Mouse genome modification and investigation of episodic disease

Louise Jane Robertson

B.Sc. (Biomedical Science)

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

Department of Molecular and Cellular Biology

School of Biological Science

University of Adelaide, Australia

May 2018

Contents

Abstract

Mouse models are essential tools for biomedical research, allowing researchers to investigate gene function and model human diseases. The mouse genome can be manipulated to ablate ("knockout") gene function, create targeted point mutations or insert transgenes or small epitope tags. Previous gene targeting methods for creating these modifications were slow and inefficient. The arrival of CRISPR/Cas9 genome editing technology has revolutionised the production of genetically modified mice as it is easy to use, efficient and cost-effective. The research comprised in this PhD thesis explores the use of genetically modified mice to better understand human diseases, as well as developing new technologies for mouse genome editing.

The first manuscript describes a phenotypic investigation of the transgenic *Prrt2* knockout (*Prrt2* KO) mouse. In humans, mutations in *PRRT2* cause an infantile epilepsy syndrome (benign familial infantile epilepsy; BFIE) and a movement disorder in adolescence (paroxysmal kinesigenic dyskinesia; PKD). We identifed a spontaneous paroxysmal phenotype in *Prrt2* KO animals, as well as premature death in HET and KO mice. Behavioural tests also revealed learning deficits and gait abnormalities in KO mice that may reflect phenotypes in homozygous patients, confirming the utility of this model for investigating *PRRT2*-related disorders.

The second manuscript examines variants of the *Streptococcus pyogenes* Cas9 endonuclease (WT SpCas9) commonly used for CRISPR genome editing. Numerous Cas9 variants have recently been characterised, each recognising different PAM sequences. Most of these variants remain untested for genome editing in mice. In this study, we tested a selection of

endonuclease variants (SpCas9 VQR, SpCas9 VRER, SaCas9 KKH and AsCpf1) for their ability to edit the mouse genome via mouse zygote injection, with the aim of expanding PAM targeting options. We showed that all variants are able to mutate the mouse genome, albeit with different efficiencies. We also highlighted the propensity of SaCas9 KKH to generate heterozygotes or mosaic offspring in which at least one allele remains unmodified. When editing using a ssDNA oligonucleotide repair template, SaCas9 KKH consistently left a wild type allele in correctly targeted offspring, whilst WT SpCas9 frequently mutated the other allele. This characteristic could be beneficial when targeting genes in which nullizygous mutations cause embryonic lethality, as the high efficiency of WT SpCas9 commonly prevents the production of viable offspring.

With CRISPR/Cas9 technology, it has become quick and efficient to tag endogenous proteins with small epitope tags. The third manuscript in this thesis compares a series of commercially available antibodies for their efficacy to detect epitope tags on *Pcdh19* HA-FLAG expression in mouse brain tissue. This model was generated within the laboratory to allow specific staining of PCDH19 for investigating protocadherin 19 girls clustering epilepsy. Of the 8 antibodies tested, only two (one HA and one FLAG) were specific for PCDH19. This data will provide guidance for researchers designing similar studies, preventing extensive optimisation of immunofluorescent staining.

Together, the data presented in this thesis demonstrate the versatility and importance of mouse models for the study of gene function and neurological disease. This research provides more options for producing genetically modified mice and streamlines the downstream applications of endogenous epitope tagged genes.

PhD thesis declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship and University of Adelaide Upgrade to PhD Scholarship.

Louise Robertson

Acknowledgements

First and most importantly, I would like to thank my supervisor Prof. Paul Thomas for your unwavering support throughout my years in the lab. Your wealth of knowledge and contagious enthusiasm for science has been a constant source of inspiration and helped to shape my approach to new projects and challenges. Without your guidance, approachability and encouragement- especially when things were not going to plan- my PhD experience would not have been such a positive one.

To my co-supervisor Dr James Hughes, thank you for being such a friendly, patient and knowledgeable teacher. You are always so willing to discuss science and results but have also been an invaluable source of wisdom and friendship.

Thank you also to other Thomas lab members who have made my time in the lab so wonderful. To Sandie for your friendship, expertise with mouse work and all those microinjection sessions. Thank you to Adi, for teaching me so much about CRISPR, our weekend lab pursuits and for bringing fun and humour into the lab. To Dan for so readily troubleshooting problems with me and to Stef for your support and encouragement. To Dale for being a good friend, a great teacher and always making yourself available to help. Thank you, Chan, for the funny chats, dry sense of humour and chocolate/brownie supply. Emily, thanks for all the helpful (scientific and not) chats we shared in the office- you have been such a supportive and kind friend. Mel, Ruby and Ella, I feel so lucky to have spent the last few years working with such wonderful and close friends. Not only have you always been happy to answer silly questions, troubleshoot strange results and provide valuable scientific feedback, you also supported me through the ups and

downs that are life and a PhD. I look forward to lifelong friendships with such strong and intelligent women.

Thank you to the collaborators both in Adelaide and interstate who took an interest in the project, contributed to the research, shared valuable insights and taught me new techniques. To all the animal services and other facility staff that assisted me both in Adelaide, I appreciate the many and varying ways in which you have made my research possible.

To my family, particularly my parents Trudy and Bill, thank you for your unconditional love and support. You have provided me with so many opportunities and always encouraged me to challenge myself and follow my passion. Thank you to Maddie and to my 'Adelaide family' of housemates/friends for all the encouragement and understanding over the years. Finally, to Geoff, thank you for the love and reassurance you have given throughout this experience. I am very lucky to have someone so supportive to share the excitements and frustrations of life and my scientific endeavours.

Presentations

Lorne Genome 2018- Lorne, February 2018

Poster presentation

Robertson, L., Adikusuma, F., Piltz,S., White, M., Pederick,.D., Ahladas, M., Thomas, P.

CRISPR variants for mouse genome editing- Do they make the cut?

ComBio 2017- Adelaide, October 2017

Oral presentation

Robertson, L., Adikusuma, F., Piltz,S., White, M., Pederick,.D., Ahladas, M., Thomas, P.

CRISPR variants for mouse genome editing- Do they make the cut?

Genome Engineering: The CRISPR/Cas Revolution- Cold Spring Harbor Laboratory, New York, July 2017

Poster presentation

Robertson, L., Adikusuma, F., Pederick, D., Piltz, S., White, M., Hughes, J., Thomas, P.

Expanding the toolkit for mouse genome editing: Analysis of CRISPR variants with different PAM recognition

Adelaide Protein Group Student Awards- Adelaide, June 2017

Poster presentation

Robertson, L., Adikusuma, F., Pederick, D., Piltz, S., White, M., Hughes, J., Thomas, P.

Expanding the toolkit for mouse genome editing: Analysis of CRISPR variants with different PAM recognition

EMBL Australia Postgraduate Symposium- Adelaide, November 2016

Poster presentation

Robertson, L., Hughes, J., Thomas, P.

Functional Analysis of PRRT2 in epilepsy and episodic disorders.

The University of Adelaide School of Biological Sciences Postgraduate Symposium- Adelaide, July 2016

Oral presentation

Robertson, L., Hughes, J., Thomas, P.

Functional Analysis of PRRT2 in epilepsy and episodic disorders.

10th Florey Postgraduate Research Conference- Adelaide, September 2016

Poster presentation

Robertson, L., Hughes, J., Thomas, P.

Functional Analysis of PRRT2 in epilepsy and episodic disorders.

Epilepsy group meeting- Melbourne, January 2015

Oral presentation

Robertson, L., Hughes, J., Thomas, P.

Elucidating the function of PRRT2 in ICCA and episodic disorders.

Awards

Adelaide Protein Group Student Awards- Adelaide, June 2017

Best poster presentation

Robinson Research Institute Travel Grant- 2017

Funding to support travel to Genome Engineering: The CRISPR/Cas Revolution in New York, USA

This page is left intentionally blank.

Chapter 1:
Introduction

1.1 Epilepsy

Epilepsy is a heterogeneous disorder characterised by recurring seizures. It affects around 3% of the population at some point in life and due to its unpredictable nature, can be extremely deleterious to quality of life [7]. Clinically, epileptic seizures can be categorised into two broad classes: focal and nonfocal epilepsy. Focal epilepsies originate from one cerebral hemisphere, whilst nonfocal is generalized, involving both hemispheres [8].

Whilst some causes for epilepsy have been known for decades (e.g. stroke, trauma), many were considered 'idiopathic', indicating an unknown cause. It is only more recently that the genetic causes have come to the fore, revealing a number of epilepsies with single-gene causes, as well as those with more complex inheritance involving multiple genetic loci [9]. It is likely that many of these epilepsies are a result of an imbalance in excitatory and inhibitory neurotransmission within neural networks of the brain. In line with this, many monogenic epilepsies are caused by mutations in genes encoding ion channels (channelopathies), whilst others involve dysfunction in regulators of neurotransmission [10, 11]. An emerging example of a single-gene epilepsy is benign familial infantile epilepsy (BFIE), which is caused by mutations in *PRRT2*. Mutations in the same gene have also been linked to movement and other episodic disorders.

1.2 Mutations in *PRRT2* cause PKD, BFIE and ICCA

Benign familial infantile epilepsy (BFIE) is an autosomal dominant epilepsy disorder that affects infants with an onset between 3-12 months of age. Seizures are usually afebrile and partial complex (beginning in one region of the brain and causing loss of consciousness) or generalised tonic-clonic in nature (affecting the entire brain and causing stiffening and uncontrolled jerking) often occurring in clusters, although some phenotypic variability exists. The disorder tends to offset by around 2 years of age and affected infants have normal neurological development thereafter [12-14].

Paroxysmal kinesigenic dyskinesia (PKD; also known as paroxysmal kinesigenic choreoathetosis) is an episodic movement disorder that affects patients in late-childhood or adolescence, becoming less severe with age [12]. The disorder is characterised by brief attacks (≈ 1 min) of dystonic (twisting or abnormal) movements and/or choreoathetosis (the combination of contractions and twisting/writhing) that are brought on by sudden voluntary movements such as standing up from a resting position. Patients do not feel pain or lose consciousness during attacks. Inheritance is autosomal dominant with incomplete penetrance, though sporadic cases have also been identified [15, 16]. PKD falls within a group of three paroxysmal movement disorders that can be discriminated by the stimuli that give rise to the attacks. The other included disorders are paroxysmal exercise-induced dyskinesia (PED) which is brought on by exercise and paroxysmal non-kinesigenic dyskinesia (PNKD) that can be induced by stress, caffeine or alcohol.

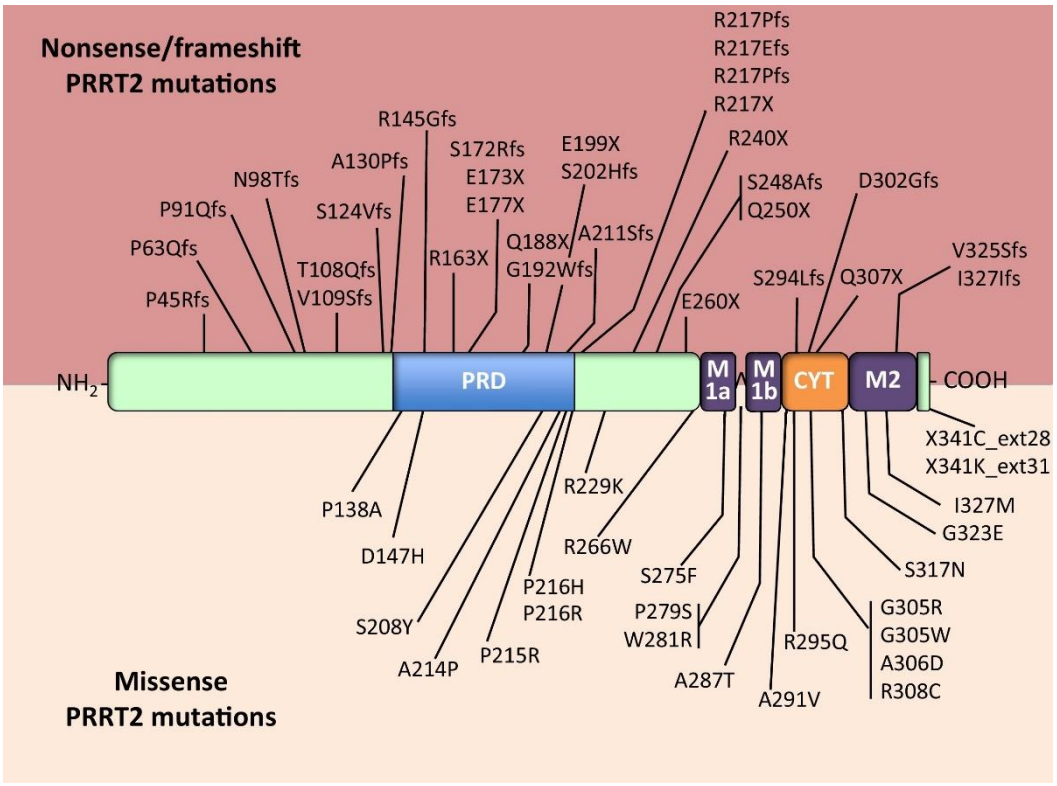
Infantile convulsions and choreoathetosis syndrome (ICCA; also infantile convulsions with PKD; PKD/IC) is a separate disorder in which an individual or a family can be affected by both BFIE

and PKD. Previously, linkage analyses showed that all three disorders link to the same region on chromosome 16, a region including *PRRT2* [17, 18].

Whilst linkage to this region was first identified in 1997 [17] and genes within it screened for mutations, it was not until 2011 that Chen et al. identified heterozygous mutations in proline rich transmembrane protein 2 (*PRRT2)* in eight Chinese families with PKD [16]. Following this discovery, the gene was shown to be causative of both BFIE and ICCA across many families with varying genetic backgrounds [13, 19, 20]. A number of other episodic disorders have also been associated with *PRRT2* mutations including hemiplegic migraine (HM) and other forms of paroxysmal dyskinesia (PED, PNKD) [21, 22]. Homozygous mutations in *PRRT2* have been observed in two consanguineous families and in both cases resulted in severe intellectual disability. Whilst one family was described as non-syndromic, two homozygous individuals from the other experienced PKD, episodic ataxia (EA) and childhood absence epilepsy (CAE) [23, 24].

More than 50 different mutations have been identified within *PRRT2* to date, including frameshift, missense, nonsense and splice site mutations, as well as 3 large deletions that span up to 30 genes, including *PRRT2* [1, 12, 13]. A breakdown of mutations, their type and location in the gene is shown in Figure 1.1a. There is no apparent genotype-phenotype correlation to explain the different disease outcomes, indicating that other environmental or genetic factors are likely influencing the phenotypic expression [12, 14].

a)



b)

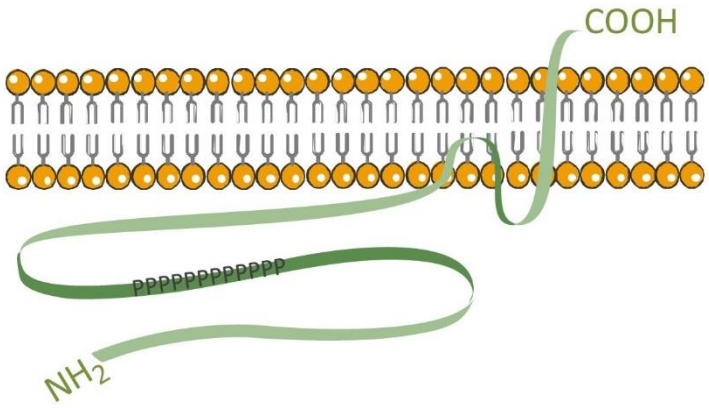


Figure 1.1: Mutations in proline rich transmembrane protein 2 (*PRRT2*) cause paroxysmal disorders. a) *PRRT2* gene structure with disease causing mutations annotated. Nonsense/frameshift mutations are detailed above with missense mutations below. Relevant domains highlighted including proline rich domain (PRD), membrane bound domains (M1a, M1b & M2) and cytoplasmic loop (CYT). b) Simple schematic of PRRT2 protein structure and placement on the intracellular side of the membrane. Proline rich region marked. Adapted from Valtorta et al. [1].

1.3 *PRRT2* structure

Before its implication in disease, *PRRT2* was a largely uncharacterised gene. It consists of four exons, which encode a polypeptide that is 340 amino acids long, containing two putative transmembrane domains near the C-terminus and what was thought to be an extracellularly located N-terminus [16]. However, using live immunolabelling, electron microscopy and computational modelling, recent studies have revealed that PRRT2 has an intracellular N-terminus and a short extracellular region at the C-terminus [25]. It is classified as a Type II transmembrane protein, with only a single transmembrane domain. The second hydrophobic domain associates with the intracellular membrane surface (Figure 1.1b).

Most mutations are predicted to lead to truncation of PRRT2, thereby causing a loss of protein function. The missense mutations identified mostly cluster around the C-terminal end of the protein which contains the hydrophobic domains (Figure 1a). These domains are highly conserved (~90% similarity with other mammals) and the entire gene shares 80-90% similarity with most mammals [20]. As most *PRRT2* mutations discovered thus far affect one or both of these highly conserved hydrophobic domains, there is strong evidence to support their importance in the function of PRRT2 and their loss of function in disease mutations.

1.4 Neuronal function of PRRT2

Analysis of *Prrt2* expression has provided clues for its role in brain-related disorders. In mice, it is largely confined to the CNS, with highest expression in the cortex, basal ganglia and cerebellum [16, 19, 20, 26, 27]. These regions of high *Prrt2* expression are consistent in the human adult brain [1, 14]. Temporally, levels of PRRT2 are low during early murine development (before embryonic day 16) and peak at postnatal day 14. Expression levels in the mouse brain are lower and persist throughout adulthood [16, 26].

Subcellular localisation has been investigated by multiple groups, beginning with colocalisation studies in mouse primary neurons, showing an overlap between overexpressed *PRRT2* and synapse protein synapsin 1 at synaptic puncta [20]. Ultrafractionation of mouse brain tissue has further confirmed its synaptic localisation, with high levels in pre-synaptic fractions and only low levels in post-synaptic densities [26, 28].

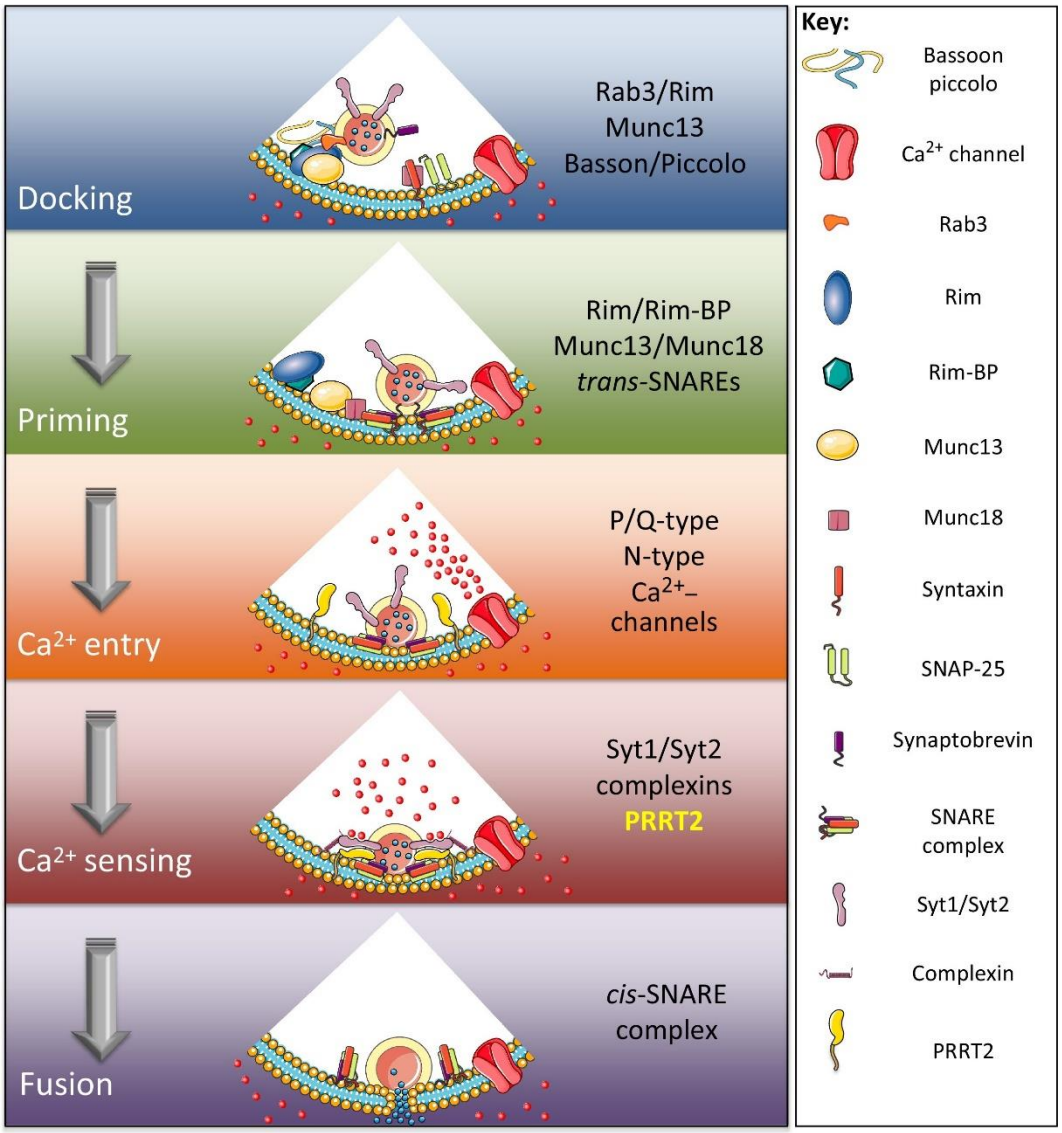
Consistent with this, coimmunoprecipitation and pulldown analyses have revealed numerous interactions with synapse proteins, suggesting a functional role for PRRT2 at the synapse. The first interaction described was with SNAP25 (synaptosomal associated protein of 25 kDa), a component of the SNARE complex, responsible for bridging synaptic vesicles to the membrane during exocytosis [20, 29]. Furthermore, Valente et al. described interactions with another SNARE protein VAMP2/synaptobrevin-2 and Ca^{2+} sensors synaptotagmin 1 and 2 (Syt 1 & 2), whilst Tan et al. observed PRRT2 interaction with Syntaxin 1A [26, 27]. When an action potential occurs along a neuron, the result is an influx of calcium (Ca^{2+}) to the presynaptic terminal. This is sensed by Syt1/2 on the vesicle surface, which then bind the SNARE proteins. The SNARE complex consists of SNAP-25 associated with membrane bound syntaxin 1 and vesicle-bound

synaptobrevin, which then catalyse vesicle fusion to the membrane and neurotransmitter release into the synapse (See Figure 1.2a) [1, 30].

Knockdown of *Prrt2 in vitro* and *in vivo* results in defects in synapse structure and function. *Prrt2* shRNA silencing in primary cultured neurons results in a decrease in synaptic density, and knockdown *in vivo* causes decreased dendritic spine density, suggesting a role in development [26, 28]. Further to this, knockdown of *Prrt2* in the developing mouse brain also results in delays in neuronal migration, though by postnatal day 1 (P1) knockdown cells reach the cortical plate as normal [28]. Electrophysiological examination of the effects of *Prrt2* knockdown *in vitro* revealed decreased synchronous release, concomitant with only subtle differences in asynchronous release (Ca^{2+} independent). *Prrt2* silenced neurons are also insensitive to increases in extracellular Ca^{2+} [26]. This, along with its association with Syt1/2, indicates a role for PRRT2 in pairing Ca^{2+} sensing machinery with the SNARE complex proteins to facilitate synchronous exocytosis of synaptic vesicles (Figure 1.2b) [1].

A recent *in vitro* study using reconstituted protein assays showed that PRRT2 can also inhibit SNARE-mediated vesicle fusion[31]. The expression of wild type *Prrt2* in PC12 cells inhibited fusion events through weak interactions with synaptic SNARE proteins. Furthermore, a non-synaptic role for PRRT2 was described by Fruscione et al., through studies on patient neurons derived from iPS cells and *Prrt2* KO mouse neurons [32]. Researchers demonstrated that PRRT2 inhibits Na^+ influx by directly interacting and limiting surface expression of sodium channels $Na_v1.2$ and $Na_v1.6$. This results in higher Na^+ currents in homozygous mutant patient neurons and KO mouse primary neurons, ultimately resulting in increased excitability.

a)



b)

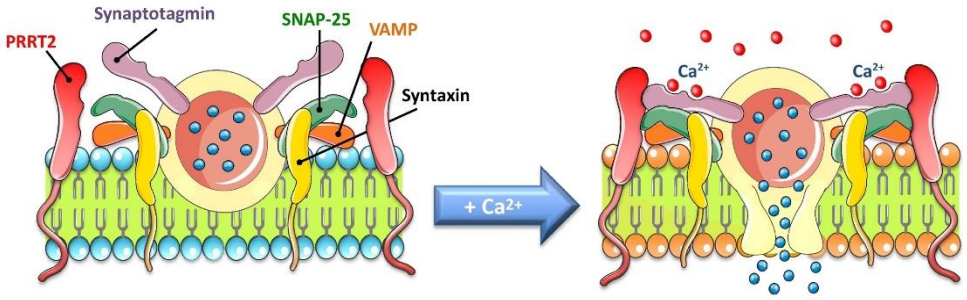


Figure 1.2: Key steps for synaptic vesicle exocytosis. a) Diagram of the processes involved in neurotransmitter exocytosis including key proteins and the proposed involvement of PRRT2. b) A mechanistic model for the role of PRRT2 in responding to intracellular calcium to mediate the final steps of exocytosis. Adapted from Valtorta et al.[1]

1.5 *Prrt2* KO mice

As human disease mutations in *PRRT2* are expected to lead to a loss of protein function, a *Prrt2* knockout mouse model (*Prrt2* KO) has the potential to be a useful system for investigating the underlying mechanisms of the related diseases. Recently, two groups have investigated mice with mutations in *Prrt2*.

The first, described by Michetti et al., was a *Prrt2* 'knockout first' (*Prrt2* KO) mouse with a splice acceptor and lacZ transgene inserted before the first coding exon, abolishing expression of the protein and allowing visualisation of tissue localisation [33]. LacZ staining in the *Prrt2* KO mouse indicated expression in concentrated regions including the cerebellum, hippocampus and cerebral cortex. *Prrt2* KO pups showed unusual spontaneous motor behaviours such as bouncing and backward walking during development, some of which persisted into adulthood.

Adult mice heterozygous (HET) or homozygous (KO) for the knockout allele were indistinguishable from wild type (WT) when subjected to locomotor, footprint/gait and accelerating rotarod testing, suggesting a normal motor/coordination profile despite motor paroxysms. Given the intellectual disability seen in patients, cognitive testing on these mice was carried out (novel object recognition and cued/contextual fear conditioning test) though no differences between genotypes were observed. Interestingly, whilst *Prrt2* KO and HET mice did not experience spontaneous seizures, KO animals (and not WT) could be triggered into wild running episodes following audiogenic stimulation. EEG activity was unaltered during these attacks indicating a motor paroxysm rather than a seizure. When probed for seizure susceptibility using proconvulsant pentylenetetrazol (PTZ), KO animals displayed significantly longer seizure duration despite no differences in threshold or latency to seizure.

A second group described a series of *Prrt2* mouse models, seeking to elucidate functional regions, as well as establish a model for the human diseases [27]. The most extensively characterised mouse in this study (*Prrt2*^{STOP}) mimics the common disease-causing mutation c649dupC (occurs in 80% of patients) by inserting two stop codons at the corresponding site in mouse *Prrt2* causing loss of function. Electron microscopy analysis on cerebellar samples from these mice reveal no differences in synaptic density or active zone length, though more synaptic vesicles were docked at the membrane. *Prrt2*^{STOP} mice showed normal gait, muscle strength and locomotion but displayed deficits in the beam walking and rotarod tests which test motor coordination and balance. Rotarod testing on the *Prrt2* KO mouse showed no difference between genotypes, making this a point of difference between the two models. The *Michetti et al* model is maintained on a C57Bl/6N substrain whilst the *Tan et al* model is maintained on C57Bl/6J. Sub-strain genomic and phenotypic differences may explain different performance on these tests[34].

Spontaneous attacks of dyskinesia were observed at low frequency in *Prrt2*^{STOP} mice but could not be induced by sudden movement or stimulation with substances such as caffeine, ethanol or cocaine. Kindling stimulation with repeated administration of sub-convulsive electrical impulses resulted in seizures followed closely by dyskinetic attacks in *Prrt2*^{STOP} homozygotes and heterozygotes but not wild type littermates. When injected with PTZ or exposed to heat, *Prrt2*^{STOP} mice had significantly shorter latency to seizure than wild type and experienced attacks of dyskinesia following these seizures. Again, the discrepancy in PTZ results between *Prrt2* ^{STOP} and *Prrt2* KO mice could be attributed to differences in the C57Bl/6 substrain. Optogenetic stimulation of neurons in the cerebellum of *Prrt2*^{STOP} homozygous mice using channel-rhodopsin-2 (ChR2) induced dyskinesia, suggesting an important role for the

cerebellum in *Prrt2* mouse phenotype. When *Prrt2* was knocked out only in granule cells of the cerebellum (using GluN2C-Cre), mice showed similar behavioural phenotypes to *Prrt2*^{STOP} animals. Conversely, mice with conditional knockout in the forebrain only (using CaMKIIa-Cre) did not experience heat induced seizures/dyskinetic attacks. Together these data indicate that the cerebellum is a key brain region in *PRRT2*-related dyskinesia.

1.6 Mouse models for disease

The utility of genetically manipulated mice as tools for biomedical research, particularly when investigating complex neurological disorders such as epilepsy, is demonstrated by the aforementioned *Prrt2* mouse studies. In general, mouse KO models provide an *in vivo* system for understanding gene or protein function, as well as investigating the underlying mechanisms of disease and can be used to perform rigorously controlled experiments that are not possible on human patients. Mice and humans share 99% of their genes and have extensive physiological similarities, allowing researchers to model human pathologies [35]. In addition to the *Prrt2* studies, several other epilepsy genes have been successfully studied using mouse models [36-38].

Compared to other mammals, mice are also relatively cheap to house and have fast reproduction times, with a total lifespan of around two years. The mouse genome is easily manipulated to mimic human disease mutations or to insert short DNA sequences and longer gene or reporter constructs. These DNA modifications have historically been achieved through a method called gene targeting. More recently however, genome editing techniques have been characterised and allow direct injection of reagents into the mouse zygote for one-step generation of mutant founders, making this the method of choice.

1.7 Gene targeting for transgenic mice

Gene targeting utilises the homologous recombination process inherent to mammalian cells to mediate the transfer of exogenous DNA into the genome. To generate mutant mice using this method, researchers deliver a targeting vector containing the donor DNA sequence along with long homology arms into mouse embryonic stem cells (ES cells). The genetically modified ES cells are implanted into the early mouse blastocyst, before transfer into a pseudopregnant female. Resulting founders are chimeric for the desired mutation and are then bred with wild type mice to create heterozygotes for starting a colony.

