate transporter gene co-segregates with schizophrenia and bipolar schizoaffective disorder in a 5-generation family. Am J Med Genet B Neuropsychiatr Genet 162B:87–95.

N'Gouemo P, Faingold CL. 1999. The periaqueductal grey is a critical site in the neuronal network for audiogenic seizures: Modulation by GABA(A), NMDA and opioid receptors. Epilepsy Res 35:39–46.

Need AC, McEvoy JP, Gennarelli M, Heinzen EL, Ge D, Maia JM, Shianna K V., He M, Cirulli ET, Gumbs CE, Zhao Q, Campbell CR, et al. 2012. Exome sequencing followed by large-scale genotyping suggests a limited role for moderately rare risk factors of strong effect in schizophrenia. Am J Hum Genet 91:303–312.

Ng BY. 2002. Psychiatric aspects of self-induced epileptic seizures. Aust N Z J Psychiatry 36:534–543.

Nguyen DK, Rouleau I, Sénéchal G, Ansaldo AI, Gravel M, Benfenati F, Cossette P. 2015. X-linked focal epilepsy with reflex bathing seizures: Characterization of a distinct epileptic syndrome. Epilepsia 56:1098–1108.

Nigro C Lo, Chong SS, Smith ACM, Dobyns WB, Carrozzo R, Ledbetter DH. 1997. Point mutations and an intragenic deletion in LIS1, the lissencephaly causative gene in isolated lissencephaly sequence and Miller-Dieker syndrome. Hum Mol Genet 6:157–164.

Nijs L De, Léon C, Nguyen L, Loturco JJ, Delgado-Escueta A V., Grisar T, Lakaye B. 2009. EFHC1 interacts with microtubules to regulate cell division and cortical development. Nat Neurosci 12:1266–1274.

O'Driscoll M, Cerosaletti KM, Girard PM, Dai Y, Stumm M, Kysela B, Hirsch B, Gennery A, Palmer SE, Seidel J, Gatti RA, Varon R, et al. 2001. DNA ligase IV mutations identified in patients exhibiting developmental delay and immunodeficiency. Mol Cell 8:1175–1185.

Okudan ZV, Özkara Ç. 2018. Reflex epilepsy: triggers and management strategies. Neuropsychiatr Dis Treat 14:327.

Ottman R, Winawer MR, Kalachikov S, Barker-Cummings C, Gilliam TC, Pedley TA, Hauser WA. 2004. LGI1 mutations in autosomal dominant partial epilepsy with auditory features. Neurology 62:1120–1126.

Ouyang S, Song Y, Tian Y, Chen Y, Yu X, Wang D. 2015. RNF8 deficiency results in neurodegeneration in mice. Neurobiol Aging 36:2850–2860.

Palma L De, Boniver C, Cassina M, Toldo I, Nosadini M, Clementi M, Sartori S. 2012. Eating-induced epileptic spasms in a boy with MECP2 duplication syndrome: Insights into pathogenesis of genetic epilepsies. Epileptic Disord 14:414–417.

Panayiotopoulos C. 2005. Reflex Seizures and Reflex Epilepsies. The Epilepsies: Seizures, Syndromes and Management, Oxfordshire (UK): Bladon Medical Publishing,.

Panayiotopoulos CP. 2008. Typical absence seizures and related epileptic syndromes: Assessment of current state and directions for future research. Epilepsia 49:2131–2139.

Pao GM, Zhu Q, Perez-Garcia CG, Chou S-J, Suh H, Gage FH, O'Leary DDM, Verma IM. 2014. Role of BRCA1 in brain development. Proc Natl Acad Sci 111:E1240–E1248.

Parker L, Howlett IC, Rusan ZM, Tanouye MA. 2011. Seizure and epilepsy: Studies of

seizure disorders in drosophila. International Review of Neurobiology, Academic Press Inc., p 1–21.

Parker R, Sheth U. 2007. P Bodies and the Control of mRNA Translation and Degradation. Mol Cell 25:635–646.