Since its development in the 1980's, this process has become an important tool for investigating gene function, particularly by knocking out a gene of interest [39]. Indeed, the International Knockout Mouse Consortium (IKMC) aims to create and phenotype knockout mouse lines for all 20,000 known mouse genes and make them publically available to researchers [40]. Many of these mice have been created using a 'knockout first' strategy that causes knockout of the gene but is reversible via the Flp-FRT recombination system. An important feature of this allele type is the insertion of the engrailed splice acceptor followed by a promoterless lacZ gene to facilitate expression of β-galactosidase in place of the gene of interest. Additionally, LoxP sites flank a coding exon to allow conditional knockout of the gene using a mouse with Cre-recombinase under a tissue-specific promoter. The *Prrt2* KO mouse discussed above and used in the study reported in Chapter 2 is an example of a gene targeting mouse created through this consortium [33].

Whilst the process of gene targeting has allowed researchers to investigate gene function and disease with flexibility *in vivo*, the process is costly, time consuming and prone to failure. The

finding that double stranded breaks (DSBs) in DNA can increase the efficiency of gene targeting/ donor DNA integration opened up the field of mouse genome editing dramatically [2, 41, 42]. Meganucleases, zinc finger nucleases (ZFN), transcription activator-like effector nuclease (TALEN) and clustered regularly interspaced short palindromic repeats (CRISPR) are all techniques that use this principal to bring about DNA modification. These technologies, in particular CRISPR/Cas9, have facilitated the rapid and efficient production of mice, overcoming many of the issues (Figure 1.3). They will be introduced here and form the basis of the study reported in Chapter 3.

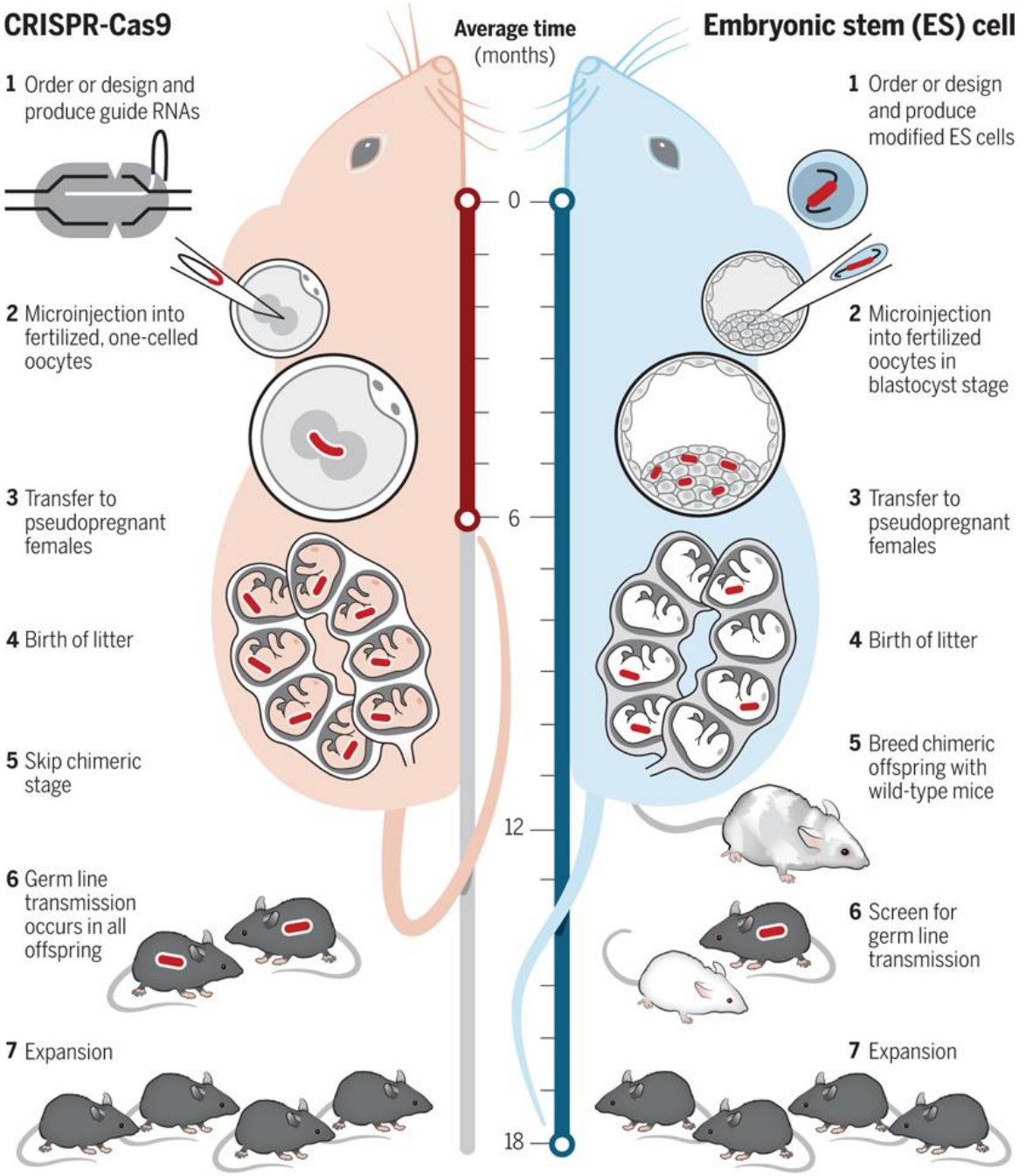


Figure 1.3: CRISPR/Cas9 genome editing technique compared to gene targeting for producing genetically modified mice. Mutations in mice are generated rapidly (less than 6 months) and with high efficiency using CRISPR/Cas9 technology. Adapted from Cohen [3].

1.8 Genome editing techniques

Meganucleases are naturally occurring endonucleases that are able to recognise 14-40bp stretches of DNA and create DSBs. A major difficulty with their use for genome editing, however is that their specific recognition sequence is unlikely to be present at the desired target locus, requiring incorporation of the sequence prior to genome editing. ZFN and TALEN technology overcome this issue as they can be tailored to target virtually any sequence in the genome (Figure 1.4). Both are engineered proteins consisting of naturally occurring DNA-binding domains (ZF or TALE) fused to the non-specific FokI nuclease for target cleavage. For ZFN, each DNA binding domain recognises 3 nucleotides, whilst each TALE domain binds a single nucleotide. These can be arranged in a modular fashion to recognise a target sequence. Since FokI requires dimerization to induce double stranded breaks, pairs of these proteins are used on opposite DNA strands for its use in genome editing [43, 44]. Both techniques have been used effectively for editing of the mouse genome [45-47]. However, each new genomic target requires the design, synthesis and validation of new nucleases. As a result, the difficulties and expenses associated have prevented their wider use [2]. Many of these issues were addressed with the discovery that the bacterial CRISPR/Cas9 system could be adapted for use in genome editing.

Target specificity for CRISPR/Cas9 nucleases is provided by short guide RNAs that form complementary base-pairing with genomic DNA. The nuclease therefore, is the same for every experiment, with the RNA guide being easily designed and produced at little cost. A key advantage of gene editing techniques over the historical gene targeting in mice is that they allow editing directly in the single-cell zygote. After microinjection or electroporation of

reagents, the zygotes are transferred to a pseudopregnant female. Founder pups resulting from these kinds of experiments can be genotyped for the desired mutation and have a high chance of germline transmission given that most mutations occur at the 1-2 cell stage.

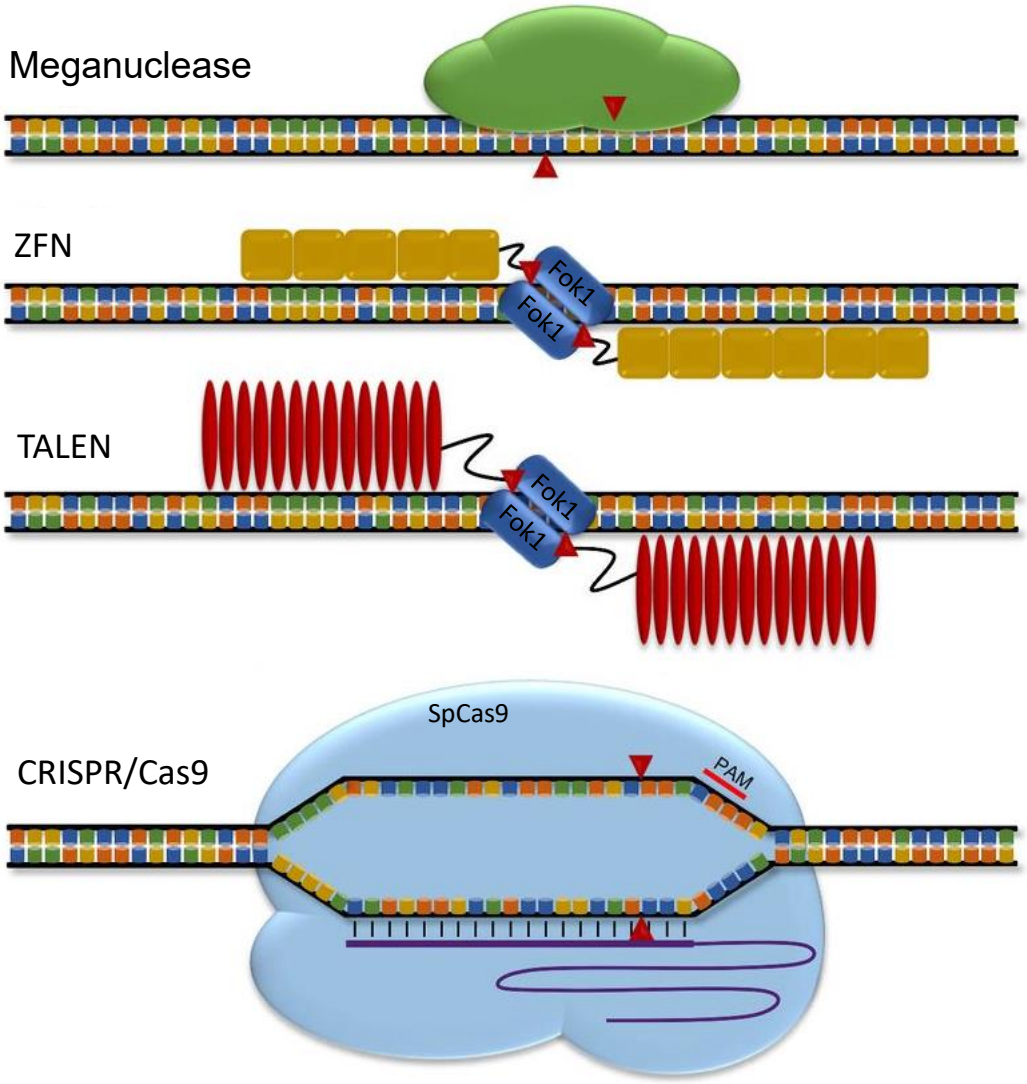


Figure 1.4: Genome editing technologies: meganuclease, zinc-finger nuclease (ZFN), transcription activator-like effector nuclease (TALEN) and CRISPR/Cas9. Meganucleases recognise long (14-40bp) specific sequences of DNA to induce a double-stranded break (indicated by red arrows). ZFN and TALENs are modular proteins that specifically bind DNA and guide non-specific Fok1 nuclease to genomic sites. CRISPR/Cas9 is an endonuclease guided to its target site by a short RNA sequence. Adapted from Hoban and Bauer [5].

1.9 CRISPR/Cas9 technology

CRISPR systems occur naturally in prokaryotes as part of the antiviral adaptive immune system[48]. The CRISPR system most commonly used comes from *Streptococcus pyogenes* (SpCas9). After insertion of exogenous DNA, the bacteria incorporate small segments of this DNA into a CRISPR locus. The CRISPR locus is transcribed into pre-crRNA that is processed into crRNAs. These crRNAs complex with another short RNA called the transactivating RNA (tracrRNA), as well as the Cas9 endonuclease. This complex is guided to the foreign DNA by the crRNA and Cas9 mediates a double-stranded break (DSB) in the DNA, facilitating its degradation [49]. Importantly, the Cas9 requires a specific sequence adjacent to the target region (known as the proto-spacer adjacent motif; PAM). In the case of SpCas9, an NGG PAM is recognised. The PAM sequence is not incorporated into the CRISPR locus, providing a point of difference to distinguish the bacterial cells' own DNA from the exogenous DNA with sequence homology, preventing self-targeting [50].

The CRISPR/Cas9 system from SpCas9 was the first to be characterised for genome editing, triggered by the finding that it could cleave DNA *in vitro*[51]. Shortly thereafter, two groups described genome editing in mammalian cell lines [52, 53]. For simplicity, the crRNA and tracrRNA components were combined, resulting in a single guide RNA (sgRNA) that complexes with Cas9 for target cleavage. This sgRNA is ~100bp in length with only around 20bp of target-specific sequence, making it simple and cheap to produce in the laboratory. The Cas9-RNA complex begins by randomly colliding with the DNA, binding where an NGG is present and rapidly dissociating from non-PAM sites. When a PAM site is bound, the Cas9-RNA complex tests the adjacent DNA for sequence complementarity with the gRNA [50]. Once PAM recognition and

sequence complementarity has been achieved, DNA cleavage is carried out by two separate nuclease domains: RuvC and HNH. These domains nick one strand of DNA each, resulting in a blunt DSB located three nucleotides upstream of the PAM (Figure 1.5) [54, 55]. As discussed later in this chapter, once a DSB has been made in the DNA, it is repaired by the cells' endogenous DNA repair mechanisms which can be exploited to generate desired mutations.

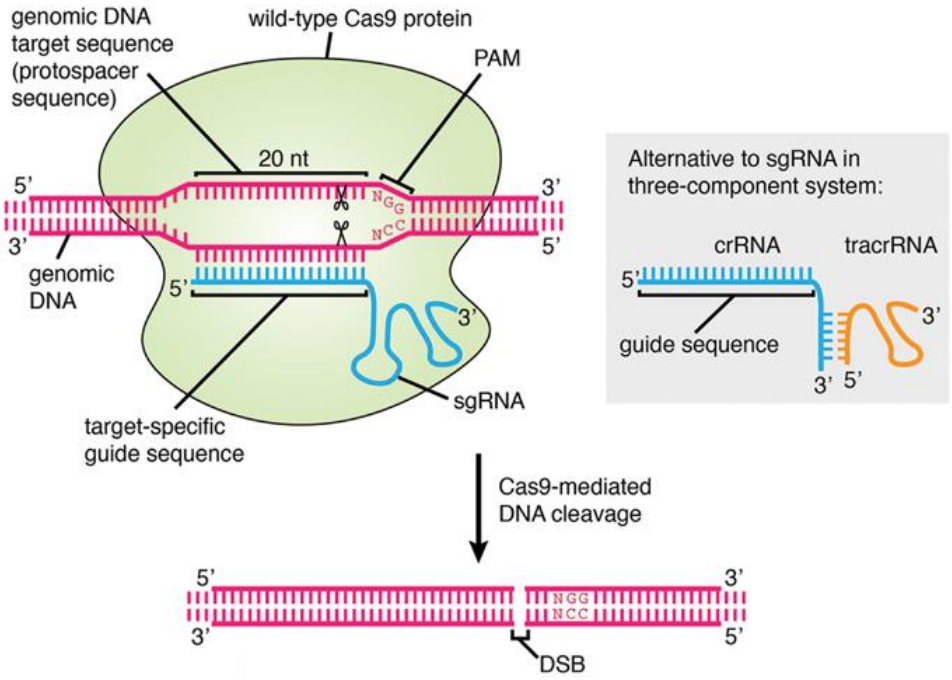


Figure 1.5: CRISPR/Cas9 genome editing technology. The Cas9 protein creates a double-stranded break (DSB) three base pairs upstream of an NGG protospacer adjacent motif (PAM). Target specificity is provided by complexing with the crRNA and tracrRNA, which can be combined into a single guide RNA (sgRNA). Adapted from Agrotis and Ketteler [4].

1.10 Variations of the CRISPR/Cas9 system

One of the limitations of SpCas9 is the requirement for an NGG PAM to be directly adjacent to the desired cleavage site. NGG PAM sites are found every 8-12bp on average. This does not pose a restraint for most CRISPR applications, however when creating a DSB with the aim to precisely insert a sequence using the homology directed repair pathway, more targeting flexibility is often desired. An inverse relationship between mutation incorporation rate and distance from DSB has been described, highlighting a need for more PAM recognition options [56]. Variations of the CRISPR/Cas9 system have been developed for genome editing and contribute to the ever-expanding toolkit of RNA-guided endonucleases(Figure 1.6) [2].

A series of Cas9 orthologues have been characterised, many of which recognise different PAM sequences (summarised in Figure 1.6). These include *Francisella novicida* Cas9 (FnCas9, NGG PAM), *Staphylococcus aureus* Cas9 (SaCas9, NNGRRT PAM), *Neisseria meningitidis* Cas9 (NmCas9, NNNNGATT PAM) and *Streptococcus thermophiles* Cas9 (St1Cas9/St3Cas9, NNAGAAW/NGGNG PAM) [57-60]. Additionally, Zetsche et al. characterised a class 2 CRISPR endonuclease called Cpf1 from both *Acidaminococcus sp. BV3L6* and *Lachnospiraceae bacterium ND2006* that recognises a TTTN PAM 5' of its target sequence and creates a staggered cut in the DNA distal to its PAM [61].

Variants with altered PAM recognition have also been engineered [2]. Kleinstiver et al. used structural information and a bacterial selection system to create the VQR, EQR and VRER variants. VQR SpCas9 contains mutations in three residues (D1135V/R1335Q/T1337R) and recognises a NGAN PAM sequence. The EQR SpCas9 variant, similar to VQR, contains a different residue at the first mutant site (D1135E/R1335Q/T1337R) and shows specificity for an

NGAG PAM sequence. The third SpCas9 variant described in this study is VRER SpCas9 and contains four mutated residues (D1135V/G1218R/R1335E/T1337R), recognising NGCG PAMs [58]. The same group also produced an engineered form of SaCas9 with a less restrictive PAM recognition [62]. The wild type form of SaCas9 recognises an NNGRRT PAM sequence, however molecular evolution of the PAM-interacting domain resulted in a variant with three mutated residues (E782K/N968K/R1015H; SaCas9 KKH) which is highly active at NNNRRT PAM sites. FnCas9 (NGG PAM) was also engineered, resulting in a variant recognising a more flexible YG PAM through mutations in three residues (E1369R/E1449H/R1556A; RHA FnCas9) [57].

Together, the Cas9/Cpf1 endonucleases from different species and engineered variants expand the targetable regions of the genome, providing more flexibility for the design of gRNAs to specific loci. FnCas9, FnCas9 RHA, As/LbCpf1 and SaCas9 have been tested via mouse zygote injection and are able to cleave and produce mutations, though many species/engineered Cas9 variants remain untested [57, 63-65].

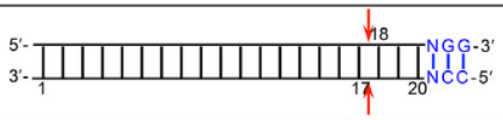
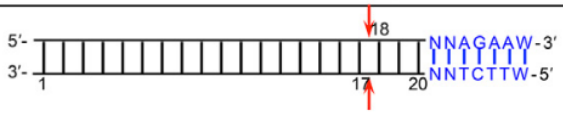
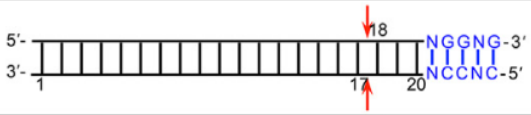
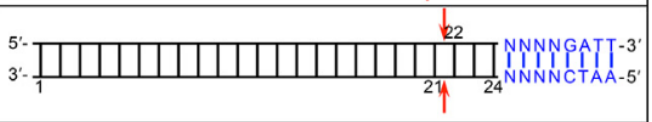
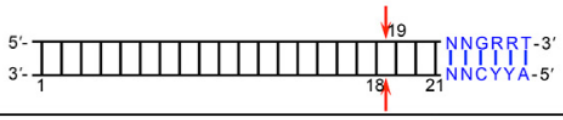
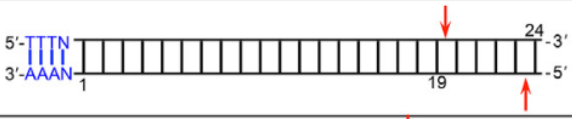
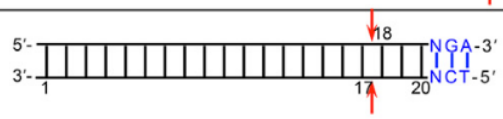
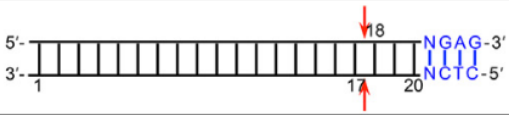
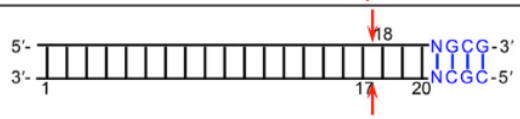
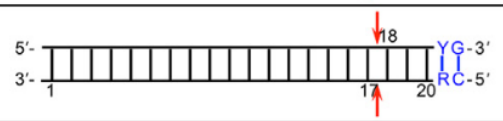
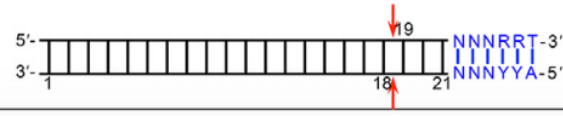
Enzyme name	Size (residues)	PAM requirement and cleavage pattern
SpCas9 / FnCas9	1368 / 1629	
St1Cas9	1121	
St3Cas9	1409	
NmCas9	1082	
SaCas9	1053	
AsCpf1 / LbCpf1	1307 / 1228	
VQR SpCas9	1368	
EQR SpCas9	1368	
VRER SpCas9	1368	
RHA FnCas9	1629	
KKH SaCas9	1053	

Figure 1.6: Naturally occurring and engineered variants of the CRISPR system.
Protospacer adjacent motif (PAM) requirement highlighted in blue with cut sites indicated by red arrow. Adapted from Komor et al. [2]

1.11 CRISPR/Cas9 target specificity

Despite the efficiency and flexibility of the CRISPR/Cas system, a major concern for the technology is off-target activity. This occurs at unintended loci with sequence similarity to the target gene, sometimes with high frequency. Non-specific cleavage can not only confound experiments but is highly problematic for therapeutic applications of the technology. Numerous methods to reduce off-target activity have been investigated, most with the aim to also maintain on-target activity of the enzyme. Bioinformatics tools that assess candidate guides for their likely off-target sites are a useful way to reduce off-target binding. Many of these tools simply require the input of a short sequence of interest and can design the guide, as well as score it for off-target binding against a reference genome. This allows researchers to avoid guides that have very few mismatches to an off-target site, especially if this falls within a coding region of another gene. Despite best efforts to avoid potential off-targets, the issue cannot be completely resolved at point of gRNA design and numerous other methods have been described to reduce these events.

Truncation of the guide sequence at the 5' end to 17 or 18 nucleotides was shown to reduce off-target activity whilst maintaining on target efficiency [66, 67]. The delivery of Cas9 protein and gRNA as a ribonuclear protein complex (RNPs), instead of a plasmid encoding the protein also helps to improve Cas9 specificity as the RNPs are able to cleave DNA immediately after delivery and are degraded quickly, limiting the time that the complex is active within the cell [68, 69].

Other strategies reduce off-target activity by modification of the Cas9 protein, meaning two simultaneous Cas9 binding events must occur for induction of a double stranded break. One such strategy involves mutating one of two catalytic residues (D10A or H840A) to make a

nickase version of Cas9 (Cas9n) that can only cut a single strand (which is easily repaired within the cell without mutation) [70, 71]. Pairing of two gRNAs on opposite DNA strands is required for Cas9n to create a DSB in the DNA and result in mutation. A similar approach fuses a catalytically inactive Cas9 (dCas9) to a non-specific FokI nuclease that requires dimerization for induction of DSB [72-74]. Whilst effective for reducing off-target activity, these strategies can be an issue given the size constraints when using viral delivery as they require two gRNAs and (in the case of FokI) an additional protein domain.

A number of engineered variants of Cas9 have been described that alter non-specific binding of the protein to DNA. These were designed using the crystal structure of Cas9. The enhanced specificity eSpCas9 1.1 variant contains mutations in residues that contact the non-target strand of DNA (K848A/K1003A/R1060A), requiring more stringent base pairing with the target strand and resulting in reduced off-targets [75]. Another high-fidelity variant, known as SpCas9-HF1 alters residues that contact the target DNA strand (N497A, R661A/Q695A/Q926A) to reduce off-target activity [76]. More recently, Jennifer Doudna's group described a new hyper-accurate variant (named HypaCas9) following analysis of domain movement in the eSpCas9 1.1 and SpCas9-HF1 endonucleases [77]. They found the REC3 domain to be responsible for sensing target accuracy before conformational changes allowing DSB induction. Analysis of clusters of mutations within REC3 resulted in the engineering of HypaCas9 with mutations in residues (N692A/M694A/Q695A/H698A). This variant appears to maintain strong on-target cutting whilst showing dramatically increased target specificity.

1.12 Mouse genome editing with CRISPR/Cas9

The presence of a DSB, such as those elicited by CRISPR/Cas9, is extremely deleterious to the cell. Therefore, endogenous DNA repair systems act quickly to repair this (Figure 1.7). The most prominent of these is non-homologous end joining (NHEJ) which ligates broken DNA ends back together [78]. The canonical NHEJ (cNHEJ) pathway does not require a repair template, but instead relies on the DNA binding factors Ku70 and Ku80, which form a heterodimer at the DNA ends for stability and recruit factors that process the DNA ends and facilitate rejoining [79]. If cNHEJ is activated after DSB induction by CRISPR/Cas9, the usual outcomes are small deletions or insertions (indels) at the break point. This has resulted in cNHEJ being termed 'error prone', however this may not be completely true; instead simply a bias as this is the most commonly observed outcome. It is more probable that cNHEJ results in high fidelity repair, however this will restore the target sequence, allowing Cas9 to re-cleave indefinitely until an error in repair eventually destroys the target sequence, preventing further cleavage [80].

Other more error-prone repair mechanisms that can occur after a DSB are single strand annealing (SSA) and alternate end joining/microhomology-mediated end joining (aNHEJ/MMEJ). They differ from cNHEJ in the enzymes that mediate repair and rely on base pairing between short (for MMEJ) or longer (SSA) regions of homology following resection of the DNA. This results in deletion of the region between the homology (Figure 1.7) [81].

Both cNHEJ and aNHEJ repair mechanisms are useful for genome editing as small indels can disrupt the open reading frame (ORF) of a gene, often resulting in nonsense mutations [51, 53, 70]. In mice, harnessing cNHEJ/aNHEJ repair is the simplest and most efficient way to generate a loss of function mutation in a gene of interest [82, 83]. Targeting CRISPR/Cas9 to a site early in the

ORF is likely to truncate the protein and abolish its function. Multiplexed genome editing has also been demonstrated in mice, allowing production of mice with mutations in multiple genes following a single zygote injection [84]. If two DSB are introduced at the same time, a large deletion between the two sites can occur, allowing deletion of whole exon, gene or regulatory region [52, 85]. To date, deletions as large as 24.4Mb have been reported in mice and up to 30Mb in a human cell line [86, 87].

Another repair pathway that can be exploited for genome editing is homology directed repair (HDR). This pathway requires a homologous template (which under normal cellular conditions is usually the sister chromatid) to mediate precise repair over the DSB [6]. To utilise this for editing the mouse genome, a single stranded DNA (ssDNA) oligo or double-stranded DNA (dsDNA) donor can be co-injected with the CRISPR/Cas9 reagents to provide a repair template. This method allows subtle alteration of the genome such as point mutations which are usually incorporated using ssDNA templates with 50-80bp homology flanking the donor sequence [84]. Short DNA sequences such as LoxP sites can also be inserted using two gRNAs and two ssDNA donors, allowing the production of conditional alleles, though many groups have experienced poor efficiency with this method [3, 88].

The CRISPR/ssDNA donor technique has also made it fast and efficient to insert a small epitope tag onto a protein of interest. Epitope tags overcome a common problem faced by researchers who lack of specific antibody for an endogenous protein. This was demonstrated by Yang et al. when they tagged endogenous Sox2 with a V5 tag resulting in tag insertion in 12 out of 35 founders [88]. High-throughput endogenous epitope tagging has also been demonstrated in the

developing mouse brain with CRISPR reagents and oligo donor delivered via *in utero* electroporation [89].

Whilst small sequences can be inserted with reasonable efficiency, the constraints of synthetic ssDNA production mean that longer inserts such as fluorescent reporters or genes are usually inserted via dsDNA templates, delivered via plasmid. These techniques have suffered low efficiency [88, 90]. Recently, a technique called *Easi*-CRISPR (Efficient additions with ssDNA inserts-CRISPR) has been described in mice and demonstrates higher efficiency in producing floxed alleles and reporter insertion. The technique uses long ssDNA donors with short homology arms and it is proposed that the SSA repair pathway might mediate integration into the genome [91].

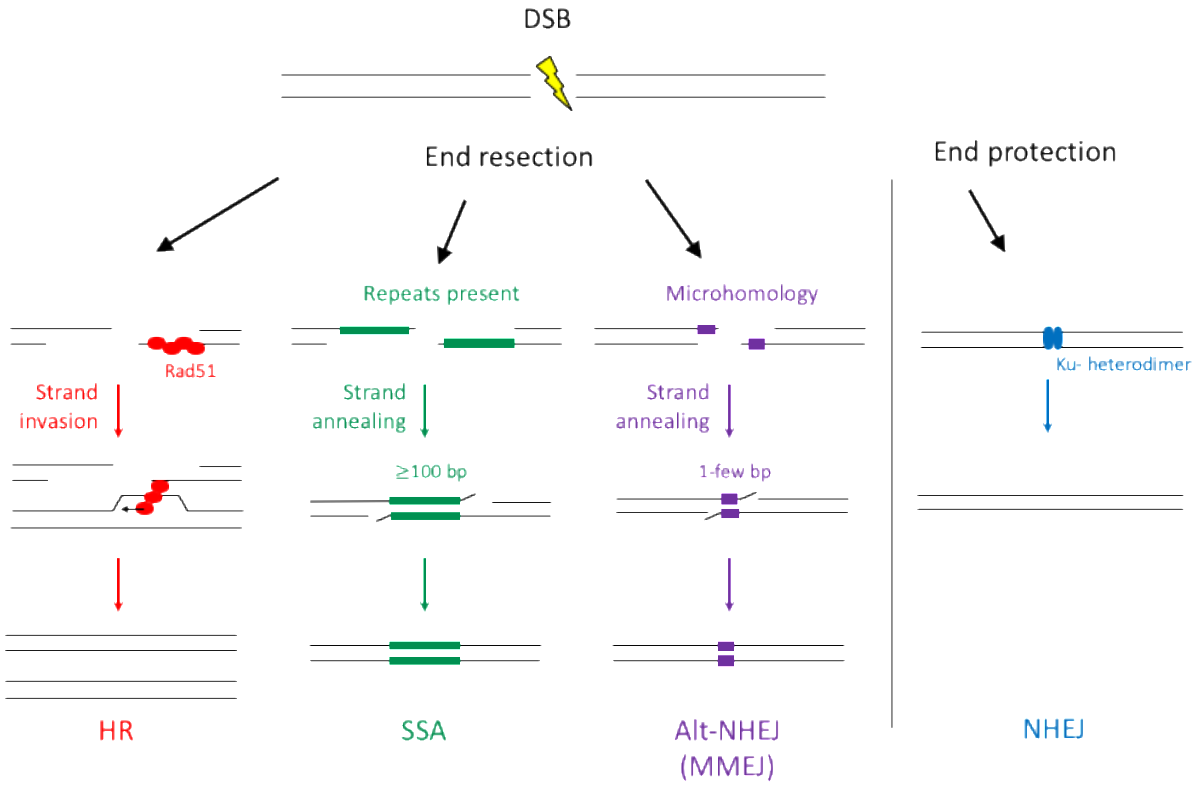


Figure 1.7 Repair mechanisms following a DSB in the DNA. The NHEJ repair mechanism does not involve resection, instead the ends are protected and ligated back together- often without error. HR, SSA and Alt-NHEJ/MMEJ all involve resection of the DNA, and repair either through local homology (resulting in deletion of intervening DNA) or from a homologous repair template such as in HDR (error-free repair). Adapted from Jasin and Rothstein [6].

1.13 Project Rationale

Mutations in *PRRT2* cause a series of paroxysmal disorders, including BFIE and PKD. At the time of its link to disease, very little was known about the function of PRRT2 or its pathogenic role in disease. Recent publications have shown a synaptic function for PRRT2, likely as part of the calcium sensing machinery facilitating exocytosis at the synapse. In Chapter 2 we aim to characterise the *Prrt2* transgenic knockout mouse as a model for these diseases.

At the outset of the project, we envisaged that investigations into the role of PRRT2 in disease using the *Prrt2* KO mouse model would constitute the entire PhD. However, the mild phenotype and numerous recent publications necessitated additional studies to be added to this thesis.

The studies presented in Chapter 3 and 4 investigate facets of CRISPR/Cas9 genome editing. The technology has revolutionised the production of genetically modified mice, making the process faster and more efficient. In Chapter 3 we aim to broaden the targetable regions of the mouse genome by testing a series of endonuclease variants with altered PAM recognition.

Given the relative ease of producing genetically modified mice using CRISPR/Cas9, more possibilities have opened up with the type of modifications that can be introduced. One such modification is the addition of an epitope tag to a protein of interest, allowing visualisation *in vivo* with a commercially available antibody. The study presented in Chapter 4 aims to compare HA and FLAG antibodies on an endogenously tagged protein to demonstrate which are most useful for immunofluorescence on mouse brain tissue.

Chapter 2:

Paroxysmal and cognitive phenotypes in *Prrt2* KO mice

2.1 Summary

Mutations in *PRRT2* cause a spectrum of episodic disorders. An understanding of the mechanism by which this occurs will help to direct treatment options for patients, improving quality of life. Investigations into the function of PRRT2 *in vitro* have provided insight into its involvement in synaptic vesicle exocytosis, as well as showing that loss of function *Prrt2* mutations result in defects in synaptic transmission [26]. To better understand how these defects in cells translate to the paroxysms observed in patients, analysis of a disease model with complex neuronal networks is particularly interesting. As such, two recent studies have investigated mice lacking PRRT2 protein. Both groups observed spontaneous and induced motor paroxysms, in addition to abnormal electrophysiology in the brains of these mice[27, 33].

In this manuscript, we performed further phenotypic characterisation of the transgenic *Prrt2* KO mouse. After showing a loss of PRRT2 protein in mice homozygous for the knockout allele (KO) and reduction in heterozygous animals (HET), we confirmed the synaptic localisation of PRRT2. We also performed transmission electron microscopy to look at synaptic density and characteristics in KO cortex compared to wild type and showed that protein levels of PRRT2 interactors Syt1/2 are unaffected by PRRT2 knockout. Most notably, we describe spontaneous phenotypes in a subset of animals, including spontaneous seizures in KO and unexpected death in both HET and KO. To further investigate the motor phenotype of this mouse model, we undertook a series of behavioural tests including locomotor, rotarod and digigait. Whilst no deficits were observed in locomotion or latency to fall in the rotarod, *Prrt2* KO mice showed differences across a number of fundamental aspects of gait. We also assessed the cognitive capabilities of the KO mice through the Continuous Y-maze and morris water maze (MWM),

including the MWM platform reversal. Mice showed learning deficits in MWM training days, though memory was unaffected. Overall, this manuscript extends the phenotype of *Prrt2* KO mice, describing new epilepsy, motor and cognitive phenotypes that reflect the disease phenotypes in patients.

Statement of Authorship

Title of Paper	Paroxysmal and cognitive phenotypes in Prrt2 mutant mice
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☑ Unpublished and Unsubmitted work written in manuscript style
Publication Details	

Principal Author

Name of Principal Author (Candidate)	Louise Robertson
Contribution to the Paper	Designed and performed experiments, analysed results, wrote manuscript.
Overall percentage (%)	80%
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date 30/4/18.

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate to include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Travis Featherby
Contribution to the Paper	Performed behavioural testing experiments, reviewed manuscript.
Signature	Date 1/5/18

Name of Co-Author	Stuart Howell
Contribution to the Paper	Performed statistical analyses on behavioural testing data, reviewed manuscript.
Signature	Date 30/4/18

Name of Co-Author	James Hughes		
Contribution to the Paper	Conceived and designed study, analysed results, reviewed manuscript.		
Signature		Date	30/4/18

Name of Co-Author	Paul Thomas		
Contribution to the Paper	Conceived and designed study, analysed results, evaluated and edited manuscript.		
Signature		Date	30/4/18

Paroxysmal and cognitive phenotypes in *Prrt2* mutant mice

Louise Robertson[1], Travis Featherby[3], Stuart Howell[4], James Hughes[1] and Paul Thomas[1,2,*]

[1]The University of Adelaide and Robinson Research Institute, Adelaide, Australia

[2]South Australia Health and Medical Research Institute, Adelaide, Australia

[3] Melbourne Brain Centre, Florey Neuroscience Institute, Parkville, Australia

[4] Adelaide Health Technology Assessment, The University of Adelaide, Adelaide, Australia

***Corresponding author**

Abstract:

Mutations in proline-rich transmembrane protein 2 (*PRRT2*) cause a range of episodic disorders that include paroxysmal kinesigenic dyskinesia (PKD) and benign familial infantile epilepsy (BFIE). Mutations are generally loss of function and include the c649dupC frameshift mutation that is present in around 80% of affected individuals. To investigate how *Prrt2* loss of function mutations cause disease, we performed a phenotypic investigation of a transgenic *Prrt2* knockout (*Prrt2* KO) mouse. We observed spontaneous paroxysmal episodes with behavioural features of both seizure and movement disorders, as well as unexpected deaths in KO and HET animals. KO mice showed spatial learning deficits in the Morris water maze, as well as gait abnormalities in the quantitative Digigait analysis; both of which may be representative of the more severe phenotypes experienced by homozygous patients. These findings extend the described phenotypes of *Prrt2* mutant mice, further confirming their utility for *in vivo* investigation of the role of *Prrt2* mutations in episodic diseases.

Introduction

Heterozygous mutations in proline-rich transmembrane protein 2 (*PRRT2)* cause a range of episodic neurological disorders, most commonly benign familial infantile epilepsy (BFIE), paroxysmal kinesigenic dyskinesia (PKD) and infantile convulsions with choreoathetosis (ICCA) [1-3]. BFIE is an infantile epileptic seizure disorder with an onset at 3-12 months and offset around 2 years of age, without developmental impediment. In contrast, PKD is a movement disorder characterised by short bursts of hyperkinetic movement. Onset occurs in adolescence and symptoms persist into adulthood. ICCA is defined by the occurrence of BFIE and PKD in the same patient. *PRRT2* mutations have also been implicated in other paroxysmal conditions, including other movement disorders, hemiplegic migraine and febrile seizures [4-8]. A number of patients with homozygous mutations in *PRRT2* have also been reported and show increased severity of episodic disorders, along with intellectual disability [9-11]. The most common disease-causing mutations result in frameshift and premature termination (for example c649dupC which occurs in almost 80% of patients) indicating a likely loss of protein function [12].

Following its association with disease, investigations into the function of PRRT2 have revealed a primarily CNS-restricted protein, with a role in synaptic function [1, 13]. PRRT2 spans 340 amino acids and contains two hydrophobic domains at the C-terminus, resulting in a membrane-bound protein of which the majority is intracellular [1, 14]. It is expressed throughout the brain, with strongest expression in the cerebellum, hippocampus and cortex [1, 15, 16]. PRRT2 interacts with SNARE complex proteins SNAP-25, STX1A and VAMP2, as well as Ca^{2+} sensors Syt1/2, all of which are important regulators of synchronous neurotransmitter release at the pre-synaptic terminal [13, 17, 18]. A study in homozygous patient-derived iPSCs revealed that PRRT2 also directly

binds and negatively regulates voltage-gated Na^+ channels $Na_v1.2$ and $Na_v1.6$, resulting in increased Na^+ currents in homozygous cells [19]. *In vitro* knockdown in primary neurons results in synaptic defects; namely decreased synaptic contacts, reduced synchronous neurotransmitter release and an insensitivity to extracellular Ca^{2+} [18]. A role for PRRT2 in inhibition of vesicle fusion at the synapse has also been described [20]. Silencing of PRRT2 *in vivo* causes delayed neuronal migration during development and decreased dendritic spine density [21].

Recently, two groups have described mouse models lacking PRRT2 protein. Both observed paroxysmal phenotypes including spontaneous PKD-like behaviours, motor abnormalities, audiogenic movement episodes and increased seizure susceptibility/severity following stimulation with pentylentetrazol (PTZ) [15, 16]. In this study, we extend the phenotypic investigation of the *Prrt2* knockout (*Prrt2* KO) mouse as a model for *Prrt2*-related disorders. We find that similar to patients with *PRRT2* mutations, *Prrt2* KO mice display a spontaneous movement disorder and seizure phenotype. We also observe premature death in both heterozygous (HET) and KO animals, along with learning deficiencies and gait abnormalities in the latter. These findings further validate this mouse model as a valuable experimental tool for investigating the underlying mechanisms of *PRRT2* mutations in a wide range of paroxysmal diseases.

Materials and methods

Animal housing and genotyping

All animal work was conducted following approval by The University of Adelaide Animal Ethics Committee (approval numbers S-2013-201, S-2014-148, S-2016-136) in accordance with the Australian code for the care and use of animals for scientific purposes. Nine heterozygous *Prrt2* founders were provided by the Australian Phenomics Network at Monash University (Melbourne, Australia) and maintained on a C57BL/6N genetic background. Animals were housed in an animal facility at the University of Adelaide on a 12-hour light/dark cycle and provided water and food *ad libitum* with meat free rat and mouse diet (Specialty Feeds, Western Australia). Genomic DNA for routine genotyping was obtained from tail tips or ear notches at weaning (3 weeks) and amplified using a multiplex PCR. The common forward primer 5'-CTGGTGCGCCTTCGAGTTGG-3' amplifies with Primer 2 5'-CTGGGACCTTCTGTCTGG-3' for a WT allele or with Primer 3 5'-CAACGGGTTCTTCTGTTAGTCC-3' for a transgenic allele. For amplification of inside the transgene to 3' outside the transgene, the forward primer 5'-GTCTGAGCTCGCCATCAGTT-3' and reverse primer 5'-GACAGCATTGAGACGTGAGC-3' were used.

RT-PCR

RNA was extracted from brain tissue using Trizol reagent (Thermofisher Scientific). RNA was converted to cDNA using high capacity RNA-cDNA kit (Thermofisher Scientific). For amplification of lacZ forward primer 5'- TACGATGCGCCCATCTACAC-3' and reverse primer 5'-AACAACCCGTCGGATTCTCC-3' were used. For amplification of the *Prrt2* ORF, forward primer 5'-

CTGTCAACATTGTGGCCTTC-3' and reverse primer 5'- GCAAAAGTGCAGGGAGAAAG-3' were used.

Protein extraction and western blotting

Brains were homogenized in extraction buffer (10mM Tris-HCl [pH 7.2], 150mM NaCl, 5mM EDTA, 1% Triton-X-100, 1% SDS, 1% Deoxycholate) with EDTA-free Protease Inhibitor Cocktail (Roche) and incubated at 4°C for 30 minutes. For chemical fractionation, Syn-PER reagent (Thermo Fisher Scientific) was used as per manufacturer's instructions. Lysates were separated on Invitrogen Bolt precast 4-12% polyacrylamide gels and transferred to PVDF membrane before blotting. Membranes were blocked for 1 hour in blocking solution (5% BSA, 5% skim milk) at room temperature and then incubated in primary antibody/blocking solution overnight at 4°C. Primary antibodies and their corresponding dilutions were: rabbit anti-PRRT2 (1/2000, Sigma #HPA014447), rabbit anti-Syt1 (1/1000, Synaptic Systems #105103), rabbit anti-Syt2 (1/1000, Synaptic Systems #105123), rabbit anti-β-actin (1/1000, Cell Signalling #4967), mouse anti-PSD95 (1/2000, Thermo Fisher Scientific #MA1046) and rabbit anti-synaptophysin (1/10000, Abcam #ab32127). Membranes were washed 3x in TBST and incubated for 2hrs at RT in secondary antibody conjugated with HRP/blocking solution at 1/400 dilution. Blots were washed and then imaged by chemiluminescence. For densitometric quantification, three biological replicates were performed and quantified using Biorad Chemidoc with Image Lab software. Each band was normalized to its β-actin control, then to the relevant WT band. One-way ANOVA was performed to analyse statistical differences using GraphPad Prism 7 software (La Jolla, California, USA).

TEM

3-4week old mice were trans-cardially perfused with ice-cold PBS followed by ice-cold 4% PFA (n=3 WT, n=3 KO). The brains were then dissected and 1mm pieces of cortex fixed in TEM fixative (4% Sucrose, 4% Paraformaldehyde, 1.25% Gluteraldehyde in PBS). Tissue was postfixed for 1 h with osmium tetraoxide (2% aqueous solution) before dehydration in increasing ethanol concentrations, washing in propylene oxide and then embedding in Procure-812. Tissue was cut into ultra-thin sections (80nm), stained with uranyl acetate and lead citrate and imaged on the FEI Tecnai G2 Spirit transmission electron microscope. Images were analysed using Image J software and two-tailed unpaired t-test used for statistical comparison using GraphPad Prism 7 software.

Nocturnal monitoring

Nocturnal monitoring was performed by recording with an infrared camera mounted above a Perspex-covered cage in which mice had free access to food and hydration gel. Each mouse was recorded for 12 hours. Video footage was reviewed for overt behavioural phenotypes.

PTZ assay

Seizure susceptibility was measured by subcutaneous injection of pentylenetetrazole (PTZ) (120 mg/kg). *Prrt2* WT, HET and KO mice (n=5 per genotype) were observed individually in clear cylinders for drug-induced seizure phenotypes culminating in tonic-clonic seizure, at which point latency to seizure was recorded.

Behavioural Testing

Behavioural testing was carried out at the Florey Institute Behaviour Core facility (Melbourne, Australia). 8-week old mice (n=10 WT and n=14 KO) were taken through a series of behavioural tests (locomotor, rotarod, Y-maze, digigait, Morris water maze and reversal) over a number of weeks.

Locomotor

27.5cm square clear Perspex locomotor arenas (Tru-Scan) were used for testing with 16 IR beams on the lower sensor ring to detect floor plane movement and 16 on an elevated ring to detect rearing. Locomotive activity and rearing behaviour was recorded by the IR beam sensors over 30 minutes.

Digigait

Animals were placed inside the testing chamber positioned on a clear Perspex treadmill belt. The treadmill was started at a low speed (10cm/sec) until the animal was comfortable walking on the treadmill. The speed was slowly increased to 15, 20 or 25cm/s and then approximately 15 steps captured before returning the mouse to its home cage.

Rotarod

Mice were pre-trained on the rotarod (Ugo Basil) for 3, 5-minute trials. The first 2 trials were run at a constant speed of 4 rpm, and the 3rd had an accelerating speed ranging from 4 - 40 rpm over 5 min. The testing phase was carried out the following day and involved 3 trials. Each

trial consists of the accelerating setting (4 – 40 rpm) across the 5-minute trial. An inter-trial interval of ~30 mins was included for each animal.

Spontaneous alternation using Y-maze

Testing took place in a Y-shaped maze with three Perspex arms at a 120° angle from each other. After introduction to the centre of the maze, the animal was allowed to freely explore the three arms. The number of arm entries and the number of triads were recorded in order to calculate the percentage of alternation. An entry was scored when all four limbs were within the arm.

Morris Water Maze and reversal

The 1.2m diameter water maze pool was filled to a depth of 40cm with 22 degree C water. This leaves the 15cm diameter submerged platform 1cm below the water level. Approximately 500ml of non-toxic white paint was added to the water to make it opaque and hide the platform from the animal. The mouse was then placed in the pool at one of the cardinal points (N, E, S, W) in a randomized fashion and allowed 1 minute to find the platform. If the mouse found the platform within this time it was allowed 10 seconds on the platform before it was removed, gently towelled down and placed under a warming lamp. If the mouse did not find the platform within the 2 minutes, it was led to the platform by trailing a hand in the water in front of the mouse. It was then allowed 10 seconds on the platform before it was removed, towelled down and placed under a warming lamp. The routine was repeated 4 times per day for 6 days with an interval time between swims of 20-30 minutes, until the mouse had clearly learnt the maze, signified by no significant improvement occurring after 3 consecutive days.

Statistical Analysis

Rotarod, morris water maze and reversal data were assessed using linear mixed effects models. Where data violated distributional assumptions, it was transformed to geometric means for analysis, however arithmetic means are reported. For tests with maximal time constraints (censored data), linear mixed effects tobit models were used for statistical analysis. Locomotor data was analysed using the Wilcoxin test. Linear mixed effects models and Wilcoxin tests were completed using SAS v9.4 (SAS Institute, Cary, NC, USA). Digigait parameters were analysed using two-way ANOVA and post-hoc Sidak tests. Unpaired two-tailed students t-tests were used for analysis of the continuous Y-maze. PTZ survival plots were analysed using the Mantel-Cox test. These analyses were performed using GraphPad Prism 7 software.

Results

Prrt2 KO mice

Prrt2 KO mice were generated using pre-targeted ES cells purchased from KOMP (part of the International Knockout Mouse Consortium - IKMC) by ES2M/Australian Phenomics Network Facility at Monash University (Melbourne, Australia). They carry a 'knockout first' allele, encoding an engrailed (En2) splice acceptor and IRES-LacZ transgene between exon 1 and 2 of the *Prrt2* locus, which is preferentially spliced resulting in *PRRT2* knockout (Supplementary Figure 1a). We first established that homozygous (KO) and heterozygous (HET) *Prrt2* mice expressed the IRES-LacZ transgene and that KO animals lack the *Prrt2* ORF using RT-PCR (Supplementary Figure 1b). Western blot analysis on adult whole brain extract confirmed reduced PRRT2 protein in HET and an absence in *Prrt2* KO animals (Supplementary Figure 1c). To confirm the location of the transgene, genomic DNA from heterozygous founders was PCR amplified using a primer outside of the transgene homology arms to within (Supplementary Figure 1d).

Chemical fractionation on adult brain samples confirmed that PRRT2 is present in the synaptosome but not the cytosolic fraction (Supplementary Figure 2a). Given that PRRT2 has been shown to interact with SYT1/2, and that knockdown of PRRT2 in primary neurons results in lower SYT1 and 2 levels[18], we assessed SYT1/2 protein levels in brain lysates by western blot. No significant change in SYT1 or SYT2 was detected in the brains of *Prrt2* HET or KO mice. The levels of other synapse proteins, synaptophysin and PSD95 were similarly unaffected by reduction or absence of PRRT2 (Supplementary Figure 2b). Transmission electron microscopy (TEM) analysis of synapses in wild type (WT) and KO cortex showed no significant difference in

synapse density (p=0.1824; Supplementary Figure 2c) or active zone length (p=0.0781; Supplementary Figure 2d).

Spontaneous paroxysms and sudden unexplained death in Prrt2 mice

PRRT2 disorders are characterised by spontaneous paroxysms, which manifest as epileptic seizures and/or a movement disorder. Consistent with these clinical features, we observed spontaneous paroxyms in *Prrt2* KO mice. These events were observed during routine handling in a small proportion of *Prrt2* KO animals, without an obvious trigger (2.53%, Figure 1). Episodes lasted for varying lengths of time and affected mice across a range of ages (5 weeks to 8 months). Episodes began with dyskinesia and were sometimes followed by secondary generalized seizures (which was accompanied by loss of awareness; Supplementary Video 1+ 2). To determine whether spontaneous paroxysms were occurring during routine housing, we performed infrared video monitoring of a small cohort during the dark cycle (the active period for mice). Overt behavioural phenotypes were not detected in either HET or KO mice (data not shown). We also assessed seizure susceptibility by performing a PTZ assay on a cohort of WT, HET and KO adult mice (n=5 per genotype) but observed no significant difference in latency to tonic-clonic seizure (Supplementary Figure 3).

Interestingly, *Prrt2* KO and HET mice also displayed a marked survival defect with 7.69% of KO and 1% of HET mice found dead in their cages, without any clear cause of death (Figure 1). Both the spontaneous paroxysms and unexplained deaths were not observed in WT animals from the same colony, confirming *Prrt2* mutation as the cause of these phenotypes.

Gait alteration in Prrt2 KO mice

To assess motor activity and coordination in *Prrt2* KO mice, we performed locomotor, digigait and rotarod behavioural testing. Given that phenotypic severity is greater in *PRRT2* homozygous patients, we compared *Prrt2* KO (n=14) and WT (n=10) mice. Locomotor analysis showed no overall difference in movement or activity between *Prrt2* KO and WT (Supplementary Figure 4a). No significant difference was observed between groups of animals on the Rotarod test, indicating that *Prrt2* mice do not suffer from coordination issues (Supplementary Figure 4b).

Quantitative comparison of gait was performed using digigait analysis across three speeds (15, 20 and 25cm/s). Interestingly, significant differences were detected between WT and KO mice across most aspects of gait including stride, swing, stance and brake. Stride length was significantly lower, whilst number of steps and step frequency were significantly higher in KO animals (Figure 2, Table 1). To further define these differences, *post hoc* Sidak multiple comparison tests were performed. This analysis revealed significant differences between WT and KO mice at 15cm/s and to a large extent at 20cm/s, but not at 25cm/s (Supplementary Table 1). To rule out the possibility that animal size was the cause of the differences, we compared animal length between WT and KO animals and saw no significant difference ($p=0.3671$, Supplementary Figure 4c). Altogether, digigait analysis on *Prrt2* KO mice show a comprehensive series of abnormalities across many components of gait.

Prrt2 KO mice display learning deficiencies

Given that intellectual disability is observed in homozygous patients, we investigated the cognitive abilities of KO mice. When tested on the continuous Y-maze, there was no significant difference between *Prrt2* KO and WT animals in percentage alternation between arms (p=0.1314; Supplementary Figure 5a).

In the Morris Water Maze, *Prrt2* KO mice displayed significantly longer latency to reach the platform than WT (p=0.014, Figure 3a). *Post hoc* analysis confirmed an overall genotype difference that was independent of trial day. When we measured distance travelled and velocity in Morris Water Maze, no differences were observed between KO and WT, indicating that differences observed were due to cognitive deficits rather than reduced or slower swimming ability (Supplementary Figure 5b-e). When spatial memory was tested during the probe trial, there was no significant difference in target quadrant time between genotypes (p=0.083, Figure 3c). Repeating the Morris Water maze with reversal of the platform resulted in no differences between genotypes in latency to reach the platform across 5 days nor in target quadrant time in a reversal probe test (p=0.906 or 0.693 respectively, Figure 3d,e). Together these data suggest learning defects in *Prrt2* KO mice, in the absence of memory deficits.

Discussion

Heterozygous loss of function mutations in *PRRT2* cause a broad spectrum of episodic diseases, most frequently BFIE and PKD. Analysis of gene function using knockdown techniques has revealed a synaptic function for the protein, as well as a possible role in neuronal development[18, 21].

In order to investigate the impact of *Prrt2* mutations *in vivo* we undertook phenotypic analysis of a *Prrt2* KO mouse. Consistent with a recent report on the same transgenic model, we observed no detectable protein or transcript in the brains of KO animals, with reduced levels in heterozygous mutants [15]. PRRT2 localised to the synapse as previously reported [13, 16, 18, 21], though its reduction or knockout did not change levels of SYT 1/2 *in vivo* as observed in PRRT2 knockdown neurons [18]. When we assessed cortical synaptic density by TEM, we saw no significant differences between WT and *Prrt2* KO brains. The same observation was recently described in the cerebellum of a *Prrt2*^{STOP} mouse that models the recurrent c649dupC mutation in patients [16]. Interestingly, Michetti et al reported decreased synaptic contacts in the dentate gyrus (DG) but not in two regions of the cerebellum. Similarly, they observed that mRNA expression levels of most synaptic markers were not altered in the cortex of *Prrt2* HET or KO animals [15]. This suggests that although *Prrt2* is expressed broadly throughout the brain, it may be playing a different role in the DG compared to other brain regions.

We describe a series of behavioural phenotypes in *Prrt2* KO mice that model clinical phenotypes observed in humans. Most notably, spontaneous paroxysms were detected in *Prrt2* KO mice, albeit at low penetrance. A similar phenotype was described by Tan *et al,* with low frequency spontaneous dyskinetic attacks observed in *Prrt2*^{STOP} mice [16]. Consistent with a

previous analysis of the *Prrt2* KO model, when we probed the seizure susceptibility by PTZ injection, we observed no difference in latency to tonic-clonic seizure for both *Prrt2* HET and KO mice, though the previous study observed more severe seizures once induced [15]. In contrast, increased seizure susceptibility was observed in *Prrt2*^{STOP} mice using both a PTZ and febrile seizure assay [16]. The models are maintained on different C57BL/6 substrains, which have genotypic and phenotypic differences that may impact on performance in behavioural tests[22]. We also identified unexplained death in a subset of *Prrt2* mutant (HET and KO) animals (Figure 1). Though not a common feature of *PRRT2* mutations in humans, a heterozygous patient with probable sudden unexpected death in epilepsy (SUDEP) has been reported [23]. Other murine models of genetic epilepsies have also been identified with a premature death phenotype, although it is not clear if the primary cause of death is related [24, 25].

Whilst most *PRRT2*-related disorders are paroxysmal in nature, homozygous patients experience increased severity of disease, along with additional intellectual and learning disabilities [10, 11]. Consistent with these phenotypes, we observed learning defects in *Prrt2* KO mice when testing their spatial learning in the Morris Water Maze (Figure 3b). However, working memory and memory plasticity were unaffected as evidenced by the probe test and platform reversal [26]. Rotarod and locomotor tests did not reveal a motor or coordination phenotype for the *Prrt2* KO mice in our study, despite the coordination differences described in *Prrt2*^{STOP} mice [16]. *Prrt2* KO mice also displayed gait abnormalities when assessed on the Digigait software. These abnormalities span many fundamental gait parameters and manifested when travelling at the slowest treadmill speed (15cm/s) during digigait analysis (Table 1, Supp Table 1). Interestingly, these differences were not as pronounced at the faster speed (20cm/s) and were virtually absent at the fastest speed (25cm/s). Differences in gait can

be speed-dependent, with faster speeds (akin to running) masking phenotypes observed at slower (walking) speeds (unpublished observations, T. Featherby). Previously reported footprint analysis of *Prrt2* mice showed no differences between genotypes, however many aspects of gait that differed here were not tested and mice were not controlled for speed [15, 16]. Movement disorders experienced by patients with *PRRT2* mutations are typically episodic in nature, making the finding of overall gait differences somewhat unexpected. We cannot rule out the possibility that subtle movement episodes were experienced by a subset of KO animals during testing, resulting in the overall gait deficit attributed to the group, however given the lack of significant differences at the fastest speed, this seems unlikely. It remains to be seen how *Prrt2* HET animals perform in these same behavioural experiments. However, we expect that learning and gait phenotypes will not be present, given that *Prrt2* heterozygosity in humans most commonly results in epileptic or motor paroxysms with an absence of other phenotypes.

Together, our phenotypic investigation of the *Prrt2* KO mouse revealed a paroxysmal phenotype in KO, along with premature death affecting HET and KO animals. Wide-ranging gait abnormalities and learning deficits were also observed in KO mice. These data support *Prrt2* KO mice as a useful model for investigating *Prrt2* function and elucidation of disease mechanisms underlying the *PRRT2*-related disease spectrum.

Acknowledgments

We would like to acknowledge the Adelaide Microscopy facility, particularly Ruth Williams, for assistance with and provision of the transmission electron microscope, as well as the Monash University ES2M facility for generating the *Prrt2* KO mouse model. We would also like to thank Kay Richards and Chris Reid for expert interpretations of the behavioural manifestations of the paroxysmal events.

Table 1: Genotype comparison across gait parameters from Digigait testing

Digigait Parameter	Stride duration		Swing		Stance		Brake		Propel		Stride Frequency		Number of Steps		Stride length		Stance Width	
Limb	Fore	Hind	Fore	Hind	Fore	Hind	Fore	Hind	Fore	Hind	Fore	Hind	Fore	Hind	Fore	Hind	Fore	Hind
WT vs KO p-value	**0.0054 ****	**0.0019 ****	0.3125 ns	**0.0449 ***	**0.0007 *****	**0.0023 ****	**0.0268 ***	**0.0013 ****	0.2663 ns	0.2076 ns	**0.0156 ***	**0.0039 ****	**0.0432 ***	**0.0099 ****	**0.0109 ***	**0.0052 ****	**0.0235 ***	0.7667 ns

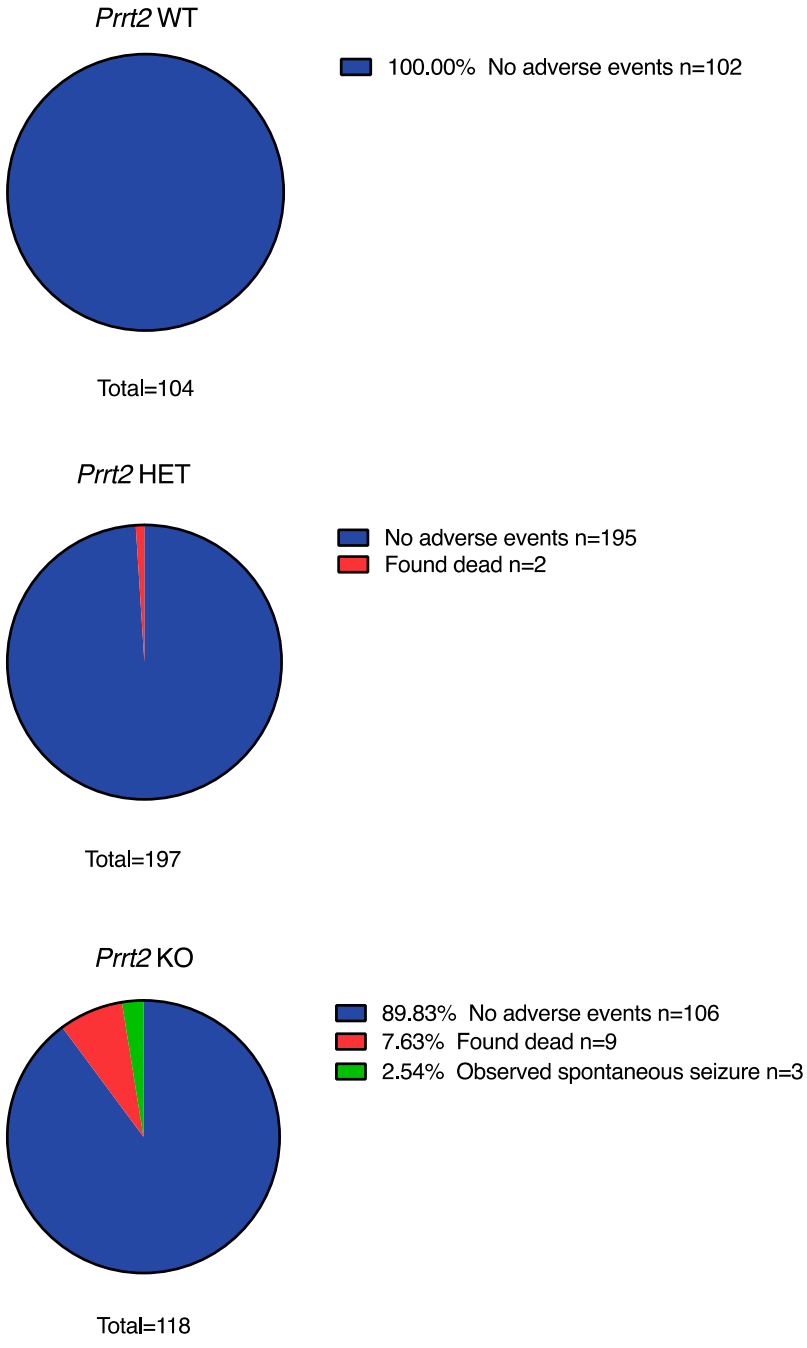


Figure 1: Proportions of *Prrt2* HET and KO animals that experienced spontaneous death and seizures. Animals that were found dead are represented by orange shading, those that experienced spontaneous paroxysms are in green and unaffected mice are indicated by blue. No adverse events were observed in WT animals from the *Prrt2* colony.