Patel, Satishchandra P, Aravinda H, Bharath RD, Sinha S. 2014. Hot water epilepsy: Phenotype and single photon emission computed tomography observations. Ann Indian Acad Neurol 17:470.

Pavlidis P, Tanouye MA. 1995. Seizures and failures in the giant fiber pathway of Drosophila bang- sensitive paralytic mutants. J Neurosci 15:5810–5819.

Peter B, Wijsman EM, Nato AQ, University of Washington Center for Mendelian Genomics, Matsushita MM, Chapman KL, Stanaway IB, Wolff J, Oda K, Gabo VB, Raskind WH, Bamshad M, et al. 2016. Genetic candidate variants in two multigenerational families with childhood apraxia of speech. PLoS One 11:e0153864.

Petsalaki E, Zachos G. 2020. DNA damage response proteins regulating mitotic cell division: double agents preserving genome stability. FEBS J 287:1700–1721.

Pinto D, Westland B, Haan GJ de, Rudolf G, Silva BM da, Hirsch E, Lindhout D, Kasteleijn-Nolst Trenité DGA, Koeleman BPC. 2005. Genome-wide linkage scan in epilepsy-related photoparoxysmal electroencephalographic response: Evidence for linkage on chromosomes 7q32 and 16p13. Hum Mol Genet 14:171–178.

Pittman DL, Schimenti JC. 2000. Midgestation Lethality in Mice Deficient for the RecA-Related Gene, Rad51d/Rad51l3. Genesis 26:167–73.

Poduri A, Lowenstein D. 2011. Epilepsy genetics-past, present, and future. Curr Opin Genet Dev 21:325–332.

Protter DSW, Parker R. 2016. Principles and Properties of Stress Granules. Trends Cell Biol 26:668–679.

Qiu LF, Hao YH, Li QZ, Xiong ZQ. 2008. Fragile X syndrome and epilepsy. Neurosci Bull 24:338–344.

Quek AML, Britton JW, McKeon A, So E, Lennon VA, Shin C, Klein CJ, Watson RE, Kotsenas AL, Lagerlund TD, Cascino GD, Worrell GA, et al. 2012. Autoimmune epilepsy: Clinical characteristics and response to immunotherapy. Arch Neurol 69:582–593.

Quirk JA, Fish DR, Smith SJM, Sander JWAS, Shorvon SD, Allen PJ. 1995. Incidence of photosensitive epilepsy: a prospective national study. Electroencephalogr Clin Neurophysiol 95:260–267.

Radhakrishnan K, Silbert PL, Klass DW. 1995. Reading epilepsyAn appraisal of 20 patients diagnosed at the Mayo Clinic, Rochester, Minnesota, between 1949 and 1989, and delineation of the epileptic syndrome. Brain 118:75–89.

Raju PK, Satishchandra P, Nayak S, Iyer V, Sinha S, Anand A. 2017. Microtubule-associated defects caused by *EFHC1* mutations in juvenile myoclonic epilepsy. Hum Mutat 38:816–826.

Ramaswami M, Tanouye MA. 1989. Two sodium-channel genes in Drosophila: Implications for channel diversity. Proc Natl Acad Sci U S A 86:2079–2082.

Ramaswami M, Taylor JP, Parker R. 2013. Altered ribostasis: RNA-protein granules in degenerative disorders. Cell 154:727–736.

Raschle M, Komen S Van, Chi P, Ellenberger T, Sung P. 2004. Multiple interactions with the rad51 recombinase govern the homologous recombination function of Rad54. J Biol Chem 279:51973–51980.

Ratnapriya R, Satishchandra P, Dilip S, Gadre G, Anand A. 2009a. Familial autosomal dominant refex epilepsy triggered by hot water maps to 4q24-q28. Hum Genet 126:677–683.

Ratnapriya R, Satishchandra P, Kumar SD, Gadre G, Reddy R, Anand A. 2009b. A locus for autosomal dominant reflex epilepsy precipitated by hot water maps at chromosome 10q21.3-q22.3. Hum Genet 125:541–549.