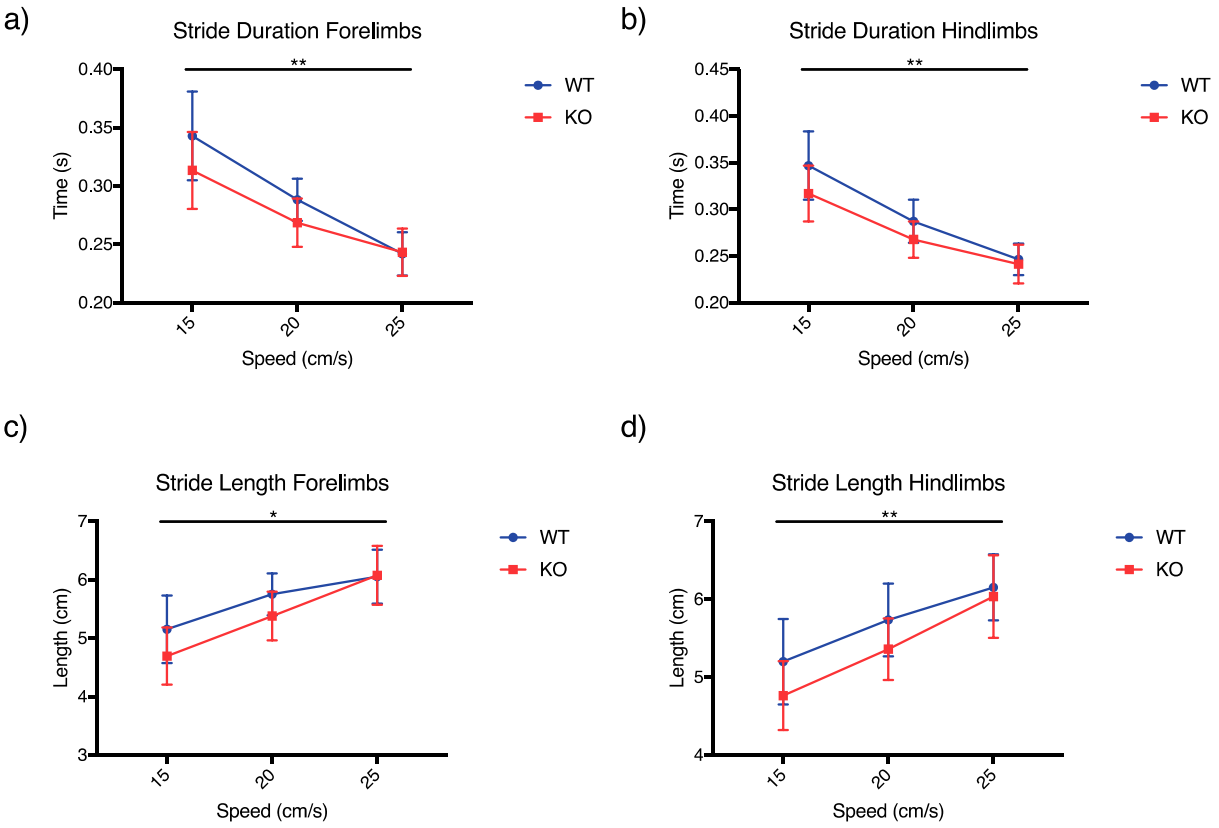


Figure 2: Gait deficiencies in *Prrt2* KO mice. (a,b) Stride time across three speeds (15, 20 & 25 cm/s) was significantly lower in KO compared to WT in both fore (p=0.0054) and hind limbs (p=0.0019). (c,d) *Prrt2* KO mice showed shorter stride length than WT in both fore (p=0.0109) and hind limbs (p=0.0052) across the three speeds. Differences between genotypes were analysed with two-way ANOVA. Error bars indicate mean ± SD. n=10 WT and n=14 KO.

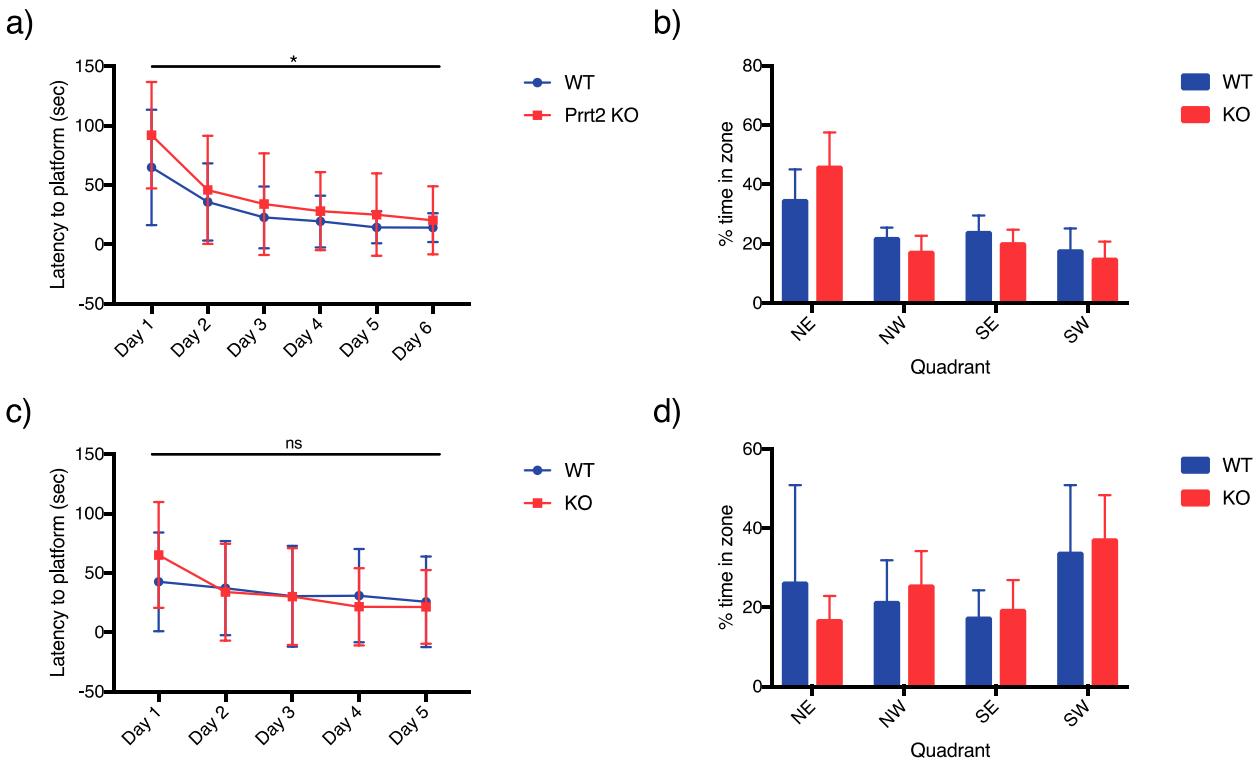


Figure 3: Learning disabilities in *Prrt2* KO mice (a) *Prrt2* KO mice show significantly longer latency to platform in the Morris Water Maze (MWM) training days (p=0.014). This difference was independent of day. (b) Percentage of time spent in each quadrant during the probe trial with the platform removed. The platform had previously been in the NE quadrant during training days. Differences between WT and KO were not significant (p=0.083). (c) The platform was moved to the SW quadrant for the MWM reversal. Latency to find platform across the 5 reversals days was not significant between WT and KO (p=0.906). (d) Percentage time spent in each quadrant during the reversal probe trial. Differences between WT and KO were not significant (p=0.693). MWM and reversal training day analyses were performed using a linear mixed effects tobit model. Probe data was analysed using a linear mixed effects model. Data expressed as mean ± SD. n=10 WT and n=14 KO.

References

1. Chen WJ, Lin Y, Xiong ZQ et al. Exome sequencing identifies truncating mutations in PRRT2 that cause paroxysmal kinesigenic dyskinesia. *Nat Genet.* 2011;43:1252-1255. DOI: 10.1038/ng.1008

2. Wang JL, Cao L, Li XH et al. Identification of PRRT2 as the causative gene of paroxysmal kinesigenic dyskinesias. *Brain.* 2011;134:3493-3501. DOI: 10.1093/brain/awr289

3. Heron SE, Grinton BE, Kivity S et al. PRRT2 mutations cause benign familial infantile epilepsy and infantile convulsions with choreoathetosis syndrome. *Am J Hum Genet.* 2012;90:152-160. DOI: 10.1016/j.ajhg.2011.12.003

4. Liu Q, Qi Z, Wan XH et al. Mutations in PRRT2 result in paroxysmal dyskinesias with marked variability in clinical expression. *J Med Genet.* 2012;49:79-82. DOI: 10.1136/jmedgenet-2011-100653

5. Cloarec R, Bruneau N, Rudolf G et al. PRRT2 links infantile convulsions and paroxysmal dyskinesia with migraine. *Neurology.* 2012;79:2097-2103. DOI: 10.1212/WNL.0b013e3182752c46

6. Gardiner AR, Bhatia KP, Stamelou M et al. PRRT2 gene mutations: from paroxysmal dyskinesia to episodic ataxia and hemiplegic migraine. *Neurology.* 2012;79:2115-2121. DOI: 10.1212/WNL.0b013e3182752c5a

7. Marini C, Conti V, Mei D et al. PRRT2 mutations in familial infantile seizures, paroxysmal dyskinesia, and hemiplegic migraine. *Neurology.* 2012;79:2109-2114. DOI: 10.1212/WNL.0b013e3182752ca2

8. Scheffer IE, Grinton BE, Heron SE et al. PRRT2 phenotypic spectrum includes sporadic and fever-related infantile seizures. *Neurology.* 2012;79:2104-2108. DOI: 10.1212/WNL.0b013e3182752c6c

9. Ebrahimi-Fakhari D, Saffari A, Westenberger A et al. The evolving spectrum of PRRT2-associated paroxysmal diseases. *Brain.* 2015;138:3476-3495. DOI: 10.1093/brain/awv317

10. Labate A, Tarantino P, Viri M et al. Homozygous c.649dupC mutation in PRRT2 worsens the BFIS/PKD phenotype with mental retardation, episodic ataxia, and absences. *Epilepsia.* 2012;53:e196-199. DOI: 10.1111/epi.12009

11. Delcourt M, Riant F, Mancini J et al. Severe phenotypic spectrum of biallelic mutations in PRRT2 gene. *J Neurol Neurosurg Psychiatry.* 2015;86:782-785. DOI: 10.1136/jnnp-2014-309025

12. Heron SE, Dibbens LM. Role of PRRT2 in common paroxysmal neurological disorders: a gene with remarkable pleiotropy. *J Med Genet.* 2013;50:133-139. DOI: 10.1136/jmedgenet-2012-101406

13. Lee HY, Huang Y, Bruneau N et al. Mutations in the gene PRRT2 cause paroxysmal kinesigenic dyskinesia with infantile convulsions. *Cell Rep.* 2012;1:2-12. DOI: 10.1016/j.celrep.2011.11.001

14. Rossi P, Sterlini B, Castroflorio E et al. A Novel Topology of Proline-rich Transmembrane Protein 2 (PRRT2): hints for an intracellular function at the synapse. *J Biol Chem.* 2016;291:6111-6123. DOI: 10.1074/jbc.M115.683888

15. Michetti C, Castroflorio E, Marchionni I et al. The PRRT2 knockout mouse recapitulates the neurological diseases associated with PRRT2 mutations. *Neurobiol Dis.* 2017;99:66-83. DOI: 10.1016/j.nbd.2016.12.018

16. Tan GH, Liu YY, Wang L et al. PRRT2 deficiency induces paroxysmal kinesigenic dyskinesia by regulating synaptic transmission in cerebellum. *Cell Res.* 2017. DOI: 10.1038/cr.2017.128

17. Valtorta F, Benfenati F, Zara F et al. PRRT2: from Paroxysmal Disorders to Regulation of Synaptic Function. *Trends Neurosci.* 2016;39:668-679. DOI: 10.1016/j.tins.2016.08.005

18. Valente P, Castroflorio E, Rossi P et al. PRRT2 Is a Key Component of the Ca(2+)-Dependent Neurotransmitter Release Machinery. *Cell Rep.* 2016;15:117-131. DOI: 10.1016/j.celrep.2016.03.005

19. Fruscione F, Valente P, Sterlini B et al. PRRT2 controls neuronal excitability by negatively modulating Na+ channel 1.2/1.6 activity. *Brain.* 2018;141:1000-1016. DOI: 10.1093/brain/awy051

20. Coleman J, Jouannot O, Ramakrishnan SK et al. PRRT2 Regulates Synaptic Fusion by Directly Modulating SNARE Complex Assembly. *Cell Rep.* 2018;22:820-831. DOI: 10.1016/j.celrep.2017.12.056

21. Liu YT, Nian FS, Chou WJ et al. PRRT2 mutations lead to neuronal dysfunction and neurodevelopmental defects. *Oncotarget.* 2016;7:39184-39196. DOI: 10.18632/oncotarget.9258

22. Simon MM, Greenaway S, White JK et al. A comparative phenotypic and genomic analysis of C57BL/6J and C57BL/6N mouse strains. *Genome Biol.* 2013;14:R82. DOI: 10.1186/gb-2013-14-7-r82

23. Labate A, Tarantino P, Palamara G et al. Mutations in PRRT2 result in familial infantile seizures with heterogeneous phenotypes including febrile convulsions and probable SUDEP. *Epilepsy Res.* 2013;104:280-284. DOI: 10.1016/j.eplepsyres.2012.10.014

24. Massey CA, Sowers LP, Dlouhy BJ et al. Mechanisms of sudden unexpected death in epilepsy: the pathway to prevention. *Nat Rev Neurol.* 2014;10:271-282. DOI: 10.1038/nrneurol.2014.64

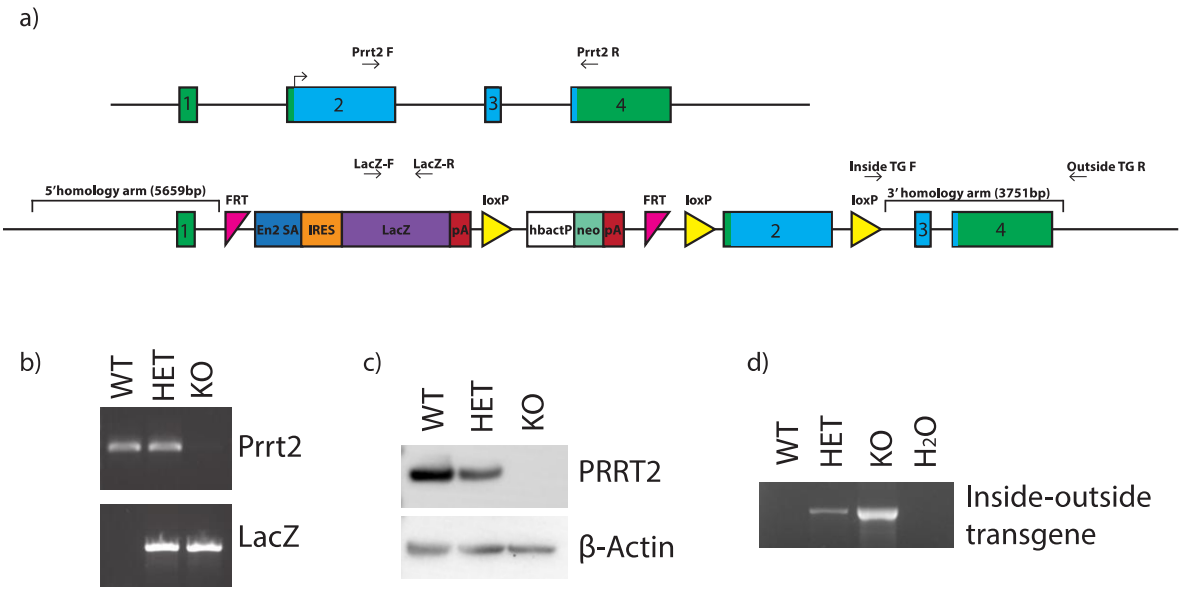
25. Palmer EE, Jarrett KE, Sachdev RK et al. Neuronal deficiency of ARV1 causes an autosomal recessive epileptic encephalopathy. *Hum Mol Genet.* 2016;25:3042-3054. DOI: 10.1093/hmg/ddw157

26. Tsetsenis T, Younts TJ, Chiu CQ et al. Rab3B protein is required for long-term depression of hippocampal inhibitory synapses and for normal reversal learning. *Proc Natl Acad Sci U S A.* 2011;108:14300-14305. DOI: 10.1073/pnas.1112237108

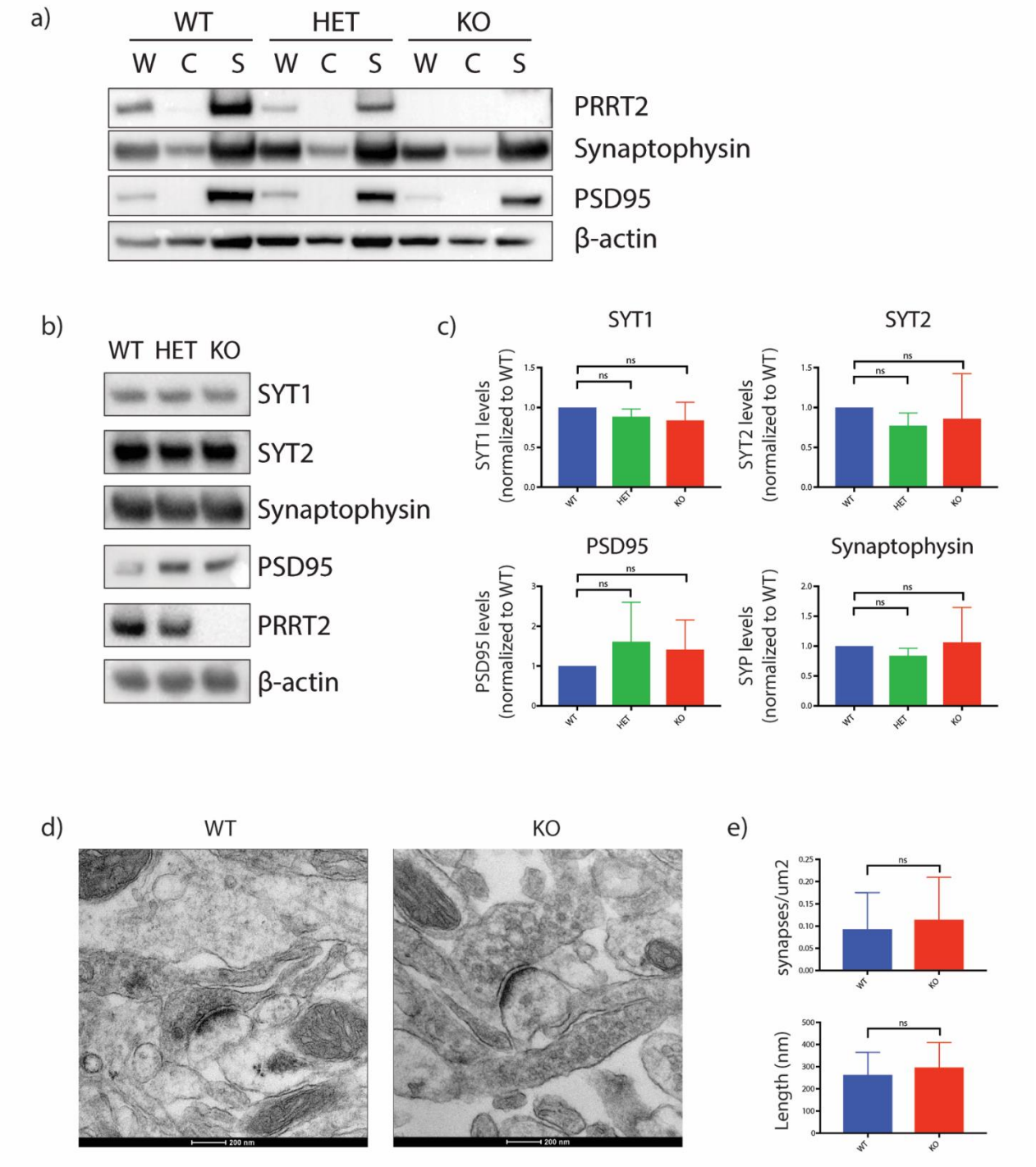
Supplementary Material

Digigait parameter	Forelimbs/ Hindimbs	Interaction p-value	WT vs KO Sidak adjusted p-value		
			Treadmill Speed		
			15cm/s	20cm/s	25cm/s
Stride Duration	**Fore**	**0.0017 ****	**0.0003 *****	**0.0253 ***	0.9962 ns
	Hind	**0.0097 ****	**0.0002 *****	**0.0214 ***	0.8475 ns
Swing	**Fore**	0.0745 ns	0.3699 ns	0.3087 ns	0.8201 ns
	Hind	**0.0126 ***	**0.0387 ***	**0.0327 ***	**09491 ***
Stance	**Fore**	**0.0014 ****	**<0.0001 ******	**0.0233 ***	0.9778 ns
	Hind	**0.0491 ***	**0.0002 *****	0.1820 ns	0.3409 ns
Brake	**Fore**	0.4309 ns	**0.0458 ***	0.7692 ns	0.3194 ns
	Hind	0.3384 ns	**0.0162 ***	0.6180 ns	**0.0200 ***
Propel	**Fore**	**0.0252 ***	0.1341 ns	0.3640 ns	0.7179 ns
	Hind	0.1325 ns	0.1187 ns	0.6421 ns	0.9997 ns
Stride Frequency	**Fore**	**0.0132 ***	0.0139 ns	**0.0128 ***	>0.9999 ns
	Hind	0.2036 ns	**0.0107 ***	**0.0126 ***	0.5089 ns
Number of Steps	**Fore**	**0.0110 ***	**0.0209 ***	0.1136 ns	0.4054 ns
	Hind	0.0640 ns	**0.0042 ****	**0.0169 ***	0.7568 ns
Stride length	**Fore**	**0.0047 ****	**0.0026 ****	**0.0190 ***	0.9972 ns
	Hind	0.0640 ns	**0.0042 ****	**0.0169 ***	0.7568 ns
Stance Width	**Fore**	0.9008 ns	0.1856 ns	0.1520 ns	0.3886 ns
	Hind	0.9937 ns	0.9857 ns	0.9956 ns	0.9938 ns

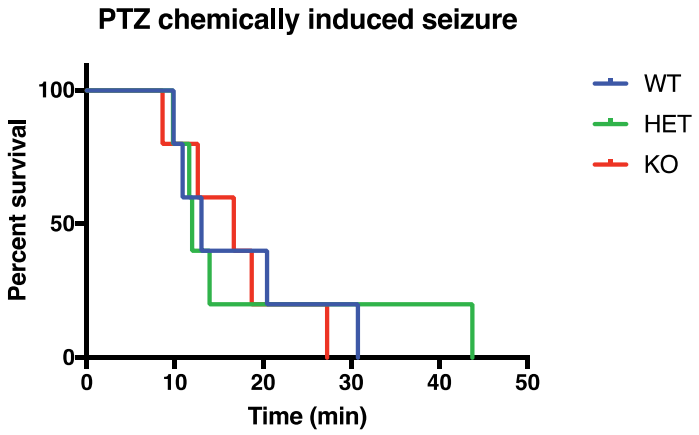
Supplementary Table 1: Genotype comparisons from two-way ANOVA analyses of Digigait parameters, showing *post hoc* Sidak test significance results across three treadmill speeds.



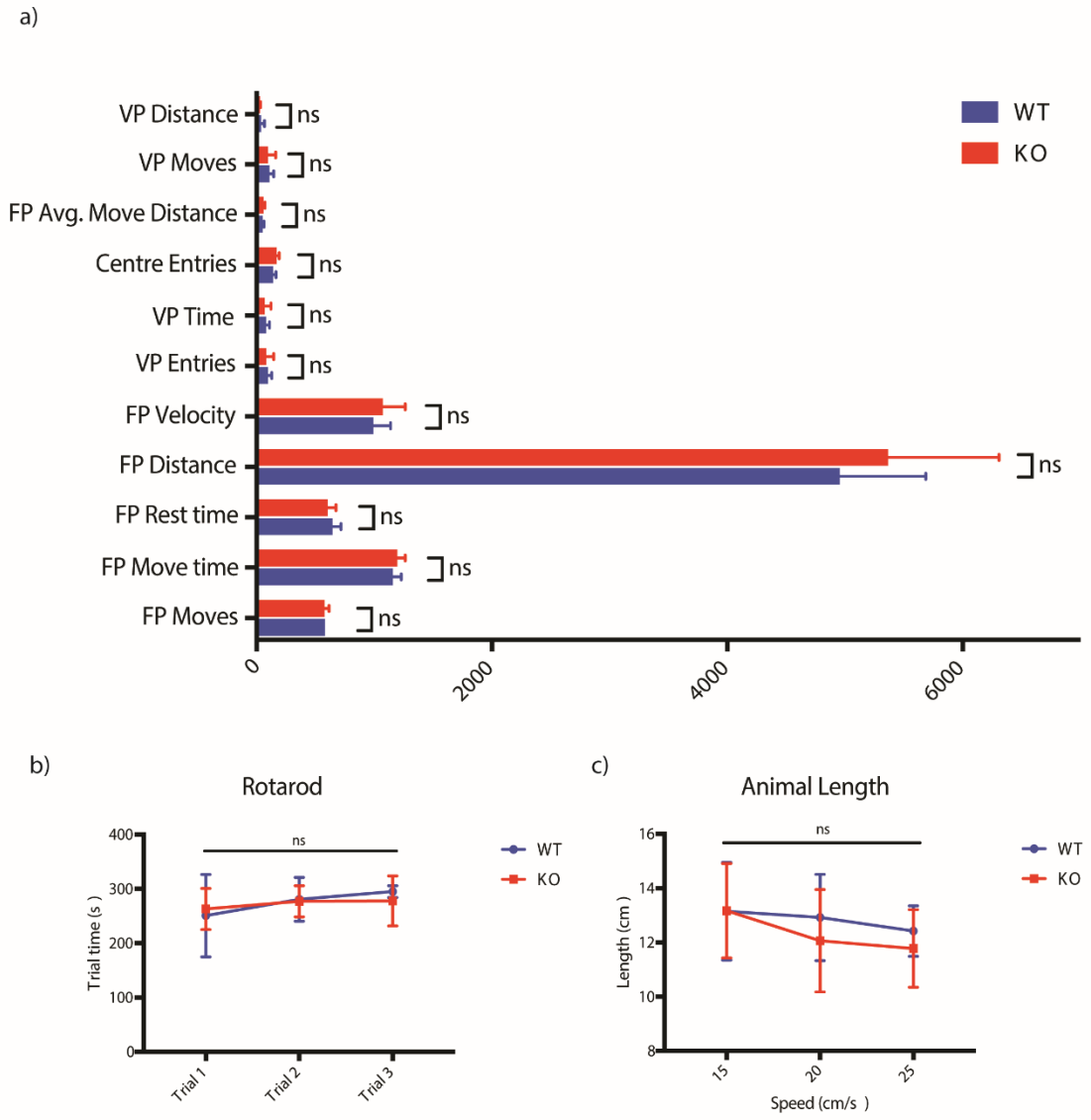
Supplementary Figure 1: Characterisation of *Prrt2* KO mice. a) Schematic of the wild type *Prrt2* and transgene alleles with primers for PCR marked above. The Engrailed 2 Splice Acceptor (En2) is preferentially spliced, resulting in transcription of the IRES-LacZ gene in place of Exon 2-4. Promoter driven neomycin resistance gene (neo) and Exon 2 are flanked by LoxP sites giving conditional potential. b) RT-PCR from Exon 2-4 of *Prrt2* confirms amplification in WT and HET, whilst LacZ amplifies only in HET and KO animals. c) Western blot analysis on adult whole brain lysates, confirming complete lack of PRRT2 in KO and reduction in HET. d) PCR analysis from inside to outside the transgene to confirm location of transgene.


a)
WT
HET
KO
W C S W C S W C S
PRRT2
Synaptophysin
PSD95
β-actin
b)
WT HET KO
SYT1
SYT2
Synaptophysin
PSD95
PRRT2
β-actin
c)
SYT1
SYT2
PSD95
Synaptophysin
SYT1 levels (normalized to WT)
SYT2 levels (normalized to WT)
PSD95 levels (normalized to WT)
SYP levels (normalized to WT)
ns
WT
HET
KO
d)
WT
KO
200 nm
e)
synapses/um2
Length (nm)

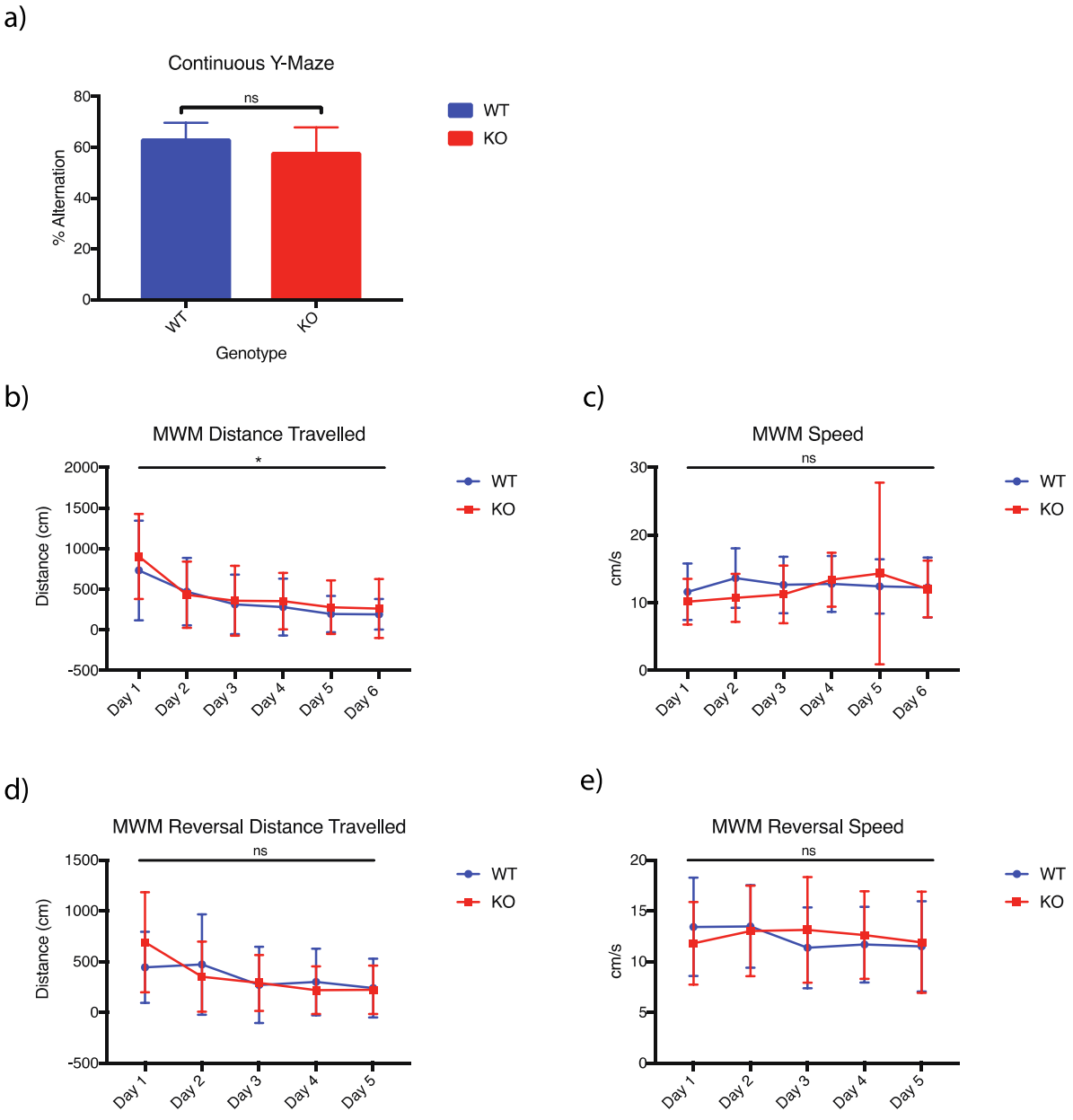
Supplementary Figure 2: Molecular and ultrastructural analysis of *Prrt2* mutant brains. a) Subcellular fractionation of adult mouse cortex, confirming localisation of PRRT2 to the synaptosome fraction, similar to synaptic markers PSD95 and Synaptophysin. W= whole cell extract, C=cytosolic fraction, S=synaptosome. n=3 biological replicates per genotype. b) Western blot analysis on adult whole brain comparing levels of Synaptotagmins 1/2 (Syt1/2), PRRT2 and synaptic markers PSD95 and Synaptophysin. c) Densitometric quantification of b) normalized to WT. All data are expressed as means $\pm$ SD and analysed using one-way ANOVA (ns=not significant). n=3 biological replicates for each genotype. d) Sample images from transmission electron microscopy on cortical sections from *Prrt2* WT and KO mice. e) Quantification of synapse number and length from n=3 biological replicates for each genotype. Data expressed as mean $\pm$ SD and analysed using unpaired t-test.



Supplementary Figure 3: *Prrt2* mice do not show increased seizure susceptibility following subcutaneous pentylenetetrazol (PTZ) injection. Data expressed as a survival curve of latency to tonic-clonic seizure and analysed using the Mantel-Cox test (p=0.9880) n=5 per genotype.



Supplementary Figure 4: Locomotor activity and coordination in Prrt2 KO mice are not significantly different to wildtype. a) WT and KO performance across locomotor parameters. Data expressed as mean ± SD and significance analysed using Wilcoxon test. VP- vertical plane, FP- Floor plane b) Latency to fall from accelerating Rotarod across three trials on consecutive days. Data expressed as mean ± SD and analysed using linear mixed effects tobit model c) Animal length during digigait testing at three different speeds. Data expressed as mean ± SD and analysed using two-way ANOVA. n=10 WT, n=14 KO.



Supplementary Figure 5: Continuous Y-Maze and Morris water maze (MWM) outcomes. a) Percentage alternation on the continuous Y-maze. Data expressed as mean ± SD and analysed using unpaired t-test. b-e) MWM and reversal distance and speed across trial days. Data expressed as mean ± SD and analysed using two-way ANOVA. n=10 WT, n=14 KO.

This page is left intentionally blank.

Chapter 3:
Expanding the RNA-guided endonuclease toolkit for mouse genome editing

3.1 Summary

As demonstrated in the previous chapter, mouse models are important tools for investigating biological questions and modelling human diseases. The advent of CRISPR/Cas technology has enabled the rapid and efficient production of mouse models through direct microinjection of genome editing agents in the zygote. Whilst the commonly used Cas9 protein from *Streptococcus pyogenes* (WT SpCas9) is incredibly efficient at introducing mutations in mice, the NGG PAM requirement restricts the targetable regions. An application particularly affected by this restriction is when introducing a point mutation via HDR, as mutation incorporation is directly affected by its distance from the double-stranded break. Variants of the CRISPR/Cas system have been described that recognise alternative PAM sequences, allowing more flexibility in targeting, though many of these have not been tested for mouse genome editing.

In this submitted manuscript, we compare a series of CRISPR variants for their ability to introduce mutations in the mouse genome editing. We tested the following endonucleases, targeted to a site in the *Ngn3* intron: SpCas9 VQR (NGAN PAM), SpCas9 VRER (NGCG PAM), AsCpf1 (TTTN PAM) and SaCas9 KKH (NNNRRT PAM), as well as the WT SpCas9 (NGG PAM). We showed efficient editing with WT SpCas9 (95%) and AsCpf1 (30%). We also validated SpCas9 VQR and SpCas9 VRER for mouse genome editing though observed lower efficiencies (11% and 6% respectively). The SaCas9 KKH variant generated mutations with high efficiency (56%), making it a useful alternative to WT SpCas9 with its flexible PAM site.

We also observed an interesting difference in SaCas9 KKH mutant embryos when compared to WT SpCas9; where WT SpCas9 would cut both alleles frequently, SaCas9 KKH left an unmodified allele 93% of the time. We tested this property whilst co-injecting with a ssDNA

template for modification through HDR. All SaCas9 KKH mediated oligo insertions were accompanied by a WT allele, whilst only 50% of embryos with oligo insertions from WT SpCas9 injections had a WT allele remaining. We propose that this property may be useful when targeting genes that cause embryonic lethality if mutated on both alleles.

We confirmed efficient introduction of mutations in a second genomic locus using SaCas9 KKH (tyrosinase gene) and showed that mutations could be carried through the germline to establish a colony. This study extends the available 'toolkit' for mouse genome editing, validating CRISPR variants that will broaden PAM recognition options. It also explores a novel property of SaCas9 KKH that may prove useful in targeting genes that cause nullizygous embryonic lethality.

Statement of Authorship

Title of Paper	Expanding the RNA-guided endonuclease toolkit for mouse genome editing
Publication Status	☐ Published ☐ Accepted for Publication ☑ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Submitted to the CRISPR Journal.

Principal Author

Name of Principal Author (Candidate)	Louise Robertson
Contribution to the Paper	Conceived, designed and performed experiments, generated reagents, analysed data and wrote manuscript.
Overall percentage (%)	85%
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date 26/4/18

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Daniel Pederick
Contribution to the Paper	Performed and analysed cloning experiments, conceived HDR experiment, revised manuscript.
Signature	Date 4/30/18

Name of Co-Author	Marie Ahladas
Contribution to the Paper	Performed a portion of embryo screening from Tyrosinase breeding and HDR experiment, revised manuscript.
Signature	Date 26/04/18

Name of Co-Author	Melissa White		
Contribution to the Paper	Performed mouse zygote injections, revised manuscript.		
Signature		Date	26/4/18

Name of Co-Author	Sandra Piltz		
Contribution to the Paper	Performed mouse zygote injections, revised manuscript.		
Signature		Date	26/04/18

Name of Co-Author	Fatwa Adikusuma		
Contribution to the Paper	Conceived and designed study, evaluated and edited manuscript.		
Signature		Date	26/04/18

Name of Co-Author	Paul Thomas		
Contribution to the Paper	Conceived, designed and supervised study, evaluated and edited manuscript.		
Signature		Date	26/4/18

Expanding the RNA-guided endonuclease toolkit for mouse genome editing

Louise Robertson[1], Daniel Pederick[1], Sandra Piltz[1,2], Melissa White[1,2], Marie Ahladas[3], Fatwa Adikusuma[1], Paul Thomas[1,2,3*]

[1]Department of Molecular and Cellular Biology and Robinson Research Institute, University of Adelaide, Adelaide, SA, Australia, 5005

[2]SA Genome Editing Facility, University of Adelaide, Adelaide, SA, Australia, 5005

[3]South Australia Health and Medical Research Institute, Adelaide, SA, Australia, 5000

***Corresponding author:** Paul Thomas

Louise Robertson	+61883135009	louise.robertson@adelaide.edu.au
Daniel Pederick	+16507255809	pederick@stanford.edu
Sandra Piltz	+61883135009	sandra.piltz@adelaide.edu.au
Melissa White	+61883135009	melissa.white@adelaide.edu.au
Marie Ahladas	+61881284087	marie.ahladas@sahmri.com
Fatwa Adikusuma		fatwa.adikusuma@adelaide.edu.au
Paul Thomas	+61883135009	paul.thomas@adelaide.edu.au

Abstract

The RNA-guided endonuclease CRISPR/Cas system from *Streptococcus pyogenes* (SpCas9) is widely used for generating genetically modified mice via zygote micro-injection. Although double-strand breaks (DSBs) are induced efficiently by SpCas9, it requires an NGG proto-spacer adjacent motif (PAM) at the target site, restricting sequence targetability. Recently, alternative RNA-guided endonucleases that utilise different PAM sequences have been characterised. Here we tested the following endonucleases for their ability to edit the mouse genome: SpCas9 VQR (NGAN PAM), SpCas9 VRER (NGCG PAM), AsCpf1 (TTTN PAM) and SaCas9 KKH (NNNRRT PAM), alongside wild type (WT) SpCas9 (NGG PAM). SaCas9 KKH generated mutations in up to 70% of embryos. Remarkably, in contrast to WT SpCas9, this variant left one allele unmodified in the vast majority of mutant embryos (6% versus 93%). Our findings not only broaden the PAM recognition options for mouse genome editing, but also identify SaCas9 KKH as an attractive alternative when targeting genes that cause null lethal phenotypes.

Introduction

Genetically modified mice are widely used for investigating gene function, disease modelling and therapeutic development. The traditional method for generating targeted modifications in mice involves creation of mutant ES cells, injection of modified ES cells into blastocysts and breeding of chimeric founders [1]. While this approach has proven useful, it is costly, time consuming and prone to failure due to the reliance on germline transmission. The development of CRISPR/Cas9 genome editing technology provides an alternative approach for the generation of mutant mice through direct modification of the zygotic genome [2-4]. This approach is efficient, fast and cheap, and has largely replaced traditional gene targeting.

The most commonly used Cas9 endonuclease is derived from *Streptococcus pyogenes* (SpCas9). Target sequence specificity is provided by two features: a gRNA molecule that contains a 20bp sequence that is complementary to the target site and forms a complex with Cas9, and an NGG PAM (proto-spacer adjacent motif) located immediately in 3' of the gRNA binding site that is essential for Cas9/DNA interaction. Cas9 cleavage generates a double-stranded break (DSB) which is typically repaired by the non-homologous end joining (NHEJ) pathway, ultimately resulting in insertions or deletions (indels) due to misrepair. Generation of indels in exonic sequences can be used to create frameshift mutations and is therefore a useful approach to generate loss of function mutations. Alternatively, the homology directed repair (HDR) pathway can be exploited to introduce a specific sequence (eg. point mutation) at the cleavage site. This approach requires a ssDNA or dsDNA template DNA containing the mutation of interest flanked by homology arms [5-7].

While CRIPSR/Cas9 technology has proven to be incredibly useful for producing targeted mutations, the availability of PAM sequences in the target region imposes a significant limitation on potential cleavage sites. This restriction is particularly relevant for HDR-mediated mutagenesis using ssDNA templates where the efficiency of HDR insertion is maximised when the DSB is generated as close as possible to the intended sequence modification[8]. Different versions of the CRISPR/Cas9 system have recently been characterised, providing alternatives to overcome this issue; for example, the Class 2 CRISPR endonuclease Cpf1, which creates a staggered cut distal to its TTTN PAM sequence[9-13]. Additionally, engineered variants of Cas9 from a number of species including SpCas9, *Francisella novicida* (FnCas9) and *Staphylococcus aureus* Cas9 (SaCas9) have been developed that use alternative PAM recognition sites [11; 13-15]. Higher fidelity Cas9 variants have also been developed to reduce potential off-target DNA mutations [16-18].

While Cpf1 and SaCas9 have been shown to produce heritable mutations in mice, several endonuclease variants remain untested [19-21]. We selected and tested a number of programmable endonuclease proteins with the potential to expand the range of targetable sequences through their use of alternative PAM sequences. Here we validate SpCas9 VQR and VRER for mouse genome editing and demonstrate efficient targeted mutagenesis using the SaCas9 KKH variant. We also show that this variant has a propensity to generate heterozygous mutations and provides a useful alternative to WT SpCas9 when targeting genes in which nullizygous mutations cause embryonic lethality.

Materials and methods

Generation of Cas9 variant/Cpf1 mRNAs

Plasmids for each variant: MSP469 (SpCas9 VQR), MSP1101 (SpCas9 VRER), MSP1830 (SaCas9 KKH), pY010 (pcDNA3.1- hAsCpf1) were obtained from Addgene (#65771,65773,70708,69982). pCMV/T7-hCas9 plasmid (Toolgen) was used for WT SpCas9. SaCas9 KKH was cloned into pBluescript to position a T7 promoter 5' of the gene. Plasmids were linearized with XhoI or SalI and mRNA generated using mMESSAGE mMACHINE® T7 ULTRA Transcription Kit (Ambion). mRNAs were purified using RNEasy Mini kit (QIAGEN).

sgRNA production

sgRNAs were designed using CCTop- CRISPR target online predictor[22]. Guide oligonucleotides (Geneworks, Australia, Supplementary Table 1) were denatured at 95 °C for 5 minutes and annealed at 25 °C before being ligated into BbsI linearised px459 vector (Addgene, #62988) or BsmBI linearised BPK2660 vector (SaCas9 sgRNA, Addgene, #70709). A template for *in vitro* transcription was generated by PCR amplification of the pX459 vector using a forward primer containing a 5' T7 promoter sequence and a reverse primer in the TracR sequence. *In vitro* transcription was performed using HiScribe T7 In Vitro Transcription Kit (NEB) and transcribed gRNA was purified using RNeasy Mini Kit (QIAGEN). For Cpf1 crRNA, an oligonucleotide (IDT) containing a T7 promoter sequence, gRNA and scaffold was used. *In vitro* transcription was performed using HiScribe T7 In Vitro Transcription Kit (NEB) and transcribed crRNA was purified using RNA Clean & Concentrator-100 (Zymo Research).

Microinjection

C57BL/6 females were superovulated by injecting 5IU PMSG (Folligon; Intervet India) followed by 5IU hCG (Chorulon; Intervet India) 47.5 hours later. Superovulated females were mated to male C57BL/6 mice and fertilised oocytes were collected from oviducts in FHM (Merck Millipore). Oocytes were maintained in $KSOM^{AA}$ (Merck Millipore) at 37 °C in 5% CO2 and were screened for the presence of 2 pronuclei to indicate fertilization. Fertilised oocytes were microinjected (cytoplasmic) with buffered solution containing Cas9 variant mRNA and sgRNA (concentrations for each variant in Supplementary Table 2) in FHM before being transferred to the oviduct of pseudopregnant Swiss female mice. ssODN with phosphorothioate modifications on four nucleotides at the 5' and 3' end was included in mRNA/gRNA mix for HDR injections (Integrated DNA Technologies)[23].

All animal work was conducted in accordance with Australian guidelines for the care and use of laboratory animals following approval by The University of Adelaide Animal Ethics committee (approval number S-2016-049).

Genomic DNA extraction

gDNA was extracted from embryo or tail tip samples using High Pure PCR Template Preparation Kit (Roche) or KAPA Mouse Genotyping Kit (Kapa Biosystems). PCR products were amplified using primers specific to each target region (sequences in Supplementary Table 1).

Heteroduplex Mobility Assay and RFLP analysis

PCR products were resolved by polyacrylamide gel electrophoresis in non-denaturing gels with 12% acrylamide-bisacrylamide solution (29:1, 30% solution, Sigma), 1X Tris-borate EDTA,

ammonium persulfate and TEMED with a low percentage stacker layer (4% acrylamide-bisacrylamide solution). Prior to loading, PCR products (unmixed and mixed with WT) were re-annealed to form heteroduplex products by heating to 95°C for 5 minutes and slowly ramped down to room temperature. RFLP analysis was performed by mixing 5uL of PCR products with restriction enzyme and appropriate buffer (NEB) in a total volume of 20uL and incubating at the appropriate temperature for 90 minutes.

Sanger sequencing and cloning

PCR products were purified using QIAquick PCR purification kit (Qiagen) before being used as a template in a BigDye Terminator v3.1 reaction (ThermoFisher) with the forward PCR primer. Sequencing was performed by the Australian Genome Research Facility, Adelaide. For cloning, PCR products were ligated into pGEM-T easy vector (Promega) and resulting clones screened for the presence of a BsiHKAI restriction site. For each sample, a plasmid with and without the BsiHKAI restriction site was sequenced using a T7 primer.

Results

CRISPR variants create targeted mutations *in vivo*

To assess emerging CRISPR and Cpf1 technologies for targeted mutagenesis in mice we microinjected mouse zygotes with mRNA encoding WT SpCas9, SpCas9 VQR, SpCas9 VRER, AsCpf1 or SaCas9 KKH and their cognate gRNAs (Table 1). To allow direct comparison of mutagenesis efficiencies, we targeted a region in the intron of *Ngn3* in which all gRNA target sequences overlapped (Figure 1). Each endonuclease cleavage site also contained a restriction site permitting RFLP analysis for mutation detection.

WT SpCas9 was initially selected for mutation analysis. Zygotes were microinjected with WT SpCas9 mRNA and *Ngn3* gRNA, transferred to pseudopregnant females and harvested at 10.5-12.5 days post transfer (dpt) for analysis (Table 2). PCR products spanning the target site were generated from 19 of the 20 embryos harvested for analysis. The failure of one sample (sample 12) to amplify across the *Ngn3* target site may reflect generation of large deletions on both alleles resulting in loss of PCR primer binding sites, a frequent occurrence of WT SpCas9 DNA DSB repair (our unpublished data). Consistent with this conclusion, we were able to successfully amplify DNA from the *Prrt2* locus (data not shown). We initially genotyped the samples using a polyacrylamide gel heteroduplex mobility assay (HMA). Analysis of PCR products revealed 12 of 19 samples harboured mutations (63.2%, Figure 2a, Supplementary Figure 1a). Interestingly, mixing samples in a 1:1 ratio with WT *Ngn3* PCR product before HMA analysis increased the mutation rate to 17 of 19 samples (89.5%, Figure 2b, Supplementary Figure 1b). These data indicate that a single mutant allele was amplified in 5 samples (presumably due to deletion of primer binding site(s) on the other mutant allele), resulting in

homoduplexes for the PCR product HMA. RFLP analysis with BsiHKAI confirmed these results, and revealed one additional mutant (sample 8) raising the total mutation rate to 18/19 (94.7%; Figure 2c, Supplementary Figure 1c). We sequenced several of these samples and observed indels ranging from 1bp insertions to 7bp deletions, confirming mutations resulting from SpCas9 DNA cleavage (data not shown). Together these data validate our screening assays and confirm that SpCas9 generates targeted mutations with high efficiency.

Next, we tested SpCas9 variants VQR and VRER, which recognise PAM sequences NGAN and NGCG, respectively. Mouse zygotes were microinjected with SpCas9 VQR or VRER mRNAs and corresponding gRNAs and transferred to pseudopregnant females. Embryos were harvested at 11.5-12.5 dpt and subjected to the same mutation screening protocol. For VRER, only 1 mutant embryo was observed from 16 samples and findings were consistent across unmixed and WT mixed HMA, as well as RFLP analysis and sequencing (Supplementary Figure 2). VQR embryo analysis showed that 2 of 18 samples were mutant, however heteroduplex bands were faint on the HMA both unmixed and in a 1:1 ratio with WT PCR product. Only 1 sample was visible by RFLP and both samples showed WT sequencing traces, suggesting mosaicism with only a small proportion of mutant cells (Supplementary Figure 3).

Mutation analysis of 36 AsCpf1/gRNA injected embryos revealed a mutation rate of 33%. Results were consistent across unmixed and WT mixed HMA assays and RFLP analysis indicating that large deletions affecting primer binding sites did not occur (Supplementary Figure 4). Selected mutations were confirmed by sequencing and mixed traces observed following the expected cleavage site in samples with strong undigested bands in RFLP analysis (data not shown).

High frequency of mono-allelic mutations using SaCas9 KKH

Zygote injection and transfer of SaCas KKH mRNA/gRNA yielded 25 embryos. Results from HMA (mixed and unmixed) were consistent with BsiHKAI RFLP analysis, revealing 14 mutants (56%, Figure 3, Supplementary Figure 5). Interestingly, almost all SaCas9 KKH mutant samples (13 of 14 mutant samples, 93%) appeared to have a remaining WT allele, which is rarely observed with WT SpCas9 injections (1 of 18 mutants- 6%, Figure 2b, Supplementary Figure 1b). To confirm the heterozygosity (single hit), we sequenced the cloned PCR products from the 13 samples. Consistent with BsiHKAI RFLP analysis, all samples contained both wild type and mutant alleles (Supplementary Figure 6).

While SpCas9 generates bi-allelic mutations with very high efficiency, this can be a disadvantage when targeting genes that cause nullizygous embryonic lethality as founder embryos die before birth. This problem is particularly acute when introducing a deleterious point mutation via HDR-mediated repair (which is less efficient than NHEJ) as the point mutation is almost always accompanied by a loss of function indel mutation on the other allele, resulting in embryonic lethality. Given the propensity of SaCas9 KKH to generate heterozygous indel mutations, we reasoned that we could exploit this property to generate mutants with an HDR-mediated mutant allele and a WT allele. To test this, we injected a ssDNA oligonucleotide repair template with 4 nucleotide differences encoding a BsmBI restriction site and “PAM killing” substitution with WT SpCas9 or SaCas9 KKH mRNAs and their cognate *Ngn3* gRNAs (Figure 4a). Embryos were collected mid-gestation for DNA extraction. BsmBI restriction digests of the *Ngn3* PCR product revealed 6 of 27 (22%) SaCas9 KKH embryos had oligonucleotide insertions alleles, whilst 4 of 9 (44%) WT SpCas9 embryos had the new

restriction site (Table 3). The presence of a WT allele in these samples was assessed by BsiHKAI digest, showing 6 of 6 (100%) of SaCas9 KKH samples retained a WT allele (Supplementary Figure 7a). Conversely, just 2 of the 4 (50%) WT SpCas9 samples with oligo insertions had a WT allele intact (Supplementary Figure 7b). Sanger sequencing confirmed the oligo insertion in all cases, showing mixed traces at the restriction sites (Figure 4b).

SaCas9 KKH cleaves effectively at different loci

We next wanted to confirm the capacity of SaCas9 KKH to create mutations at a different genomic location. To assess this, we selected the tyrosinase (*Tyr*) gene for targeted mutagenesis. Tyrosinase is essential for the production of melanin which is the pigment responsible for black coat colour. Cells lacking tyrosinase appear white while cells containing at least one functional copy are black. Thus, *Tyr* provides a useful marker to identify nullizygous and mosaic null founders. Mouse zygotes were injected with (300ng/uL) SaCas9 KKH mRNA and a gRNA targeting exon 1 and transferred to pseudopregnant females. Screening of founders via HMA revealed that 9 of 35 (25%) had mutations at the *Tyr* locus (Table 4). Of the 9 resulting founders only 1 had phenotypic evidence of nullizygous mosaicism (patchy black/white coat colour, Supplementary Figure 8d) whilst the remaining 8 mutants had fully black coats, providing further independent evidence that SaCas9 KKH generates mostly heterozygous founders. Sequencing the 9 mutants confirmed indels at the SaCas9 KKH cut site, as well as a remaining WT allele. Finally, to confirm that mutations created by SaCas9 KKH could be transmitted to the next generation, 5 founders (male and female) were bred and offspring from one female and one male (40%) had mutations based on HMA analysis and/or Sanger sequencing (Table 4, Supplementary Figure 8) confirming germline transmission.

Discussion

CRISPR/Cas9 technology has enabled rapid production of genetically modified mice for modelling human physiology and disease. However, the limitation imposed by the NGG PAM requirement means that some regions of the genome remain untargetable. This is particularly pertinent for HDR-mediated mutagenesis where efficiency is enhanced by cleaving as close as possible to the site of the intended mutation [8]. Whilst a number of CRISPR variants with different PAM recognition have become available, their ability to generate mutations in mice has not yet been reported. Validating these variants will extend the available toolkit for genome editing in mice and potentially other mammals, increasing flexibility and broadening the targetable regions of the genome.

In this study, we assessed the mutagenic activity of AsCpf1, SpCas9 VQR, SpCas9 VRER and SaCas9 KKH CRISPR variants via mouse zygote injection, using WT SpCas9 as a "positive control". All variants successfully generated mutations at the target site, although significant variations in efficiency were observed (Table 1).

Previous studies have shown that injection of zygotes with AsCpf1 generated mutations in 18-79% of offspring [21]. Consistent with these data, we observed an AsCpf1 mutation rate within this range (33%). Notably, a recent report showed higher editing efficiency with pronuclear injection and high AsCpf1 mRNA concentrations, which may be a useful strategy for future applications [24]. We have also shown for the first time that SpCas9 VQR and VRER variants are functional in mouse zygotes, though the mutation efficiency was lower than that of WT SpCas9 (11% and 6%, respectively). These variants will require further testing at additional genomic loci to fully determine their cutting efficiencies. It should also be noted that the SpCas9 VQR

variant was previously shown to have maximum efficiency when targeting a site with a NGAG PAM, with the NGAT PAM (as used in this study) being its second preferred PAM sequence [15].

The most striking finding of this study was the activity of the previously untested SaCas9 KKH variant. Mutations were detected in 25-70% of mouse embryos enabling steady generation of mutant founders from a single injection session. Importantly, unlike WT SpCas9, the SaCas9 KKH variant left an intact WT allele in 93% of mutant embryos. Furthermore, when introducing a point mutation via HDR, a WT allele remained in 100% of HDR edited embryos while WT SpCas9-injected samples had a WT allele in only 50%. We propose that SaCas9 KKH may prove a useful alternative to WT SpCas9 when introducing deleterious point mutations into genes essential for embryonic survival.

We showed that this feature of SaCas9 KKH was applicable to other genomic locations and confirmed that mutations in the tyrosinase gene induced by SaCas9 KKH were transmitted through the germline, making it suitable for colony generation. The lack of white coat colour in mutant pups provides further evidence that SaCas9 KKH does not generate mutations in both alleles. This result also argues against the unlikely possibility that SaCas9 KKH generates mosaic mice composed of WT and homozygous mutant cells (which would be indistinguishable from heterozygous mice in molecular genotyping assays). The increase of SaCas9 KKH mRNA concentration when injecting the tyrosinase gRNA did not increase efficiency of mutation, nor seem to affect the monoallelic mutagenic activity of the variant. At the higher and lower mRNA concentration, SaCas9 KKH showed higher transfer to offspring rates than WT SpCas9. The same was true for the other variants tested (SpCas9 VQR and VRER and AsCpf1), suggesting

that they do not have any greater toxic effect on the zygote than the commonly used WT SpCas9.

The mechanism by which the SaCas9 KKH variant fails to modify all available alleles remains unclear, although we hypothesize that it is simply a consequence of lower nuclease efficiency. While the mutagenic activity is robust, it appears that the variant is not as active as WT SpCas9, reducing the probability of multiple mutations. The kinetics of mRNA translation and assembly of the SaCas9 KKH/gRNA RNP complex may be responsible, resulting in later induction of the DSB. *In vivo* degradation of the protein, mRNA or RNP may also be more rapid, preventing further cleavage of the target DNA. Indeed, many of our SaCas9 KKH mutant sample sequencing traces showed mutant allele dosage of around 1 in 4, suggesting that mutations occurred after the single cell stage, resulting in mosaicism (eg. Figure 4). Mosaicism in founders is well established when using SpCas9 and has also been observed with mouse zygote injection of WT SaCas9 [4; 19; 25].

Conclusion

Our results confirm the utility of a number of new variants for targeting the mouse genome by zygote microinjection. The novel PAM sequences of the variants tested provide more flexibility when designing gRNAs to target specific genetic loci; of particular importance when targeting for HDR insertion. We have also characterised SaCas9 KKH as a useful alternative to WT SpCas9 when targeting point mutations into genes that cause nullizygous embryonic lethality. Further investigation of SaCas9 KKH mRNA and gRNA optimal concentrations is required and could increase its targeting efficiency or affect its tendency to edit just a single allele. Similarly, microinjection of ribonuclear protein (RNP) complex may result in higher efficiency of editing both or a single allele. A number of CRISPR variants that have been characterised in human cells still remain to be tested in the context of mouse genome editing, potentially enabling further expansion of the mouse genome editing 'toolkit'.

Acknowledgements

The authors acknowledge the facilities, and the scientific and technical assistance of the South Australian Genome Editing (SAGE) Facility, University of Adelaide/South Australian Health and Medical Research Institute. SAGE is supported by the Australian Phenomics Network (APN). The APN is supported by the Australian Government through the National Collaborative Research Infrastructure Strategy (NCRIS) program.

Author confirmation statement

F.A., P.T. and L.R. conceived and designed the study. Mouse zygote injections were performed by S.P. and M.W; D.P. performed and analysed cloning and also conceived HDR experiment. All other experiments performed by L.R. and M.A. and analysed by L.R. and P.T. Manuscript was prepared by L.R. and P.T with critical advice from F. A.; all authors revised and approved the manuscript prior to submission. This manuscript has been submitted solely to this journal and is not published, in press, or submitted elsewhere.

Author Disclosure Statement

No competing financial interests exist.

Table 1: Naturally occurring and engineered CRISPR variants assessed in this study.