Ribak C, Khurana V, Lien N. 1994. The effect of midbrain collicular knife cuts on audiogenic seizure severity in the genetically epilepsy-prone rat. J Hirnforsch 35:303–11.

Ribak CE, Morin CL. 1995. The role of the inferior colliculus in a genetic model of audiogenic seizures. Anat Embryol (Berl) 191:279–295.

Rieckhof GE, Yoshihara M, Guan Z, Littleton JT. 2003. Presynaptic N-type Calcium Channels Regulate Synaptic Growth. J Biol Chem 278:41099–41108.

Rimmer EM, Milligan NM, Richens A. 1987. A comparison of the acute effect of single doses of vigabatrin and sodium valproate on photosensitivity in epileptic patients. Epilepsy Res 1:339–346.

Robinson R, Gardiner M. 2004. Molecular basis of Mendelian idiopathic epilepsies. Ann Med 36:89–97.

Rodrigue A, Coulombe Y, Jacquet K, Gagné JP, Roques C line, Gobeil S, Poirier G, Masson JY. 2013. The RAD51 paralogs ensure cellular protection against mitotic defects and aneuploidy. J Cell Sci 126:348–359.

Rogers MF, Shihab HA, Mort M, Cooper DN, Gaunt TR, Campbell C. 2018. FATHMM-XF: Accurate prediction of pathogenic point mutations via extended features. Bioinformatics 34:511–513.

Rojano E, Seoane P, Ranea JAG, Perkins JR. 2019. Regulatory variants: from detection to predicting impact. Brief Bioinform 20:1639–1654.

Roos RAC, Dijk JG Van. 1988. Reflex-epilepsy induced by immersion in hot water: Case report and review of the literature. Eur Neurol 28:6–10.

Roy Choudhury S, Satishchandra P, Sinha S, Anand A. 2019. Gene variants in ZGRF1 implicated for a rare sensory reflex epilepsy. bioRxiv Genet 728188.

Roy R, Chun J, Powell SN. 2012. BRCA1 and BRCA2: Different roles in a common pathway of genome protection. Nat Rev Cancer 12:68–78.

Sahoo PK, Lee SJ, Jaiswal PB, Alber S, Kar AN, Miller-Randolph S, Taylor EE, Smith T, Singh B, Ho TSY, Urisman A, Chand S, et al. 2018. Axonal G3BP1 stress granule protein limits axonal mRNA translation and nerve regeneration. Nat Commun 9:1–14.

Salkoff L, Kelly L. 1978. Temperature-induced seizure and frequency-dependent neuromuscular block in a ts mutant of Drosophila. Nature 273:156–158.

Sambrook J, Russell D. 2001. Molecular Cloning: A Laboratory Manual, 3rd Edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Sanchez-Carpintero R, Patiño-Garcia A, Urrestarazu E. 2013. Musicogenic seizures in Dravet syndrome. Dev Med Child Neurol 55:668–670.

Santos-Silva R, Passas A, Rocha C, Figueiredo R, Mendes-Ribeiro J, Fernandes S, Biskup S, Leão M. 2015. Bilateral frontoparietal polymicrogyria: A novel GPR56 mutation and an unusual phenotype. Neuropediatrics 46:134–138.

Satishchandra P. 2003a. Geographically Specific Epilepsy Syndromes in India Hot-Water Epilepsy. Epilepsia 44:29–32.

Satishchandra P. 2003b. Hot-Water Epilepsy. Epilepsia 44:29–32.

Satishchandra P, Kallur K, Jayakumr P. 2000. Inter-ictal and ictal 99 m TC ECD SPECT scan in hot-water epilepsy. Epilepsia 42:158.

Satishchandra P, Shivaramakrishana A, Kaliaperumal VG, Schoenberg BS. 1988a. Hot-Water Epilepsy: A Variant of Reflex Epilepsy in Southern India. Epilepsia 29:52–56.

Satishchandra P, Shivaramakrishana A, Kaliaperumal VG, Schoenberg BS. 1988b. Hot-Water Epilepsy: A Variant of Reflex Epilepsy in Southern India. Epilepsia 29:52–56.

Savitha MR, Krishnamurthy B, Ashok DA, Ramachandra NB. 2007. Self Abortion of Attacks in Patients with Hot Water Epilepsy. INDIAN Pediatr 44:296–298.

Scheffer IE, Berkovic S, Capovilla G, Connolly MB, French J, Guilhoto L, Hirsch E, Jain S, Mathern GW, Mosh SL, Nordli DR, Perucca E, et al. 2017. ILAE classification of the epilepsies: Position paper of the ILAE Commission for Classification and Terminology. Epilepsia 58:512–521.

Scheffer IE, Berkovic SF. 1997. Generalized epilepsy with febrile seizures plus. A genetic disorder with heterogeneous clinical phenotypes. Brain 120:479–490.

Schweers BA, Walters KJ, Stern M. 2002. The Drosophila melanogaster translational repressor pumilio regulates neuronal excitability. Genetics 161:1177–1185.

Scimemi A, Tian H, Diamond JS. 2009. Neuronal Transporters Regulate Glutamate Clearance, NMDA Receptor Activation, and Synaptic Plasticity in the Hippocampus. J Neurosci 29:14581–14595.

Senanayake N. 1990. ‘Eating epilepsy’- a reappraisal. Epilepsy Res 5:74–79.

Seneviratne U. 2001. Bathing epilepsy. Seizure 10:516–517.

Shang Y, Huang EJ. 2016. Mechanisms of FUS mutations in familial amyotrophic lateral

sclerosis. Brain Res 1647:65–78.

Shaw NJ, Livingston JH, Minns RA, Clarke M. 1988. Epilepsy Precipitated by Bathing. Dev Med Child Neurol 30:108–111.

Shelkovnikova TA, Dimasi P, Kukharsky MS, An H, Quintiero A, Schirmer C, Buée L, Galas M, Buchman VL. 2017. Chronically stressed or stress-preconditioned neurons fail to maintain stress granule assembly. Nat Publ Gr 8:e2788.

Shen J, Gilmore EC, Marshall CA, Haddadin M, Reynolds JJ, Eyaid W, Bodell A, Barry B, Gleason D, Allen K, Ganesh VS, Chang BS, et al. 2010. Mutations in PNKP cause microcephaly, seizures and defects in DNA repair. Nat Genet 42:245–249.

Shu Z, Smith S, Wang L, Rice MC, Kmiec EB. 1999. Disruption of muREC2/RAD51L1 in Mice Results in Early Embryonic Lethality Which Can Be Partially Rescued in a p53 −/− Background . Mol Cell Biol 19:8686–8693.

Siddiqi O, Benzer S. 1976. Neurophysiological defects in temperature sensitive paralytic mutants of Drosophila melanogaster. Proc Natl Acad Sci U S A 73:3253–3257.

Sigurdsson S, Komen S Van, Bussen W, Schild D, Albala JS, Sung P. 2001. Mediator function of the human Rad51B-Rad51C complex in Rad51/RPA-catalyzed DNA strand exchange. Genes Dev 15:3308–3318.

Silva-Barrat C, Ménini C, Bryére P, Naquet R. 1986. Multiunitary activity naalysis of cortical and subcortical structures in paroxysmal discharges and grand mal seizures in photosensitve baboons. Electroencephalogr Clin Neurophysiol 64:455–468.

Silva S, Altmannova V, Luke-Glaser S, Henriksen P, Gallina I, Yang X, Choudhary C, Luke B, Krejci L, Lisby M. 2016. Mte1 interacts with Mph1 and promotes crossover recombination and telomere maintenance. Genes Dev 30:700–717.

Sirsi D, Armstrong D, Muñoz-Bibiloni J, Redondo B, Park J. 2017. SYN1 Gene Mutation in a Child with Focal Epilepsy and Reflex Bathing Seizures. J Pediatr Epilepsy 06:119–124.