Species	Mutations (from wildtype)	Abbreviation	PAM sequence	Optimal guide length (nucleotides)
Streptococcus pyogenes	--	SpCas9	NGG	17-20
	D1335V, R1335Q and T1337R	SpCas9 VQR	NGAN	
	D1135V, G1218R, R1335E and T1337R	SpCas9 VRER	NGCG	
Staphylococcus aureus	E782K, N968K and R1015H	SaCas9 KKH	NNNRRT	21-23
Acidaminococcus sp. BV3L6	--	AsCpf1	TTTN	18-22

Table 2: Generation of mutant embryos using CRISPR endonuclease variants

Variant	Zygotes injected	Zygotes transferred	Age collected (dpt)	Number collected (% of transferred)	Mutant
WT SpCas9	121	95	11.5-12.5	20 (21%)	18 of 19 (95%)
AsCpf1	71	61	10.5	36 (59%)	12 (33%)
SpCas9 VQR	39	35	12.5	18 (51%)	2 (11%)
SpCas9 VRER	58	55	11.5-12.5	16 (29%)	1 (6%)
SaCas9 KKH	65	58	10.5-11.5	25 (43%)	14 (56%)

Table 3: Oligonucleotide insertion with WT SpCas9 and SaCas9 KKH

Variant	Zygotes injected	Zygotes transferred	Age Collected (dpt)	Number collected (% of transferred)	Number mutant	Oligo Insertion	Oligo ins + WT allele detected
WT SpCas9	121	98	12.5	9 (9.2%)	9 (100%)	4 (44%)	2 (50%)
SaCas9 KKH	140	103	12.5	27 (26.2%)	19 (70.4%)	6 (22%)	6 (100%)

Table 4: Generation of tyrosinase mutant founders and breeding for germline transmission

Variant	Zygotes injected	Zygotes transferred	Number of founders (% of transferred)	Mutant	% Mutant	White coat phenotype	Mutant founders bred	Germline transmission in offspring
SaCas9 KKH	108	89	35 (39.3%)	9	25.7%	0 (1 mosaic)	5	2 (40%)

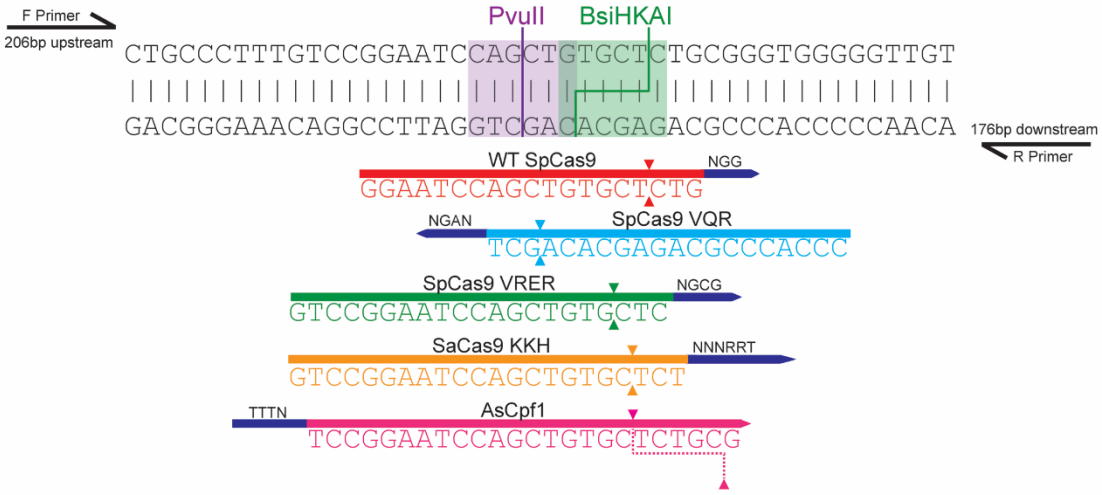


Figure 1: Schematic of the *Ngn3* intronic target site with gRNA targets and PAM sites for each variant beneath. Cut site for each endonuclease indicated by arrowheads corresponding to their colour. PvuII (purple) and BsiHKAI (green) restriction endonuclease recognition sites are highlighted, and their respective cut sites marked within. PCR primer distance from cut sites indicated above and below.

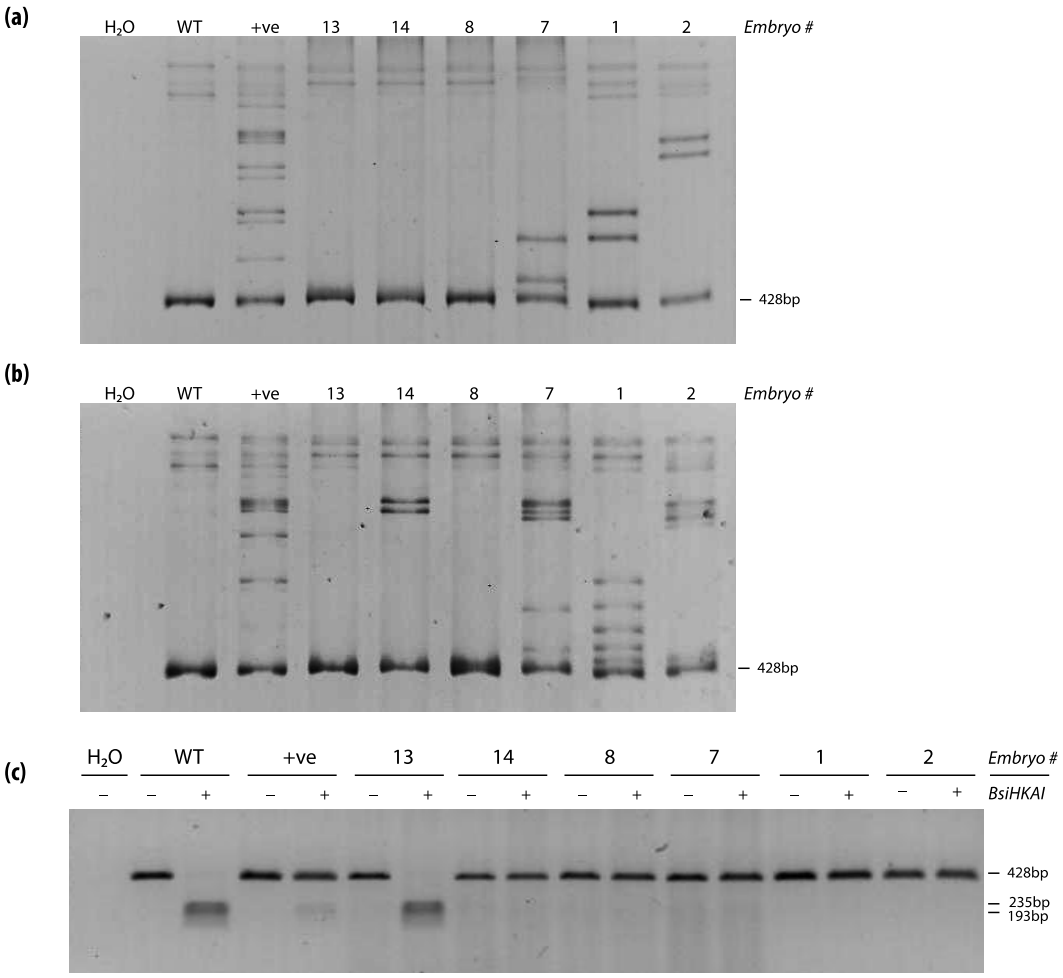


Figure 2: Representative results from 11.5-12.5dpt embryos following microinjection of WT SpCas9 and *Ngn3* intron gRNA. **(a)** A region spanning the *Ngn3* target site was PCR amplified and reannealed. Heteroduplex formation was assessed by polyacrylamide gel electrophoresis. **(b)** PCR products mixed with WT PCR product before annealing, then run on polyacrylamide gel. **(c)** Restriction digest with BsiHKAI, with no enzyme (-) and enzyme (+) conditions run side by side. Wild type (WT), positive control (+ve) and negative controls (H_2O) included.

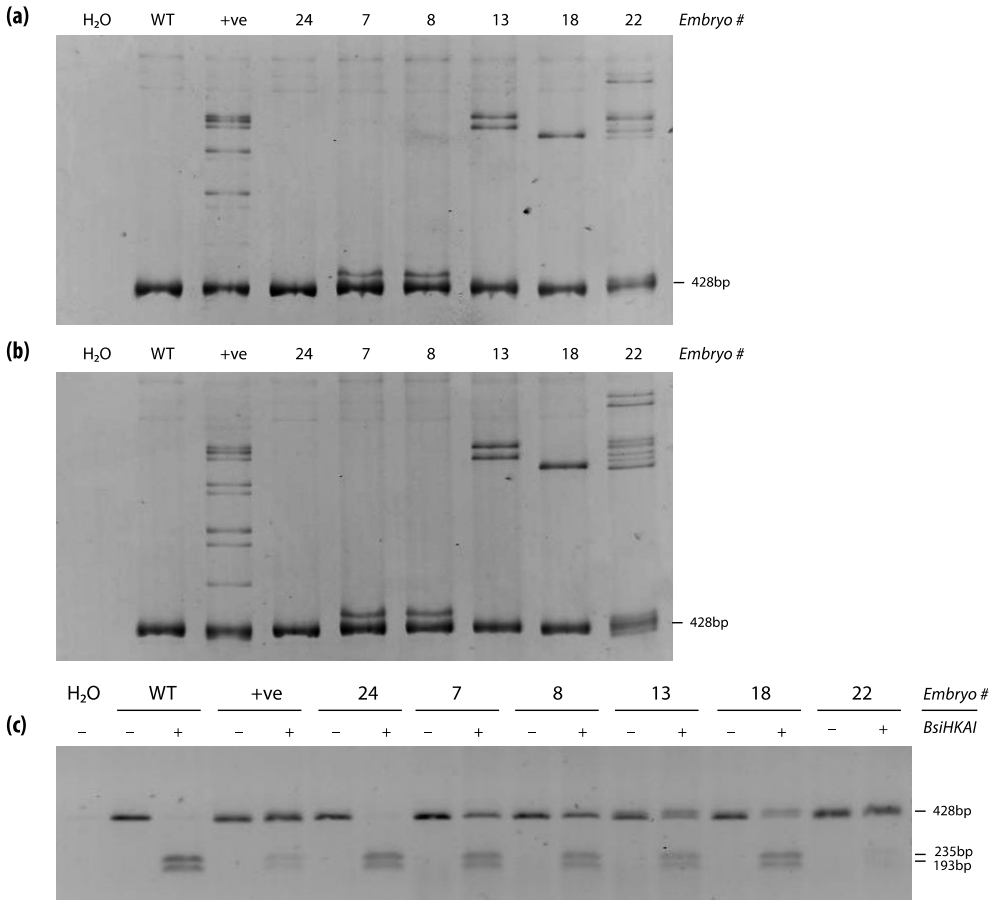


Figure 3: Representative results from 10.5-11.5dpt embryos following microinjection of WT SaCas9 KKH and *Ngn3* intron gRNA. **(a)** A region spanning the *Ngn3* target site was PCR amplified and reannealed. Heteroduplex formation was assessed by polyacrylamide gel electrophoresis. **(b)** PCR products mixed with WT PCR product before annealing, then run on a polyacrylamide gel. **(c)** Restriction digest with BsiHKAI, with no enzyme (-) and enzyme (+) conditions run side by side. Wild type (WT), positive control (+ve) and negative controls (H_2O) included.

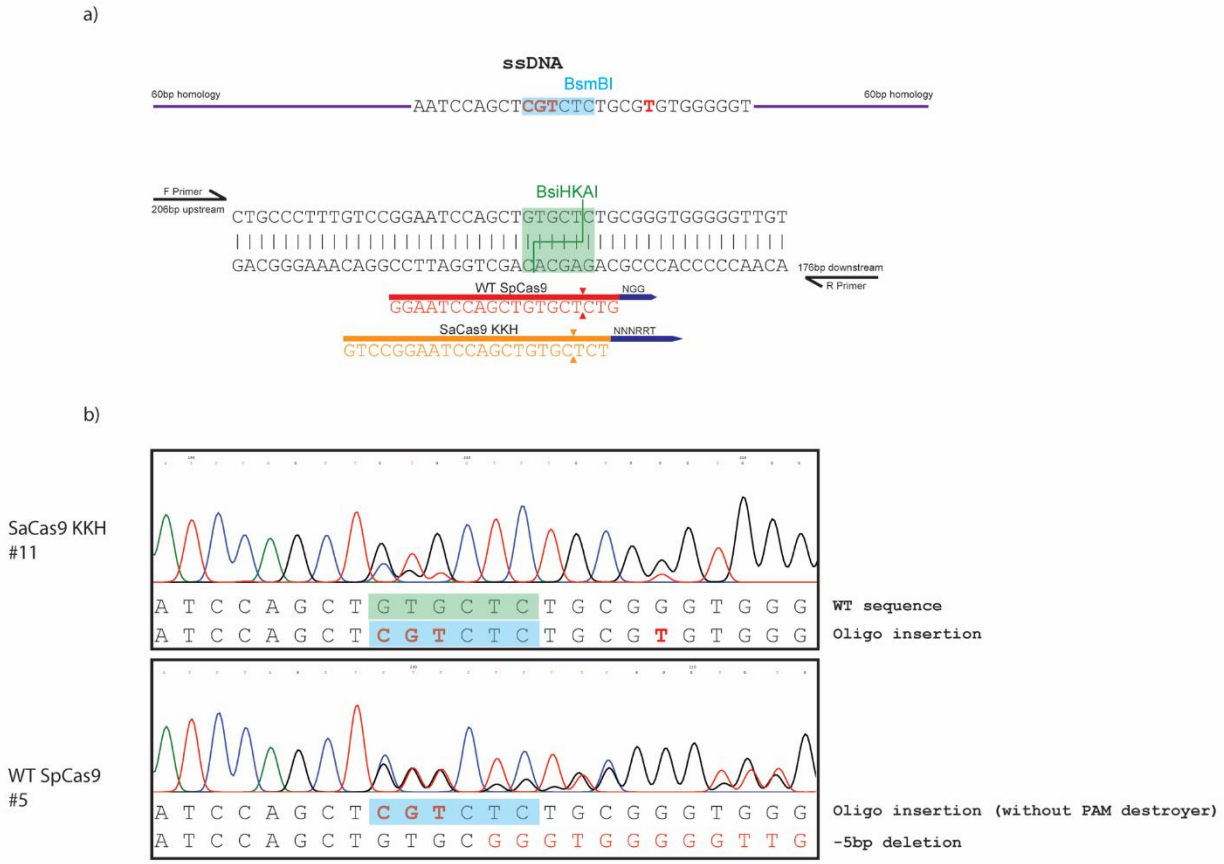


Figure 4: Generating a point mutation at the *Ngn3* intronic site using WtSpCas9 and SaCas9 KKH. **(a)** A schematic of the *Ngn3* target site, including gRNA sequences and endonuclease cut sites. The ssDNA oligonucleotide donor is shown above with base changes marked in red. The newly created BsmBI restriction site is highlighted in blue and existing BsiHKAI site highlighted in green. **(b)** Example of Sanger sequencing traces across the *Ngn3* intron region from SaCas9 KKH (sample 11) and WTSpCas9 embryos with oligonucleotide insertions. Alleles separated below to show restriction site insertion/preservation/loss.

References

1. Griep AE, John MC, Ikeda S, et al. Gene targeting in the mouse. *Methods Mol Biol.* 2011;770;293-312. DOI: 10.1007/978-1-61779-210-6_11

2. Wang H, Yang H, Shivalila CS, et al. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell.* 2013;153;910-8. DOI: 10.1016/j.cell.2013.04.025

3. Yang H, Wang H, Shivalila CS, et al. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell.* 2013;154;1370-9. DOI: 10.1016/j.cell.2013.08.022

4. Singh P, Schimenti JC, Bolcun-Filas E. A mouse geneticist's practical guide to CRISPR applications. *Genetics.* 2015;199;1-15. DOI: 10.1534/genetics.114.169771

5. Jinek M, Chylinski K, Fonfara I, et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science.* 2012;337;816-21. DOI: 10.1126/science.1225829

6. Cong L, Ran FA, Cox D, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science.* 2013;339;819-23. DOI: 10.1126/science.1231143

7. Mali P, Yang L, Esvelt KM, et al. RNA-guided human genome engineering via Cas9. *Science.* 2013;339;823-6. DOI: 10.1126/science.1232033

8. Paquet D, Kwart D, Chen A, et al. Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature.* 2016;533;125-9. DOI: 10.1038/nature17664

9. Zetsche B, Gootenberg JS, Abudayyeh OO, et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell.* 2015;163;759-71. DOI: 10.1016/j.cell.2015.09.038

10. Ran FA, Cong L, Yan WX, et al. In vivo genome editing using Staphylococcus aureus Cas9. *Nature.* 2015;520;186-91. DOI: 10.1038/nature14299

11. Esvelt KM, Mali P, Braff JL, et al. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat Methods.* 2013;10;1116-21. DOI: 10.1038/nmeth.2681

12. Komor AC, Badran AH, Liu DR. CRISPR-Based Technologies for the Manipulation of Eukaryotic Genomes. *Cell.* 2017;168;20-36. DOI: 10.1016/j.cell.2016.10.044

13. Hirano H, Gootenberg JS, Horii T, et al. Structure and Engineering of Francisella novicida Cas9. *Cell.* 2016;164;950-61. DOI: 10.1016/j.cell.2016.01.039

14. Kleinstiver BP, Prew MS, Tsai SQ, et al. Broadening the targeting range of Staphylococcus aureus CRISPR-Cas9 by modifying PAM recognition. *Nat Biotechnol.* 2015;33;1293-1298. DOI: 10.1038/nbt.3404

15. Kleinstiver BP, Prew MS, Tsai SQ, et al. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature.* 2015;523;481-5. DOI: 10.1038/nature14592

16. Kleinstiver BP, Pattanayak V, Prew MS, et al. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature.* 2016;529;490-5. DOI: 10.1038/nature16526

17. Slaymaker IM, Gao L, Zetsche B, et al. Rationally engineered Cas9 nucleases with improved specificity. *Science.* 2016;351;84-8. DOI: 10.1126/science.aad5227

18. Chen JS, Dagdas YS, Kleinstiver BP, et al. Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature.* 2017; DOI: 10.1038/nature24268

19. Zhang X, Liang P, Ding C, et al. Efficient Production of Gene-Modified Mice using Staphylococcus aureus Cas9. *Sci Rep.* 2016;6;32565. DOI: 10.1038/srep32565

20. Hur JK, Kim K, Been KW, et al. Targeted mutagenesis in mice by electroporation of Cpf1 ribonucleoproteins. *Nat Biotechnol.* 2016;34;807-8. DOI: 10.1038/nbt.3596

21. Kim Y, Cheong SA, Lee JG, et al. Generation of knockout mice by Cpf1-mediated gene targeting. *Nat Biotechnol.* 2016;34;808-10. DOI: 10.1038/nbt.3614

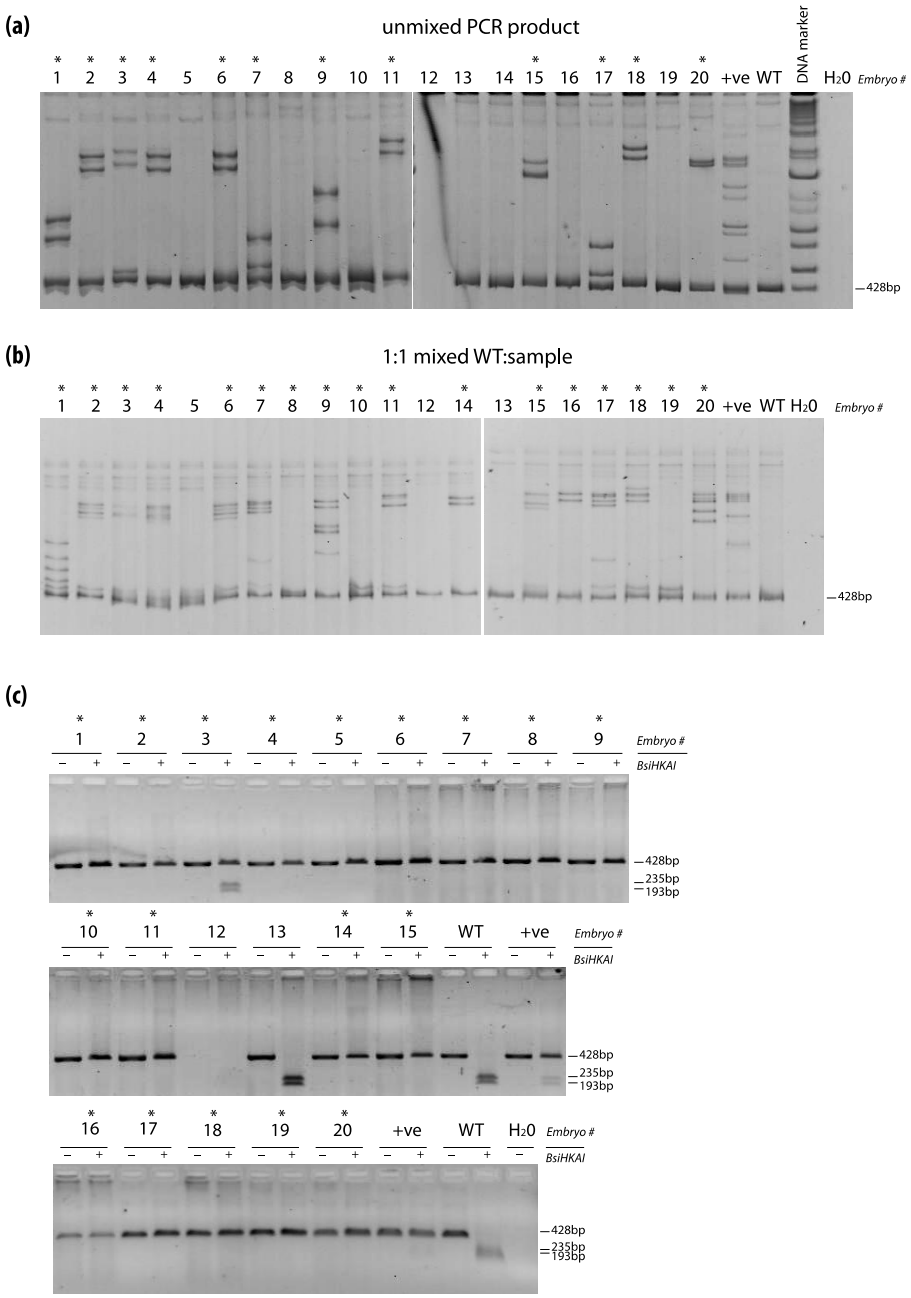
22. Stemmer M, Thumberger T, Del Sol Keyer M, et al. CCTop: An Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool. *PLoS One.* 2015;10;e0124633. DOI: 10.1371/journal.pone.0124633

23. Renaud JB, Boix C, Charpentier M, et al. Improved Genome Editing Efficiency and Flexibility Using Modified Oligonucleotides with TALEN and CRISPR-Cas9 Nucleases. *Cell Rep.* 2016;14;2263-2272. DOI: 10.1016/j.celrep.2016.02.018

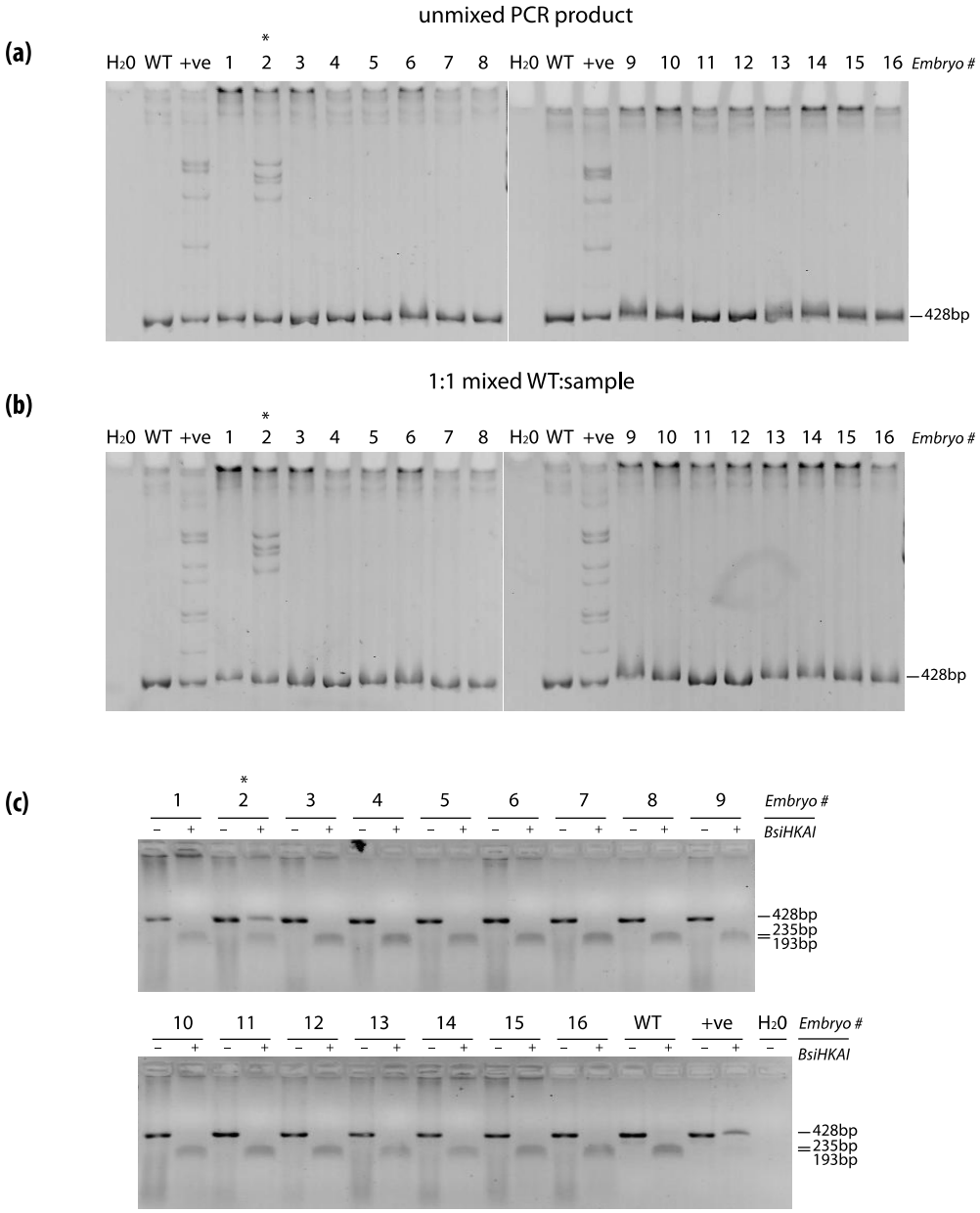
24. Watkins-Chow DE, Varshney GK, Garrett LJ, et al. Highly Efficient Cpf1-Mediated Gene Targeting in Mice Following High Concentration Pronuclear Injection. *G3 (Bethesda).* 2017;7;719-722. DOI: 10.1534/g3.116.038091

25. Yen ST, Zhang M, Deng JM, et al. Somatic mosaicism and allele complexity induced by CRISPR/Cas9 RNA injections in mouse zygotes. *Dev Biol.* 2014;393;3-9. DOI: 10.1016/j.ydbio.2014.06.017

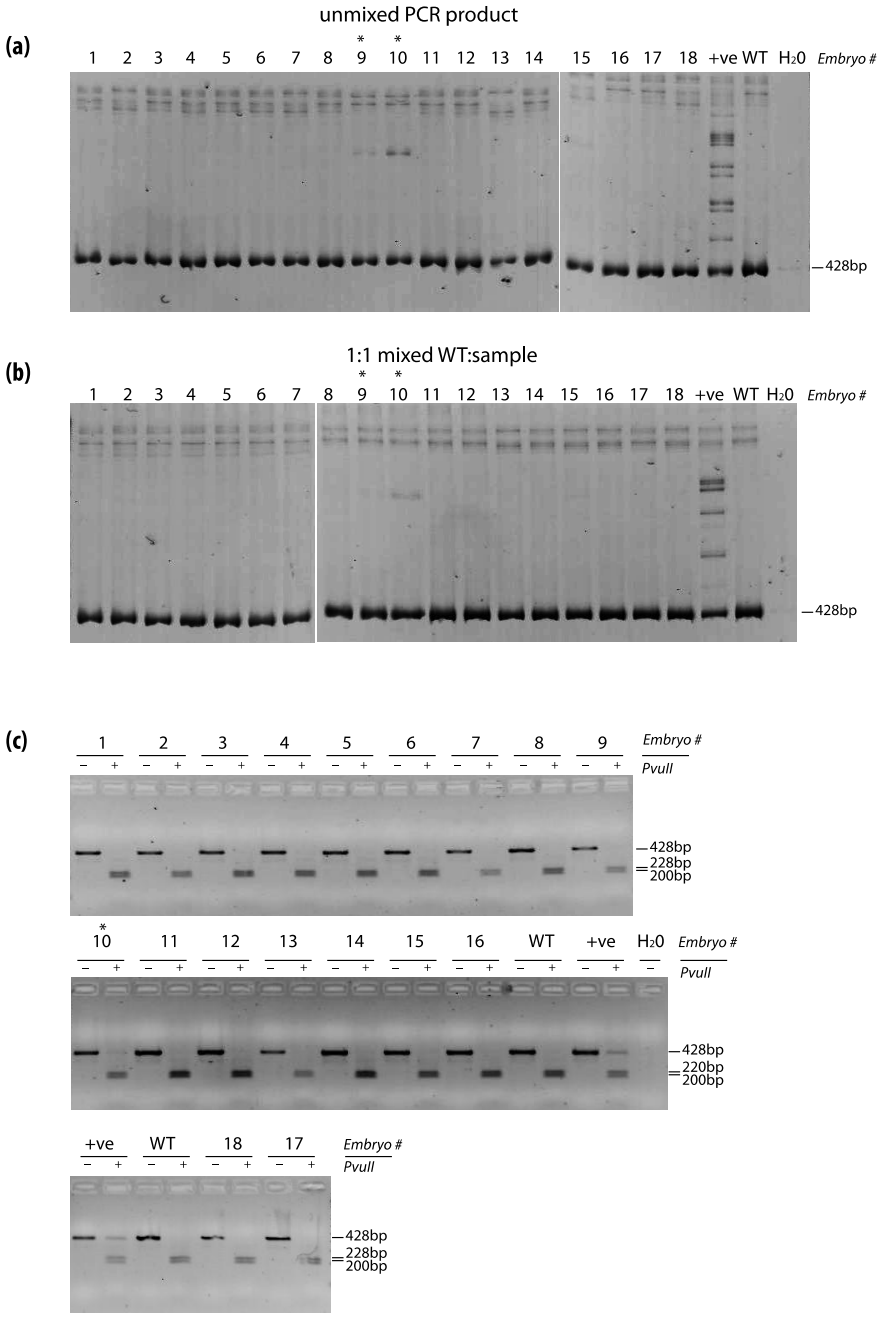
Supplementary Material



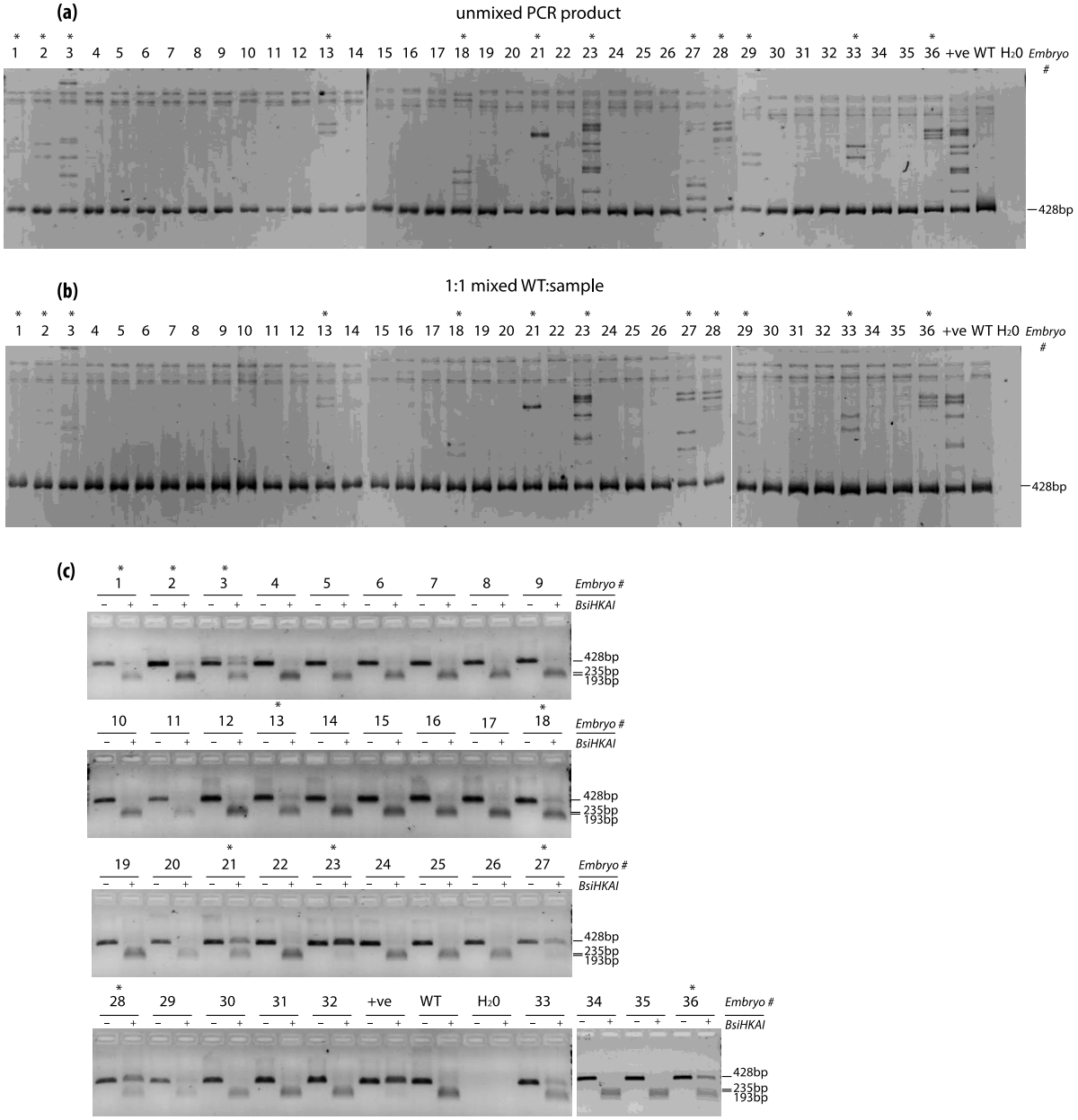
Supplementary Figure 1: Embryo screening results for 20 samples from 11.5-12.5dpt embryos following microinjection of WT SpCas9 and Ngn3 intron gRNA. (a) A region spanning the Ngn3 target site was PCR amplified and reannealed. Heteroduplex formation was assessed by polyacrylamide gel electrophoresis. (b) PCR products mixed with WT PCR product before annealing, then run on polyacrylamide gel. (c) Restriction digest with BsiHKAI, with no enzyme (-) and enzyme (+) conditions run side by side. Mutant samples marked with *, wild type (WT), positive control (+ve) and negative controls (H_2O) included.



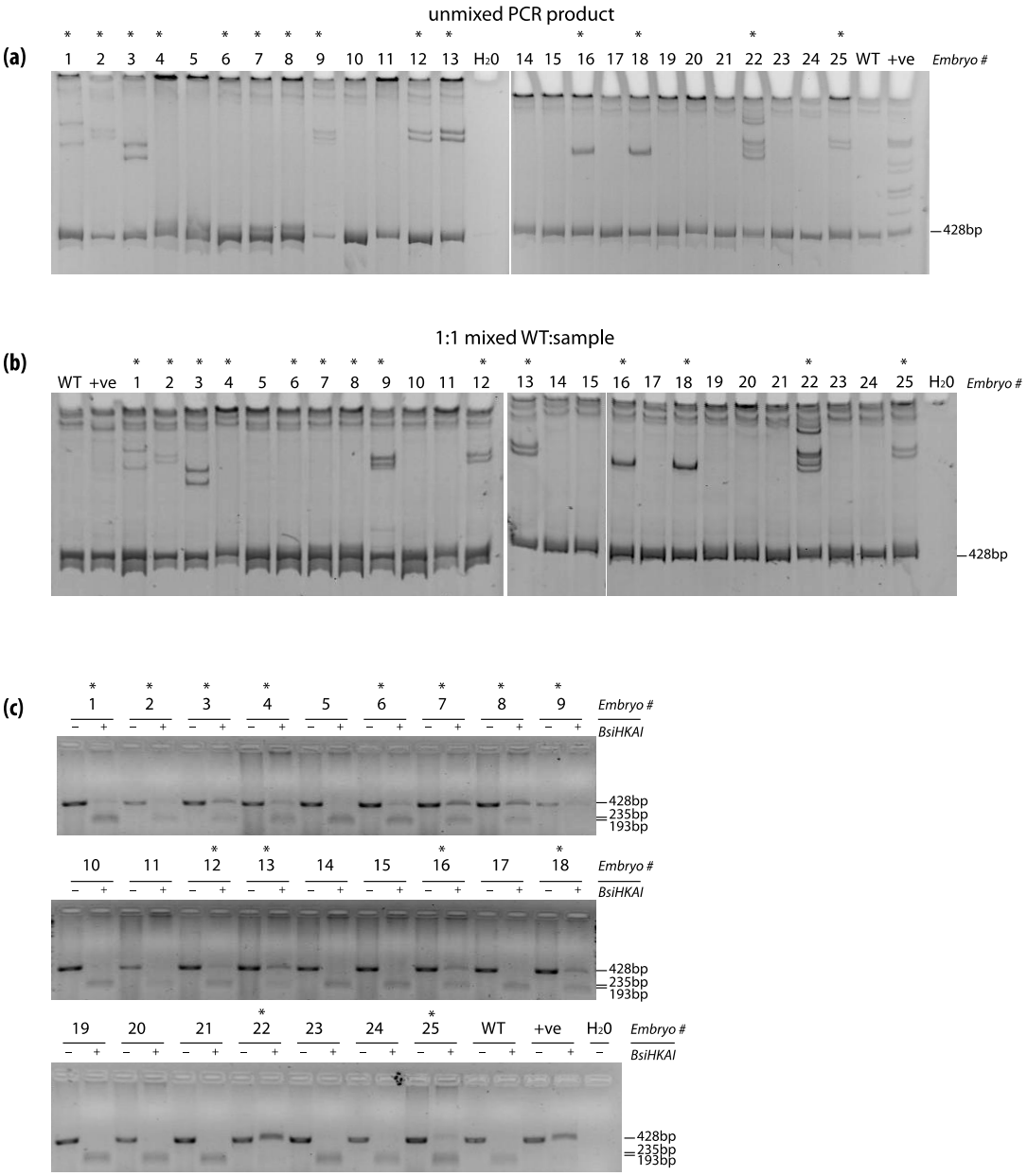
Supplementary Figure 2: Embryo screening results for 16 samples from 11.5-12.5dpt embryos following microinjection of SpCas9 VRER and Ngn3 intron gRNA. (a) A region spanning the Ngn3 target site was PCR amplified and reannealed. Heteroduplex formation was assessed by polyacrylamide gel electrophoresis. (b) PCR products mixed with WT PCR product before annealing, then run on polyacrylamide gel. (c) Restriction digest with BsiHKAI, with no enzyme (-) and enzyme (+) conditions run side by side. Mutant samples marked with *, wild type (WT), positive control (+ve) and negative controls (H_2O) included.



Supplementary Figure 3: Embryo screening results for 18 samples from 12.5dpt embryos following microinjection of SpCas9 VQR and Ngn3 intron gRNA. (a) A region spanning the Ngn3 target site was PCR amplified and reannealed. Heteroduplex formation was assessed by polyacrylamide gel electrophoresis. (b) PCR products mixed with WT PCR product before annealing, then run on polyacrylamide gel. (c) Restriction digest with PvuII, with no enzyme (-) and enzyme (+) conditions run side by side. Mutant samples marked with *, wild type (WT), positive control (+ve) and negative controls (H_2O) included.

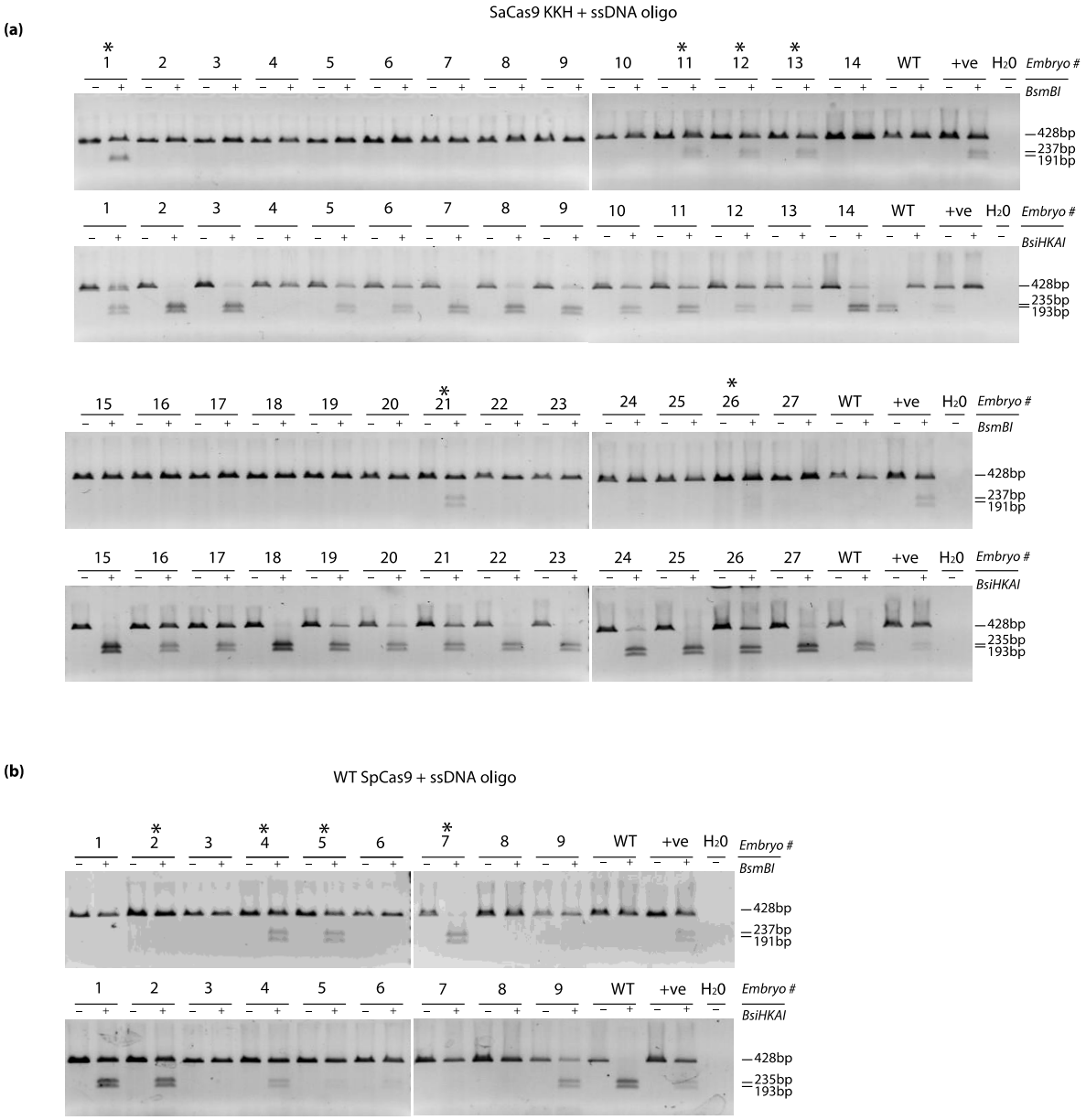


Supplementary Figure 4: Embryo screening results for 36 samples from 10.5dpt embryos following microinjection of AsCpf1 and Ngn3 intron crRNA. (a) A region spanning the Ngn3 target site was PCR amplified and reannealed. Heteroduplex formation was assessed by polyacrylamide gel electrophoresis. (b) PCR products mixed with WT PCR product before annealing, then run on polyacrylamide gel. (c) Restriction digest with BsiHKAI, with no enzyme (-) and enzyme (+) conditions run side by side. Mutant samples marked with *, wild type (WT), positive control (+ve) and negative controls (H_2O) included.

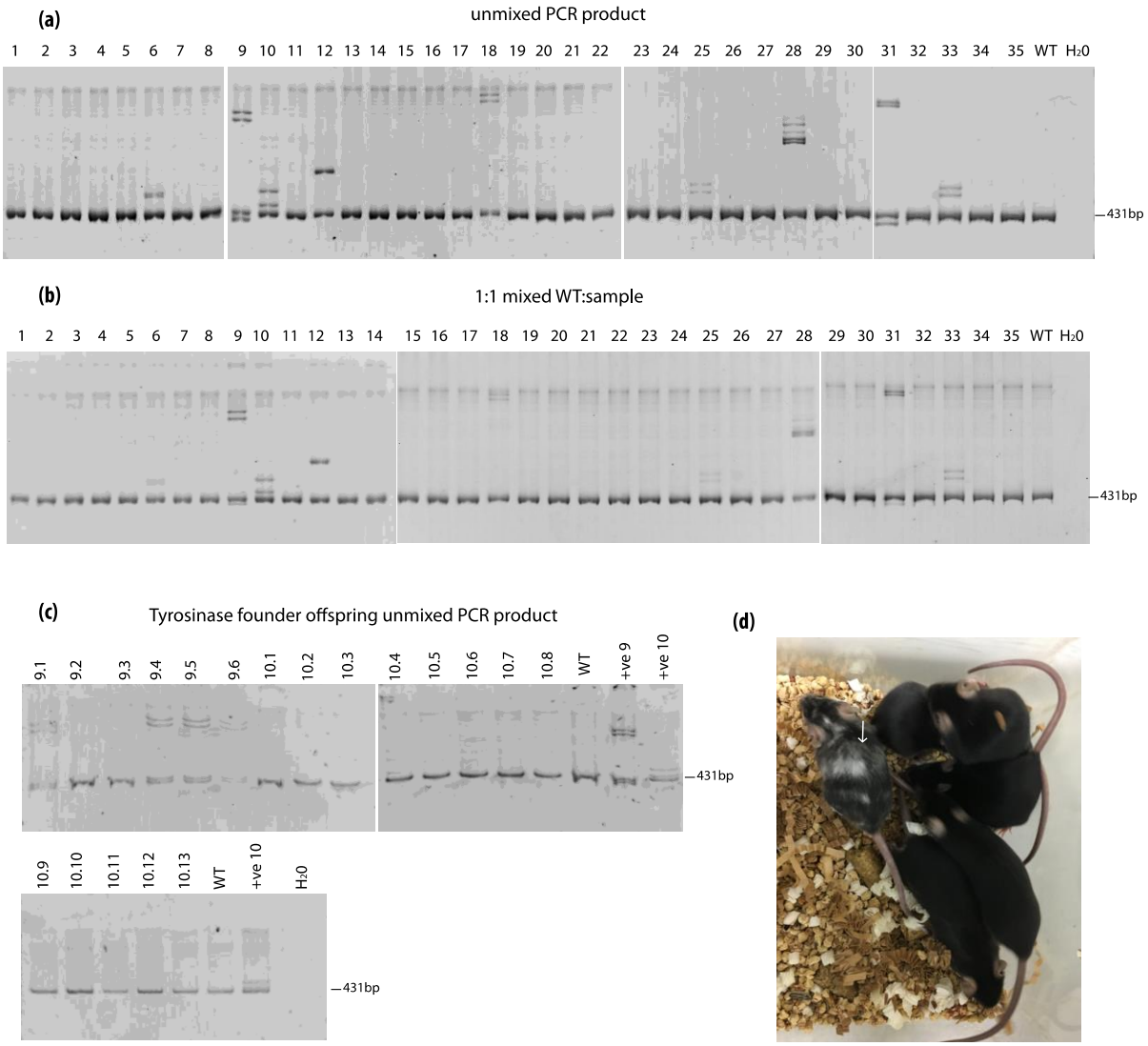


Supplementary Figure 5: Embryo screening results for 25 samples from 10.5-11.5dpt embryos following microinjection of SaCas9 KKH and Ngn3 intron gRNA. (a) A region spanning the Ngn3 target site was PCR amplified and reannealed. Heteroduplex formation was assessed by polyacrylamide gel electrophoresis. (b) PCR products mixed with WT PCR product before annealing, then run on polyacrylamide gel. (c) Restriction digest with BsiHKAI, with no enzyme (-) and enzyme (+) conditions run side by side. Mutant samples marked with *, wild type (WT), positive control (+ve) and negative controls (H_2O) included.

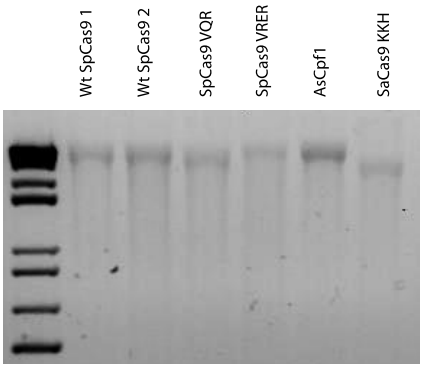
Supplementary Figure 6: Sanger sequencing results from clones of all SaCas9 KKH (a) and WT SpCas9 (b) Ngn3 PCR products positive for a BsiHKAI restriction site in the initial screen. For each sample, a clone that was positive and negative for BsiHKAI was selected and sequenced. BsiHKAI restriction site is highlighted in green and nature of allele described to the right of sequence.



Supplementary Figure 7: BsmBI and BsiHKAI restriction digest screens for oligo insertion at the Ngn3 intronic site following microinjection of endonuclease, gRNA and donor template. 27 SaCas9 KKH (a) and WT SpCas9 (b) samples were collected at 12.5 dpt and PCR amplified across the Ngn3 region. Restriction digests with BsiHKAI (upper panel) and BsmBI (lower panel) shown, with no enzyme (-) and enzyme (+) conditions run side by side. Wildtype (WT), positive control (+ve) and negative controls (H_2O) included. Samples positive for a BsmBI restriction site indicated by asterisk.



Supplementary Figure 8: Heteroduplex screening for 35 founder samples following cytoplasmic microinjection of SaCas9 KKH and Tyrosinase Exon1 gRNA. (a) A region spanning the Tyrosinase target site was PCR amplified and reannealed. Heteroduplex formation was assessed by polyacrylamide gel electrophoresis. (b) PCR products mixed with WT PCR product before annealing, then run on polyacrylamide gel. (c) Unmixed PCR samples from the offspring of two founders (9 and 10) after breeding with WT C57BL/6 females confirming germline transmission from founder #9. Wildtype (WT), positive control (+ve) and negative controls (H_2O) included. (d) Mosaic founder #18 (indicated by white arrow) with littermates (five pictured, with black coat phenotype).



Supplementary Figure 9: mRNA for all endonucleases used in microinjection confirming RNA integrity. Approximately 780ng of each mRNA sample loaded.

This page is left intentionally blank.

Chapter 4:

Evaluation of antibodies for immunofluorescent detection of endogenous HA-FLAG tagged protein in the mouse brain

4.1 Summary

CRISPR/Cas9 technology has opened up countless possibilities for genome editing and dramatically simplified the production of genetically modified mice. Previously, creating subtle mutations or small sequence insertions in mice was impractical due to the time required and low efficiency of the gene targeting method. Now it is simple, fast and efficient. Our laboratory has recently used genome editing to produce the PCDH19$^{HA\text{-}FLAG}$ mouse, containing HA and FLAG small epitope tags at the C terminal end of the protein[92]. This mouse was produced to overcome our lack of specific antibody for endogenous WT PCDH19 protein.

Protocadherin 19 (*PCDH19*) is an X-linked gene, meaning that one copy of the gene randomly undergoes X-inactivation early in development, resulting in expression of a single allele. *PCDH19*-related disease has unusual inheritance; patients with homozygous mutations in *PCDH19* are unaffected, whilst heterozygous mutation carriers suffer from protocadherin 19 girls clustering epilepsy (*PCDH19*-GCE). Through crossing the PCDH19$^{HA\text{-}FLAG}$ mouse with a transgenic PCDH19 knockout (PCDH19$^{\beta\text{-}Geo}$), we were able to stain for wild type PCDH19 cells with a commercial HA antibody and for knockout cells with an antibody against β-galactosidase. We described striking cell sorting defects in mice heterozygous for the PCDH19$^{\beta\text{-}Geo}$/PCDH19$^{HA\text{-}FLAG}$ and not in homozygous mutants. The PCDH19$^{HA\text{-}FLAG/+}$ mouse also served as a useful control to demonstrate normal X-inactivation for comparison to the disease-related abnormal cell sorting.

Whilst localising endogenous PCDH19$^{HA\text{-}FLAG}$, we noticed that many of the commercial antibodies trialled were not suitable for staining specifically on mouse brain. This manuscript tests a series of commercially available antibodies against either the HA or FLAG epitope for

their specificity. We began by overexpressing a PCDH19$^{HA\text{-}FLAG}$ construct in Cos cells and compared each antibody to an empty vector control, showing that all antibodies work on overexpression protein levels. We next tested the antibodies on PCDH19$^{HA\text{-}FLAG}$ embryonic brain tissue and compared to PCDH19 WT tissue. Only two antibodies (one HA and one FLAG) were suitable for specific staining of endogenous PCDH19$^{HA\text{-}FLAG}$ *in vivo*.

Given that CRISPR/Cas9 technology has made the tagging of proteins fast and simple, more researchers are likely to use this technique to overcome a lack of specific antibody. This manuscript will serve as a useful guide to help researchers in their choice of epitope tag and commercial antibody combination to localise their endogenous protein of interest.

Statement of Authorship

Title of Paper	Evaluation of antibodies for immunofluorescent detection of endogenous HA-FLAG tagged protein mouse brain
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☑ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	

Principal Author

Name of Principal Author (Candidate)	Louise Robertson
Contribution to the Paper	Conceived, designed and performed experiments, generated reagents, analysed data and wrote manuscript.
Overall percentage (%)	85%
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date 30/4/18

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Daniel Pederick
Contribution to the Paper	Conceived and designed study, performed cell transfections and sample preparation, evaluated manuscript.
Signature	Date 4/30/18

Name of Co-Author	Paul Thomas
Contribution to the Paper	Conceived, designed and supervised study, evaluated and edited manuscript.
Signature	Date 30/4/18

Evaluation of antibodies for immunofluorescent detection of endogenous HA-FLAG tagged protein in the mouse brain

Louise Robertson[1], Daniel Pederick[1], Paul Thomas[1,2*]

[1]The University of Adelaide and Robinson Research Institute, Adelaide, Australia

[2]South Australian Health and Medical Research Institute, Adelaide, Australia

***Corresponding author**

Louise Robertson louise.robertson@adelaide.edu.au

Daniel Pederick pederick@stanford.edu

Paul Thomas paul.thomas@adelaide.edu.au

Abstract:

Background: Localisation of proteins *in vivo* is integral to the understanding of gene function and human disease. However, specific antibodies are often not available for the protein of interest, particularly when immunostaining complex tissue such as the developing mouse brain. The advent of CRISPR/Cas9 genome editing technology has enabled efficient modification of endogenous proteins with small epitope tags such as HA or FLAG facilitating specific detection using validated commercially-available antibodies.

Results: In this study we compared eight commercially-available antibodies against either the HA or FLAG epitope for their ability to detect C-terminal endogenously-tagged PCDH19 protein ($PCDH19^{HA\text{-}FLAG}$) in the developing brain. We first confirmed the utility of each antibody on overexpressed $PCDH19^{HA\text{-}FLAG}$ in COS cells, showing that all antibodies stained specifically under these conditions. We next compared them by immunostaining developing brain tissue. While most antibodies showed immunofluorescence signal above background, only two (one HA and one FLAG) could detect the epitope tags specifically and exhibited negligible background on wild type tissue.

Conclusions: We identified two commercially-available antibodies that are effective tools for immunofluorescence detection of endogenous tagged protein complex mouse brain tissue. We also identified several antibodies that are not specific under these staining conditions. These data will aid researchers in their design of endogenous tagged protein alleles and the selection of antibodies for immunofluorescence detection.

Background

Mice are widely used to investigate gene and protein function and to model human genetic disorders. The ability to specifically detect proteins *in vivo* under normal and pathological conditions is critical for mouse model characterisation. One method for detecting a protein of interest is through immunostaining using an antibody against endogenous protein. However, antibodies are not always available to detect a protein of interest. Additionally, those that are available often generate background signal through non-specific binding to other proteins, particularly in a complex tissue such as the mouse brain. Alternative approaches include overexpression of fusion proteins with a fluorescent reporter or small epitope tag. A disadvantage of this approach is that both overexpression and large protein fusions (eg. GFP) can alter the function and/or subcellular localisation of the protein [1].

To overcome these limitations, mouse models expressing an endogenous fusion protein with a small epitope tag have been developed [2-7]. Epitope tags such as haemagglutinin (HA), FLAG, myc, and V5, which are commonly used for protein detection *in vitro*, can also be used to tag proteins endogenously. Key advantages of this approach are that these tags can be identified using well-characterised commercially available antibodies and the small size of the tag is unlikely to affect protein function in most contexts. However, until recently, this approach was rarely used because targeted insertion of the tag sequence into the mouse genome using traditional gene targeting methods was laborious, time consuming and expensive [2, 5].

Development of genome editing techniques, particularly CRISPR/Cas9, has enabled efficient targeted insertion of short sequences into the mouse genome [8]. The CRISPR/Cas9 system is composed of an endonuclease (Cas9) complexed with a single-guide RNA (sgRNA) which

directs Cas9 to the intended genomic region via complementary base pairing across a 20-nucleotide sequence [9-11]. In the absence of a repair template, Cas9-induced DNA cleavage is typically repaired by non-homologous end joining to generate small indels. However, through provision of a ssDNA donor template, an alternative repair mechanism termed homology-directed repair (HDR) can be used to insert a sequence of interest such as an epitope tag at the site of the DNA break. This technique was first demonstrated in mice through insertion of a V5 tag onto endogenous Sox2 with high efficiency [3].

We have recently used CRISPR/Cas9 mutagenesis to generate a *Pcdh19* "knock-in" (KI) mouse model in which the modified X-linked locus encodes a C-terminal HA-FLAG protein [12]. This model was used to investigate the pathological mechanism of *PCDH19*-epilepsy in which, unusually, heterozygous females develop the disease whereas hemizygous males do not. As a specific antibody for endogenous PCDH19 was unavailable, we used the HA-FLAG KI approach to label "wild type" (WT) cells expressing PCDH19. Through pairing this mouse with a PCDH19 knockout (KO) strain expressing a LacZ reporter cassette, we demonstrated dramatic cell sorting differences in genotypes that replicate the human disease pattern (i.e. heterozygous for PCDH19 KO allele). While optimising immunofluorescence experiments on this mouse, we tested a range of commercially available HA and FLAG antibodies with variable results.

Given that CRISPR/Cas9 technology has made the production of endogenously tagged mice quick and highly efficient, it is likely that many groups will adopt a similar approach [4]. Currently, there is a paucity of data to inform the choice of epitope or antibody for *in vivo* tagging of endogenous proteins [13, 14]. Here we compared the efficacy of several commercially available antibodies for detection of HA and FLAG epitopes using PCDH19$^{HA\text{-}FLAG}$ tissue. Of the

eight antibodies tested, we identified two that enabled robust specific detection of endogenous tagged PCDH19 in the mouse brain.

Methods

PCDH19 mice

PCDH19$^{HA\text{-}FLAG}$ mice were generated using CRISPR/Cas9 as previously described[12]. The HA-FLAG sequence (TACCCATACGATGTTCCAGATTACGCTGACTACAAAGACGATGACGACAAG encoding YPYDVPDYADYKDDDDK) is inserted immediately 5' of the *Pcdh19* stop codon. All animal work was conducted in accordance with institutional and national guidelines for the use of animals for scientific purposes following approval by the University of Adelaide animal ethics committee (approval number S-2013-187).

Cell culture and transfections

PCDH19$^{HA\text{-}FLAG}$-GFP plasmid was prepared by subcloning the mouse PCDH19 cDNA into the pIRES2.eGFP plasmid. The GFP-only plasmid is "empty" pIRES2.eGFP. Cos-1 cells (ATCC, CRL-1650) were grown in 24-well trays on glass coverslips in DMEM supplemented with 10% FCS and 1% glutamax. Cells were transfected with 0.5ug plasmid and 1.5uL 1mg/ml polyethylenimine (PEI; Sigma) and cultured for 48hrs. Cells were washed with PBS, fixed with 4% PFA/PBS and washed 3 times in PBS before blocking.

Immunofluorescence

Pregnant female mice were perfused with 4% PFA/PBS, embryos were harvested and their brains subsequently removed, postfixed for a further 4 hours, cryoprotected in 30% sucrose, and embedded in OCT medium. Sections (12 µm) were prepared using a Leica CM1900 cryostat. Sections were permeabilised with PBS/0.3% Triton X-100 for 10 minutes. Blocking in PBS/0.3% Triton X-100/10% horse serum was performed for 1 hour, followed by overnight

incubation at 4°C in the same solution containing diluted primary antibody. Anti-Nestin antibody (Abcam ab82375) was used as a positive control. Sections or coverslips were washed 3 times for 10 minutes in PBS and incubated in diluted secondary antibody at room temperature for 3 hours. After washing 3 times for 10 minutes each in PBS, slides were mounted in Prolong Gold Antifade plus DAPI (Molecular Probes, Invitrogen) and imaged using a Nikon Eclipse Ti microscope and Nikon Digital Sight DS-Qi1 camera. Antibodies and corresponding dilution are summarised in Table 1.

Statistical Analysis

Quantification of COS cell staining intensity was performed by selecting all cells positive for GFP, then analysing the mean TxRed intensity for GFP-only and PCDH19$^{HA\text{-}FLAG}$GFP conditions using NIS Elements Software. Unpaired T-tests were performed between empty and tag for each antibody using GraphPad Prism 7 software, providing p-values. Intensities were normalized for graphical purposes by dividing each PCDH19$^{HA\text{-}FLAG}$GFP intensity value by the corresponding GFP-only condition mean.