Skradski SL, Clark AM, Jiang H, White HS, Fu YH, Ptáček LJ. 2001. A novel gene causing a mendelian audiogenic mouse epilepsy. Neuron 31:537–544.

Smogorzewska A, Desetty R, Saito TT, Schlabach M, Lach FP, Sowa ME, Clark AB, Kunkel TA, Harper JW, Colaiácovo MP, Elledge SJ. 2010. A Genetic Screen Identifies FAN1, a Fanconi Anemia-Associated Nuclease Necessary for DNA Interstrand Crosslink Repair. Mol Cell 39:36–47.

Snider J, Thibault G, Houry WA. 2008. The AAA+ superfamily of functionally diverse proteins. Genome Biol 9:216.

Stafstrom CE, Carmant L. 2015. Seizures and epilepsy: An overview for neuroscientists. Cold Spring Harb Perspect Biol 7:1–19.

Stalder L, Mühlemann O. 2009. Processing bodies are not required for mammalian nonsense-mediated mRNA decay. RNA 15:1265–1273.

Stein JL, Medland SE, Vasquez AA, Hibar DP, Senstad RE, Winkler AM, Toro R, Appel

K, Bartecek R, Bergmann Ø, Bernard M, Brown AA, et al. 2012. Identification of common variants associated with human hippocampal and intracranial volumes. Nat Genet 44:552.

Stensman R, Ursing B. 1971. Epilepsy precipitated by hot water immersion. Neurology 21:559–562.

Stephani U, Tauer U, Koeleman B, Pinto D, Neubauer BA, Lindhout D. 2004. Genetics of Photosensitivity (Photoparoxysmal Response): A Review. Epilepsia 45:19–23.

Stes E, Laga M, Walton A, Samyn N, Timmerman E, Smet I De, Goormachtig S, Gevaert K. 2014. A COFRADIC protocol to study protein ubiquitination. J Proteome Res 13:3107–3113.

Stevens J. 1962. Central and peripheral factors in epileptic discharge. Clinical studies. Arch Neurol 26:409–19.

Stolz A, Ertych N, Kienitz A, Vogel C, Schneider V, Fritz B, Jacob R, Dittmar G, Weichert W, Petersen I, Bastians H. 2010. The CHK2-BRCA1 tumour suppressor pathway ensures chromosomal stability in human somatic cells. Nat Cell Biol 12:492–499.

Striano S, Coppola A, Gaudio L del, Striano P. 2012. Reflex seizures and reflex epilepsies: Old models for understanding mechanisms of epileptogenesis. Epilepsy Res 100:1–11.

Striano S, Meo R, Bilo L, Soricellis M, Ruosi P. 1993. Epilepsia arithmetices: study of four cases. Seizure 2:35–43.

Sun L, Gilligan J, Staber C, Schutte RJ, Nguyen V, O'Dowd DK, Reenan R. 2012. A knock-in model of human epilepsy in Drosophila reveals a novel cellular mechanism associated with heat-induced seizure. J Neurosci 32:14146–14155.

Suzuki T, Delgado-Escueta V A, Aguan K, Alonso ME, Shi J, Haras Y, Nishidas M, Numata T, Medina MT, Takeuchi T, Morita R, Bai D, et al. 2004a. Mutations in EFHC1 cause juvenile myoclonic epilepsy. Nat Genet 36:842–849.

Suzuki T, Delgado-Escueta V A, Aguan K, Alonso ME, Shi J, Haras Y, Nishidas M, Numata T, Medina MT, Takeuchi T, Morita R, Bai D, et al. 2004b. Mutations in EFHC1 cause juvenile myoclonic epilepsy. Nat Genet 36:842–849.

Szymonowicz W, Meloff KL. 1978. Hot Water Epilepsy. Can J Neurol Sci 5:247–251.

Tabarés-Seisdedos R, Rubenstein JLR. 2009. Chromosome 8p as a potential hub for developmental neuropsychiatric disorders: implications for schizophrenia, autism and cancer. Mol Psychiatry 2009 146 14:563–589.

Tassinari CA, Rubboli G, Rizzi R, Gardella E MR. 1998. Self-induction of visually-induced seizures. Adv Neurol 75:179–92.

Tauer U, Lorenz S, Lenzen KP, Heils A, Muhle H, Gresch M, Neubauer BA, Waltz S, Rudolf G, Mattheisen M, Strauch K, Nürnberg P, et al. 2005. Genetic dissection of photosensitivity and its relation to idiopathic generalized epilepsy. Ann Neurol 57:866–873.

Tezer FI, Ertas N, Yalcin D, Saygi S. 2006. Hot water epilepsy with cerebral lesion: A report of five cases with cranial MRI findings. Epilepsy Behav 8:672–676.

Thakran S, Guin D, Singh P, Singh P, Kukal S, Rawat C, Yadav S, Kushwaha SS, Srivastava AK, Hasija Y, Saso L, Ramachandran S, et al. 2020. Genetic landscape of common epilepsies: Advancing towards precision in treatment. Int J Mol Sci 21:7784.

The Epi4K Consortium. 2012. Epi4K: Gene discovery in 4,000 genomes. Epilepsia 53:1457.

Thomas RH, Berkovic SF. 2014. The hidden genetics of epilepsy—a clinically important new paradigm. Nat Rev Neurol 2014 105 10:283–292.

Tourrière H, Chebli K, Zekri L, Courselaud B, Blanchard JM, Bertrand E, Tazi J. 2003. The RasGAP-associated endoribonuclease G3BP assembles stress granules. J Cell Biol 160:823–831.

Tsuzuki T, Fujii Y, Sakumi K, Tominaga Y, Nakao K, Sekiguchi M, Matsushiro A, Yoshimura Y, MoritaT. 1996. Targeted disruption of the Rad51 gene leads to lethality in embryonic mice. Proc Natl Acad Sci U S A 93:6236.

Tyson X, Graber E, Freemantle E, Anadolu MN, Hébert-Seropian S, Macadam RL, Shin U, Hoang H-D, Alain T, Lacaille J-C, Sossin WS. 2017. Cellular/Molecular UPF1 Governs Synaptic Plasticity through Association with a STAU2 RNA Granule. J Neurosci 37:9116–9131.

Ullal GR, Satishchandra P, Kalladka D, Rajashekar K, Archana K, Mahadevan A, Shankar SK. 2006. Kindling & mossy fibre sprouting in the rat hippocampus following hot water induced hyperthermic seizures. Indian J Med Res 124:331–342.

Ullal GR, Satishchandra P, Shankar SK. 1996. Hyperthermic seizures: An animal model for hot-water epilepsy. Seizure 5:221–228.

Vance C, Scotter EL, Nishimura AL, Troakes C, Mitchell JC, Kathe C, Urwin H, Manser C, Miller CC, Hortobágyi T, Dragunow M, Rogelj B, et al. 2013. ALS mutant FUS disrupts nuclear localization and sequesters wild-type FUS within cytoplasmic stress granules. Hum Mol Genet 22:2676–2688.

Verrotti A, Beccaria F, Fiori F, Montagnini A, Capovilla G. 2012. Photosensitivity: Epidemiology, genetics, clinical manifestations, assessment, and management. Epileptic Disord 14:349–362.

Vezzani A, Fujinami RS, White HS, Preux PM, Blümcke I, Sander JW, Löscher W. 2016. Infections, inflammation and epilepsy. Acta Neuropathol 131:211–234.

Wagnon JL, Briese M, Sun W, Mahaffey CL, Curk T, Rot G, Ule J, Frankel WN. 2012. CELF4 Regulates Translation and Local Abundance of a Vast Set of mRNAs, Including Genes Associated with Regulation of Synaptic Function. PLoS Genet 8:e1003067.

Wang H, Dictenberg JB, Ku L, Li W, Bassell GJ, Feng Y. 2008. Dynamic association of the fragile X mental retardation protein as a messenger ribonucleoprotein between microtubules and polyribosomes. Mol Biol Cell 19:105–114.

Wek RC, Jiang HY, Anthony TG. 2006. Coping with stress: EIF2 kinases and

translational control. Biochem Soc Trans 34:7–11.

Whelan CD, Altmann A, Botía JA, Jahanshad N, Hibar DP, Absil J, Alhusaini S, Alvim MKM, Auvinen P, Bartolini E, Bergo FPG, Bernardes T, et al. 2018. Structural brain abnormalities in the common epilepsies assessed in a worldwide ENIGMA study. Brain 141:391–408.

Wilkins AJ, Bonanni P, Porciatti V, Guerrini R. 2004. Physiology of Human Photosensitivity. Epilepsia 45:7–13.

Wolf P. 2005. From Precipitation to Inhibition of Seizures: Rationale of a Therapeutic Paradigm. Epilepsia 46:15–16.

Wolf P, Goosses R. 1986. Relation of photosensitivity to epileptic syndromes. J Neurol Neurosurg Psychiatry 49:1386–1391.

Wreden C, Verrotti AC, Schisa JA, Lieberfarb ME, Strickland S. 1997. Nanos and pumilio establish embryonic polarity in Drosophila by promoting posterior deadenylation of hunchback mRNA. Development 124:3015–3023.

Wu CF, Ganetzky B. 1980. Genetic alteration of nerve membrane excitability in temperature-sensitive paralytic mutants of drosophila melanogaster. Nature 286:814–816.

Wu KY, Hengst U, Cox LJ, Macosko EZ, Jeromin A, Urquhart ER, Jaffrey SR. 2005a. Local translation of RhoA regulates growth cone collapse. Nature 436:1020–1024.

Wu M, Reuter M, Lilie H, Liu Y, Wahle E, Song H. 2005b. Structural insight into poly(A) binding and catalytic mechanism of human PARN. EMBO J 24:4082–4093.

Yan S, Song M, Ping J, Lai S, Cao X, Bai C-J, Xie D-F, Guan H, Gao S, Zhou P-K. 2021. ZGRF1 promotes end resection of DNA homologous recombination via forming complex with BRCA1/EXO1. Cell Death Discov 2021 71 7:1–12.

Yavuz EN, Ozdemir O, Catal S, Bebek N, Ozbek U, Baykan B. 2012. Bromodomain-Containing Protein 2 gene in photosensitive epilepsy. Seizure 21:646–648.

Zhang YQ, Bailey AM, Matthies HJG, Renden RB, Smith MA, Speese SD, Rubin GM, Broadie K. 2001. Drosophila Fragile X-Related Gene Regulates the MAP1B Homolog Futsch to Control Synaptic Structure and Function. Cell 107:591–603.

Web sources

BLAST tool at National Centre for Biotechnology Information, National Institute of Health: http://blast.ncbi.nlm.nih.gov/Blast.cgi

ClustalW2 at European Bioinformatics Institute: http://www.ebi.ac.uk/Tools/msa/clustalw2/

Entrez Gene database at National Centre for Biotechnology Information (NCBI, National Institute of Health (NIH): http://www.ncbi.nlm.nih.gov/gene

Entrez SNP database at NCBI, NIH: http://www.ncbi.nlm.nih.gov/snp/

GenBank database at NCBI, NIH: http://www.ncbi.nlm.nih.gov/genbank/

gnomAD database: https://gnomad.broadinstitute.org/

Human Genome Map Viewer database at NCBI, NIH: http://www.ncbi.nlm.nih.gov/mapview/

Primer3 Input software: http://frodo.wi.mit.edu/

Ensembl genome browser: https://asia.ensembl.org/index.html

1000 Genomes dataset: http://browser.1000genomes.org/index.html

Exome Aggregation Consortium (ExAC) dataset: http://exac.broadinstitute.org/

Exome Variant Server (EVS) NHLBI GO Exome Sequencing Project (ESP): http://evs.gs.washington.edu/EVS/

PolyPhen-2: http://genetics.bwh.harvard.edu/pph2/

SIFT: http://sift.bii.a-star.edu.sg/

FATHMM: http://fathmm.biocompute.org.uk/