Results

To compare the efficacy of antibodies to detect epitope-tagged endogenous protein in brain tissue, we selected five commercially available antibodies that recognise the HA tag (amino acid sequence: YPYDVPDYA) and three that bind the FLAG tag (DYKDDDDK), generated from a variety of species (Table 1). To confirm the immunoreactivity of each antibody, we transfected COS cells with a mPCDH19$^{HA\text{-}FLAG}$-GFP bicistronic expression plasmid. Negative controls were transfected with CMV-IRES-GFP (GFP only). Antibodies were initially trialled at a uniform dilution (1:300), unless the supplier recommended lesser dilution. Two antibodies showed high levels of background staining at 1:300 so were diluted to 1:3000 where they showed specific signal (data not shown). Immunostaining on COS cells confirmed all eight antibodies bound the HA or FLAG epitope (Figure 1a). For each antibody, staining was quantified by selecting GFP positive cells, quantifying TxRed fluorescence, then normalizing the tag condition to the empty vector control. All antibodies exceeded a mean 2.5 fold intensity difference, with p values <0.0001 (Figure 1b).

Given that all antibodies could detect overexpressed tagged PCDH19, we next tested their ability to detect endogenous PCDH19$^{HA\text{-}FLAG}$ in embryonic brains. To confirm the immunocompetency of the brain tissue, we used an antibody against the intermediate filament protein Nestin, which is expressed by radial glia in the developing brain. All samples that were subsequently used for anti-tag immunostaining (n=3 PCDH19$^{HA\text{-}FLAG}$, n=3 wild type) displayed robust Nestin immunoreactivity (Supplementary Figure 2). Next, CNS tissue from 13.5dpc PCDH19$^{HA\text{-}FLAG}$ and wild type were immunostained using the same antibody dilutions that were optimised in COS cells (Table 1). All anti-tag antibodies showed increased staining in PCDH19$^{HA\text{-}FLAG}$ versus wild type cortical tissue, although many exhibited significant background

staining on wild type tissue (Figure 2). Secondary-only controls confirmed no background staining from goat, rabbit and chicken secondary antibodies though background was evident with mouse secondary (Table 2, Supplementary Figure 3). Each antibody was scored across three biological repeats for its ability to stain specifically in addition to the extent of background staining (Table 2).

Interestingly, only two of the eight antibodies generated a robust signal in PCDH19$^{HA\text{-}FLAG}$ sections with little to no background on wild type tissue. These antibodies displayed similar staining patterns throughout the developing cortex, further confirming specificity. The most effective was HA3 rabbit antibody, which stained specifically across a coronal section of the 13.5dpc developing brain. The FLAG 6 goat antibody also proved useful for staining endogenous FLAG-tagged protein (Figure 3b). Together these data indicate that (i) several commercially available anti-tag antibodies appear unsuitable for endogenous protein detection applications and (ii) HA-FLAG epitope tagged protein can be detected with specific antibodies against the HA or FLAG tag.

Discussion

The inability to detect endogenous proteins is a vexed issue for many researchers. Genome editing technologies have greatly simplified the generation of mice with small epitope tags on a protein of interest, helping to overcome this issue. These advances include the recently reported "SLENDR" technique that utilises CRISPR/Cas9 mediated HDR to insert a HA or FLAG tag, resulting in high-resolution and multiplexed mapping of protein localisation *in vivo*[4]. Others have used zinc finger nuclease (ZFN) or TALEN genome editing techniques to insert HA tags onto endogenous mouse genes and assess protein expression in either embryos or brain tissue[6, 7]. A significant issue with endogenous tagging is that it relies heavily upon the availability of antibodies that bind specifically to the epitope tag. Our experience with the PCDH19$^{HA\text{-}FLAG}$ mouse demonstrated that many commercially available antibodies for HA and FLAG epitope tags are not suitable for staining endogenous tagged protein on mouse tissue[12].

In this study, we compared HA and FLAG antibodies for their ability to specifically detect PCDH19$^{HA\text{-}FLAG}$ both overexpressed in COS cells and at endogenous levels in mouse embryonic brain tissue. All antibodies were effective at staining overexpressed PCDH19$^{HA\text{-}FLAG}$ in COS cells, with mean staining intensity ranging from 2.5-13.5-fold over negative controls, confirming their expected immunoreactivity. Many of these antibodies have been used extensively *in vitro* and are also tested in an overexpression context by suppliers.

However, when we examined the antibodies on tagged PCDH19 at endogenous levels in embryonic brain tissue, many antibodies did not perform effectively. Aside from HA3 and FLAG6, all remaining antibodies displayed significant background staining *in vivo*. At the dilution tested, mouse antibodies (HA1, HA5, FLAG7) stained with high background though this

can be attributed partly to cross-reactivity of secondary antibody (as evidenced in secondary-only controls, Supplementary Figure 3). The remaining three antibodies (HA2, HA4 and FLAG8) also showed non-specific staining on wild type tissue despite secondary only controls showing no staining on the same tissue.

HA3 and FLAG6 antibodies proved to be the most useful for detecting the PCDH19$^{HA\text{-}FLAG}$ protein on mouse embryonic tissue. Though slight cross-reactivity occurred on wild type tissue, particularly for FLAG6, both provide specific staining patterns consistent with *Pcdh19* expression in the developing brain without producing significant background staining. HA3, along with HA5 and FLAG8 have been used effectively for immunofluorescence on mouse brain tissue with single cells containing endogenous epitope tagged protein, though with HA3 and HA5 at slightly higher dilutions (1:1000) [4].

Whilst the comparison shown here provides two antibodies that provide robust detection of endogenous HA or FLAG epitopes, many factors can influence the success of immunofluorescence, making the choice of antibody (and epitope tag) for *in vivo* endogenous tagging experiments a complex one. Antibody performance will undoubtedly vary depending on the expression level of the endogenous protein of interest, the location/accessibility of the epitope tag and the tissue type analysed. It is also possible that further optimisation of staining conditions including antibody dilution for those that showed background on wild type tissue may improve their signal to noise ratio.

Our work and that of and others clearly demonstrates the utility of epitope tags for visualizing a protein of interest *in vivo* [4, 12]. Here, we compared a series of commercial antibodies for their ability to stain endogenous small epitope tagged protein, showing that one HA (HA3) and one

FLAG (FLAG6) antibody were effective in the context of the PCDH19$^{HA\text{-}FLAG}$ mouse brain. The staining comparison presented here provides a simple reference guide for researchers undertaking similar experiments and provides not just two candidate antibodies, but a logical starting point for trialling antibodies and/or epitope tags in a tissue of interest.

Author Contributions

DP conceived and designed the study, performed cell transfections and tissue preparation. LR designed study, performed tissue preparation and all immunofluorescence experiments, analysed results and wrote manuscript. PT designed study, analysed results and wrote manuscript. All authors reviewed and approved manuscript.

Funding

This work was supported by funding from the Australian National Health and Medical Research Council.

Table 1: Commercially available antibodies for HA and FLAG included in this study

Antibody number	Company	Source	Type	Catalogue number	Recommended dilution	Dilution used
HA1	Cell Signalling	Mouse	Monoclonal 6E2	2367	1:100	1:100
HA2	Abcam	Goat	Polyclonal	ab9134	Not provided	1:300
HA3	Cell Signalling	Rabbit	Monoclonal C29F4	3724	1:1600	1:300
HA4	Abcam	Chicken	Polyclonal	ab9111	Not provided	1:300
HA5	Covance	Mouse	Monoclonal HA.11	MMA-101R	1:1000	1:300
FLAG6	Abcam	Goat	Polyclonal	ab1257	1:200-1:2000	1:3000
FLAG7	Sigma	Mouse	Monoclonal M2	F3165	1:200	1:200
FLAG8	Sigma	Rabbit	Polyclonal	F7425	1:100	1:3000

Table 2: Summary of staining pattern on tissue for antibodies tested.

Antibody	Correct staining pattern in tag condition	WT tissue background	Secondary only condition background
HA1	****	**++++**	**+++**
HA2	****	**+++**	**+**
HA3	****	**–**	**–**
HA4	****	**++**	**–**
HA5	****	**+++**	**+++**
FLAG6	****	**+**	**+**
FLAG7	****	**+++**	**++**
FLAG8	**	**++**	**+**

Key:

****	very specific	-	no background
**	less specific	+	low background
		++	medium background
		+++	high background
		++++	very high background

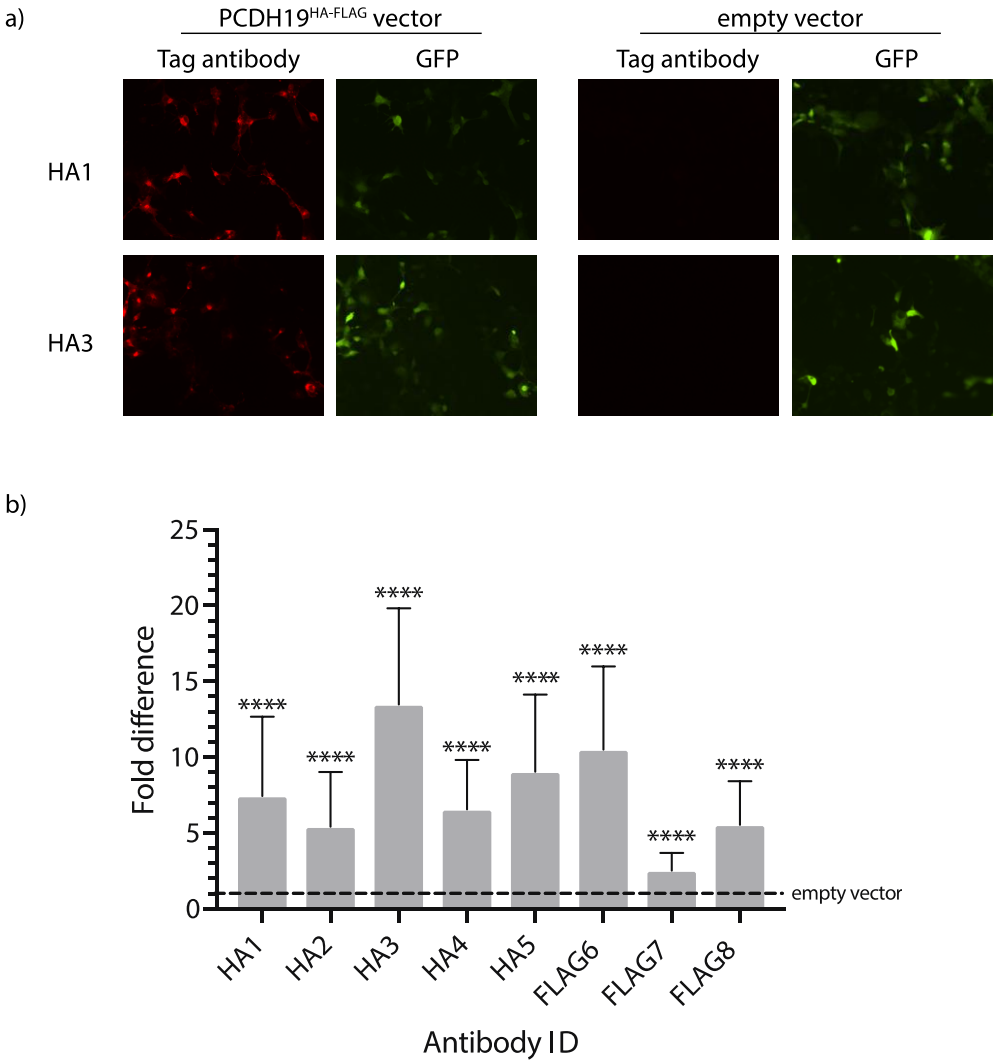


Figure 1: Immunofluorescent staining on COS cells transfected with CMV-PCDH19$^{HA\text{-}FLAG}$ or empty vector. a) Example of COS cell stains demonstrating strong staining in PCHD19$^{HA\text{-}FLAG}$ conditions and not in empty vector controls. Corresponding GFP fluorescence demonstrates transfected cells. b) Quantification of COS cell stains for fold TxRed intensity over empty vector controls. Unpaired t tests revealed p-values<0.0001 for all antibodies tested.

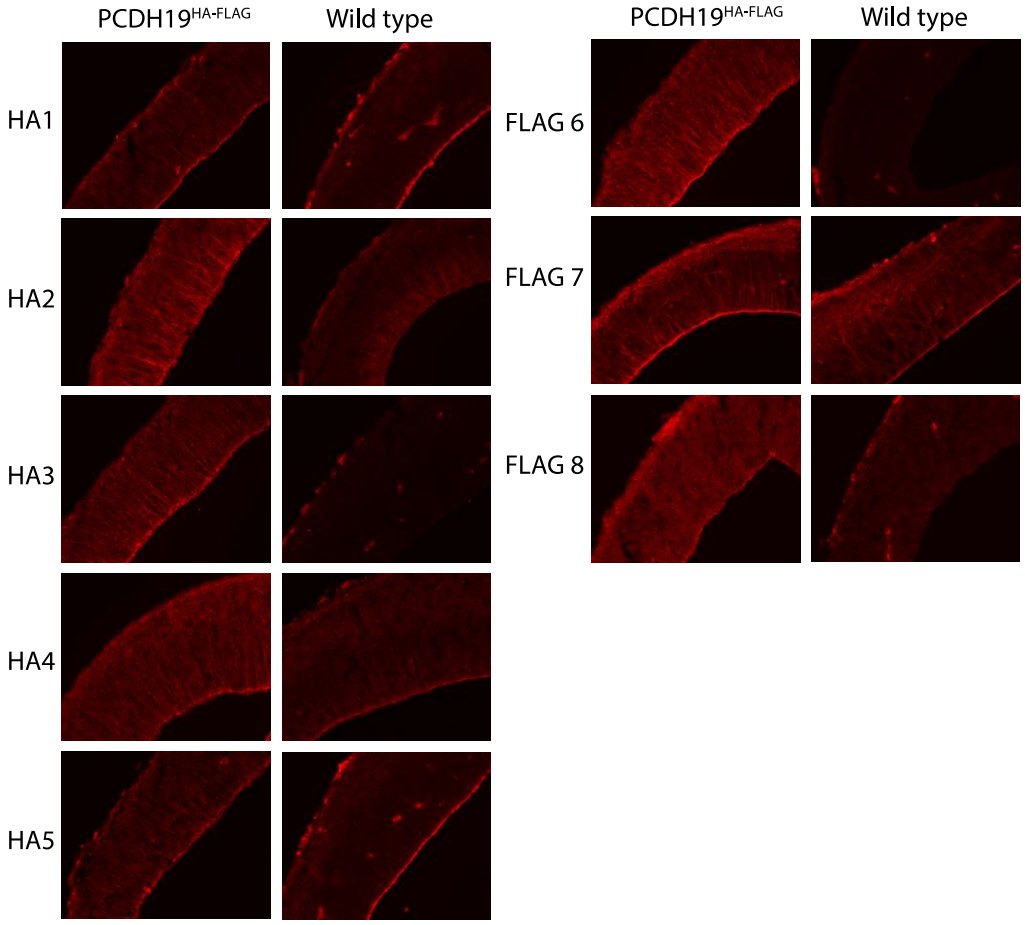


Figure 2: Comparison of immunofluorescence on mouse cortex from PCDH19$^{HA\text{-}FLAG}$ and wild type mice using HA and FLAG commercial antibodies. Brain tissue from 13.5dpc embryos n=3 PCDH19$^{HA\text{-}FLAG}$, n=3 wild type.

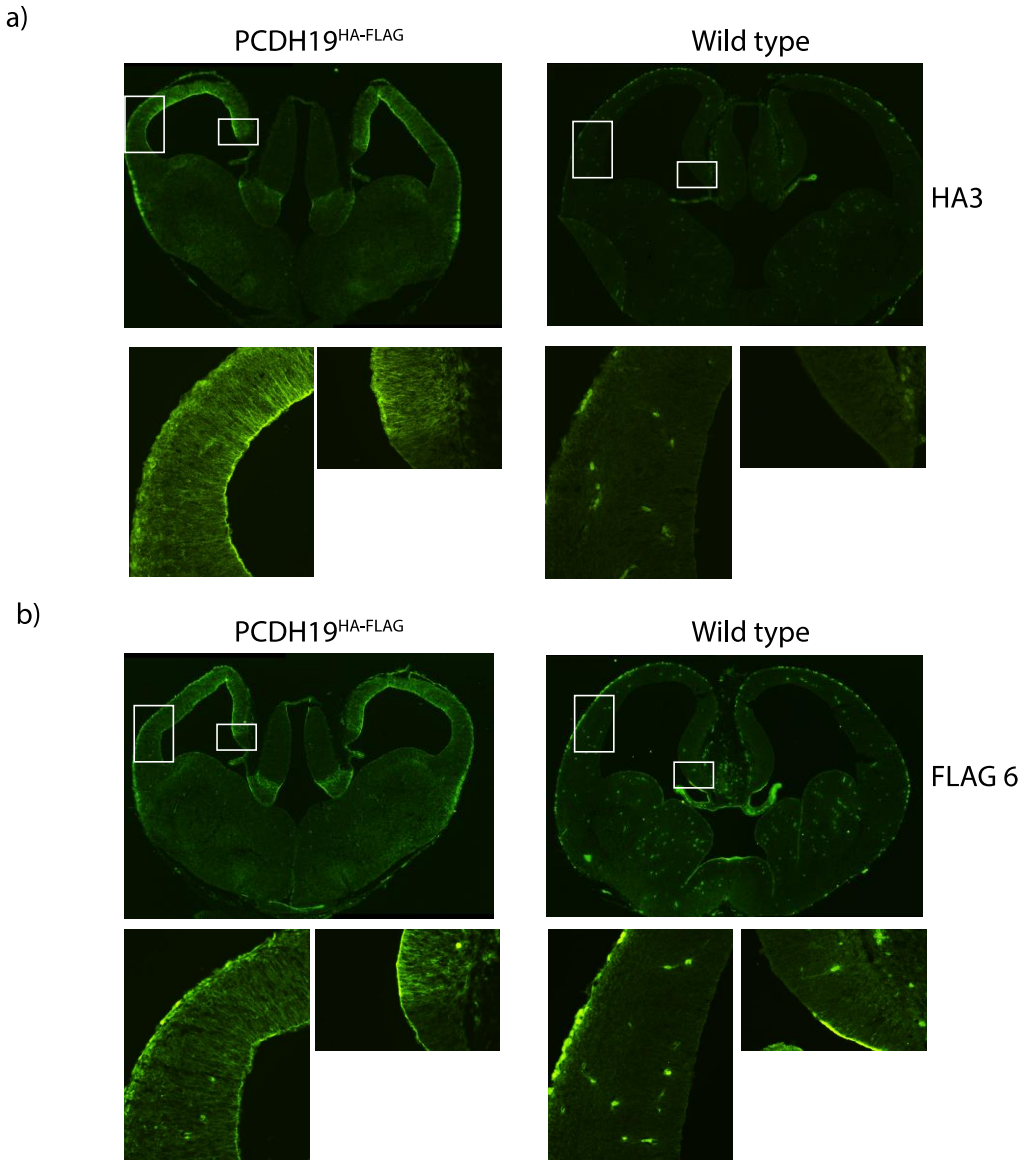


Figure 3: Immunofluorescence on mouse cortex from PCDH19$^{HA\text{-}FLAG}$ and wild type mice demonstrating specific staining by HA3 and FLAG6 commercial antibodies. Brain tissue from 13.5dpc embryos.

References

1. Stadler C, Rexhepaj E, Singan VR et al. Immunofluorescence and fluorescent-protein tagging show high correlation for protein localization in mammalian cells. *Nat Methods.* 2013;10:315-323. DOI: 10.1038/nmeth.2377

2. Lee WJ, Kraus P, Lufkin T. Endogenous tagging of the murine transcription factor Sox5 with hemaglutinin for functional studies. *Transgenic Res.* 2012;21:293-301. DOI: 10.1007/s11248-011-9531-9

3. Yang H, Wang H, Shivalila CS et al. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell.* 2013;154:1370-1379. DOI: 10.1016/j.cell.2013.08.022

4. Mikuni T, Nishiyama J, Sun Y et al. High-Throughput, High-Resolution Mapping of Protein Localization in Mammalian Brain by In Vivo Genome Editing. *Cell.* 2016;165:1803-1817. DOI: 10.1016/j.cell.2016.04.044

5. Rao A, Richards TL, Simmons D et al. Epitope-tagged dopamine transporter knock-in mice reveal rapid endocytic trafficking and filopodia targeting of the transporter in dopaminergic axons. *FASEB J.* 2012;26:1921-1933. DOI: 10.1096/fj.11-196113

6. Su D, Wang M, Ye C et al. One-step generation of mice carrying a conditional allele together with an HA-tag insertion for the delta opioid receptor. *Sci Rep.* 2017;7:44476. DOI: 10.1038/srep44476

7. Wen D, Noh KM, Goldberg AD et al. Genome editing a mouse locus encoding a variant histone, H3.3B, to report on its expression in live animals. *Genesis.* 2014;52:959-966. DOI: 10.1002/dvg.22827

8. Wang H, Yang H, Shivalila CS et al. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell.* 2013;153:910-918. DOI: 10.1016/j.cell.2013.04.025

9. Jinek M, Chylinski K, Fonfara I et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science.* 2012;337:816-821. DOI: 10.1126/science.1225829

10. Cong L, Ran FA, Cox D et al. Multiplex genome engineering using CRISPR/Cas systems. *Science.* 2013;339:819-823. DOI: 10.1126/science.1231143

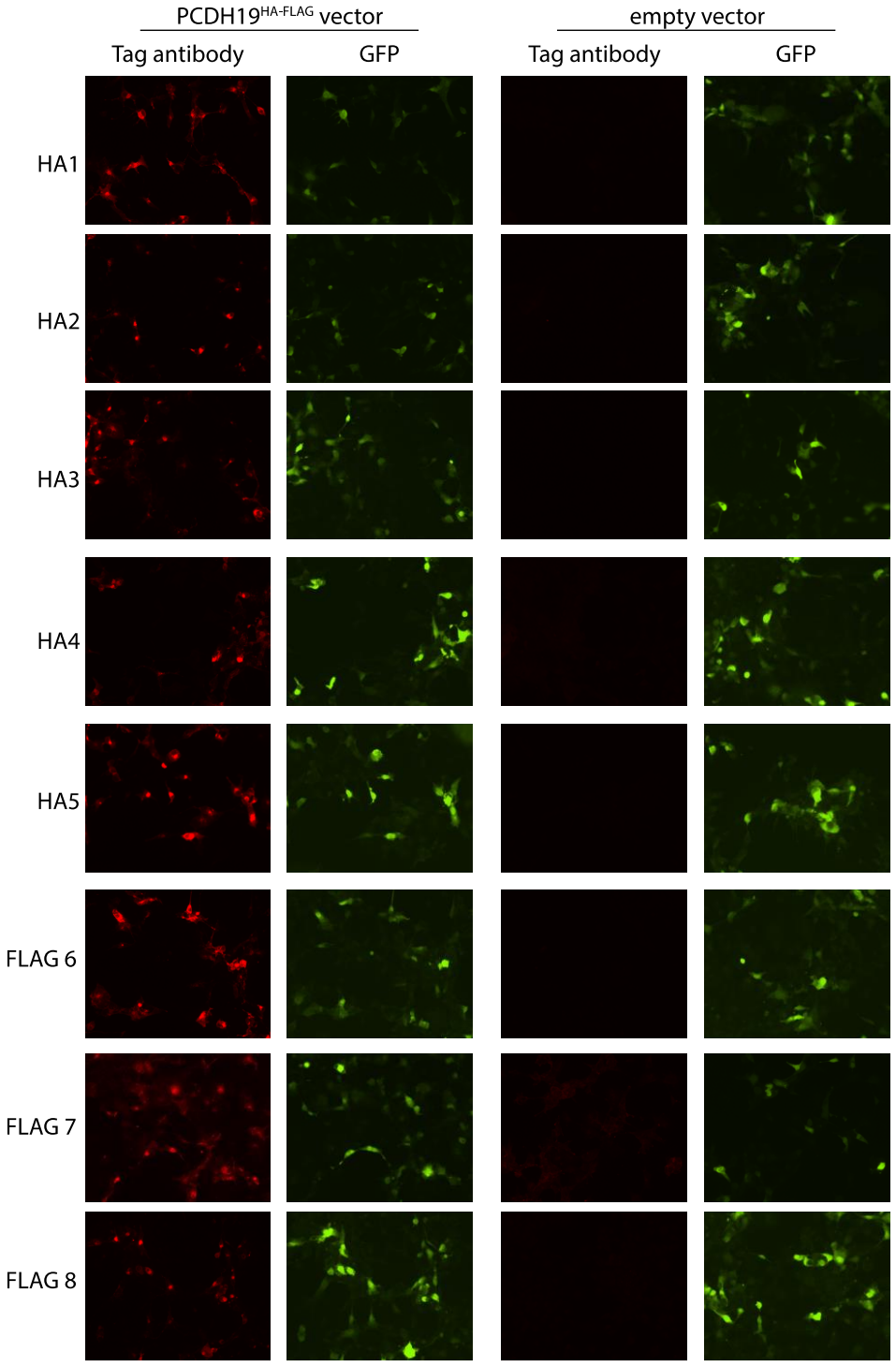
11. Mali P, Yang L, Esvelt KM et al. RNA-guided human genome engineering via Cas9. *Science.* 2013;339:823-826. DOI: 10.1126/science.1232033

12. Pederick DT, Richards KL, Piltz SG et al. Abnormal Cell Sorting Underlies the Unique X-Linked Inheritance of PCDH19 Epilepsy. *Neuron.* 2018;97:59-66 e55. DOI: 10.1016/j.neuron.2017.12.005

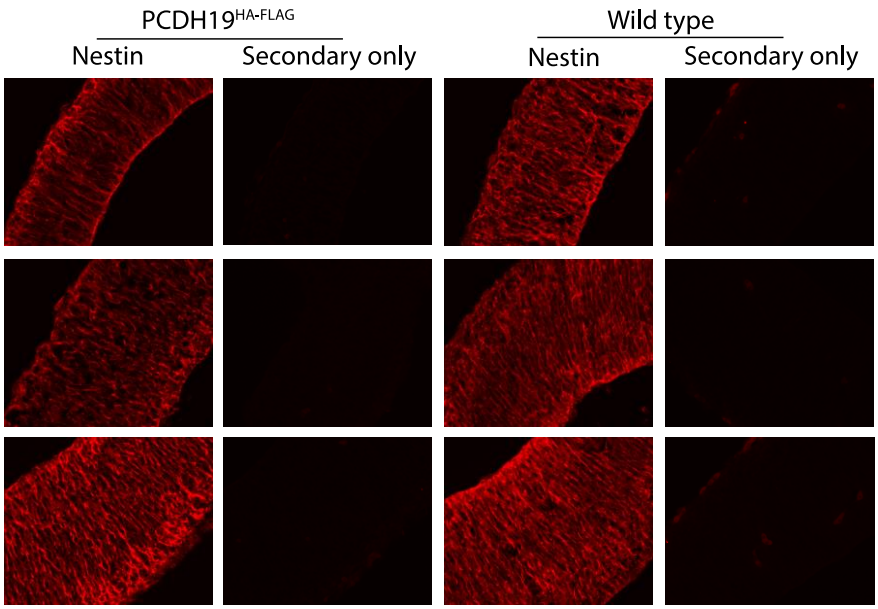
13. Lobbestael E, Reumers V, Ibrahimi A et al. Immunohistochemical detection of transgene expression in the brain using small epitope tags. *BMC Biotechnol.* 2010;10:16. DOI: 10.1186/1472-6750-10-16

14. Shevtsova Z, Malik JM, Michel U et al. Evaluation of epitope tags for protein detection after in vivo CNS gene transfer. *Eur J Neurosci.* 2006;23:1961-1969. DOI: 10.1111/j.1460-9568.2006.04725.x

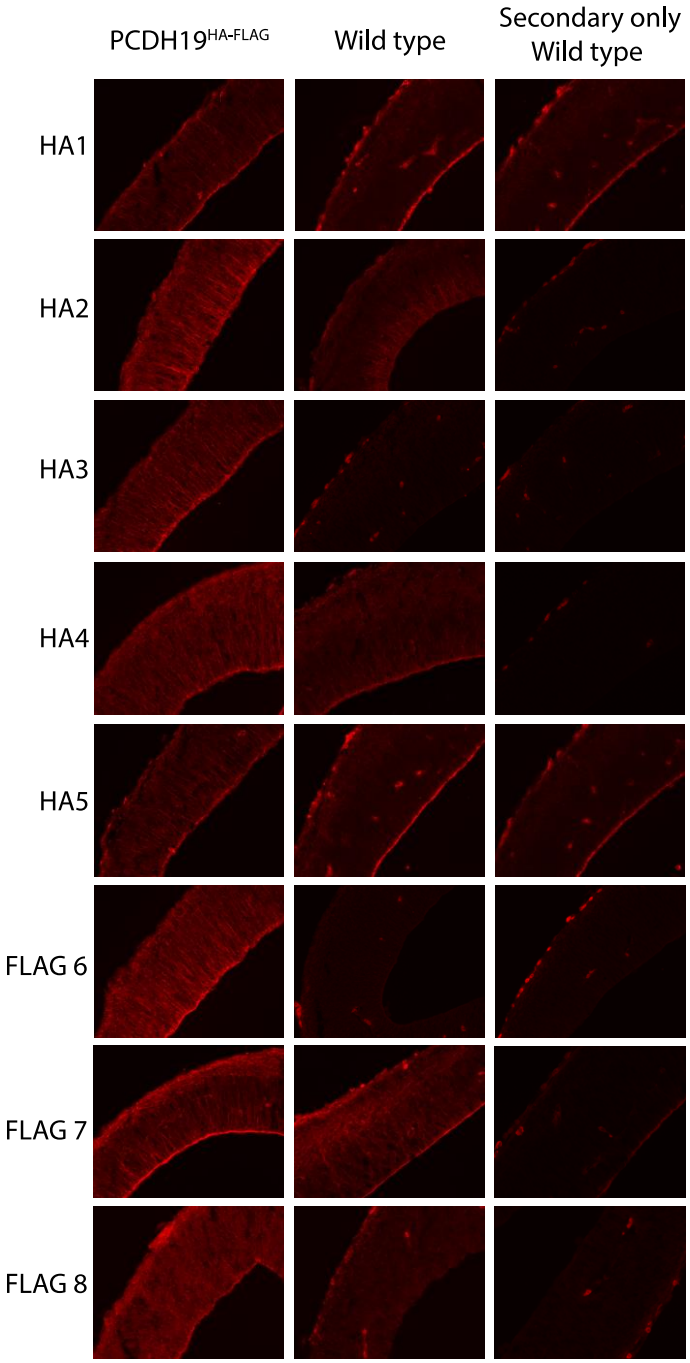
Supplementary Material



Supplementary Figure 1: Immunofluorescent staining on Cos cells transfected with CMV-PCDH19$^{HA\text{-}FLAG}$ or empty vector using each of the selected commercial antibodies. GFP staining indicates transfected cells.



Supplementary Figure 2: Immunofluorescent staining using Nestin antibody on tissue samples used for screening. n=3 per condition.



Supplementary Figure 3: Immunofluorescence on mouse cortex from PCDH19$^{HA\text{-}FLAG}$ and wild type mice using HA and FLAG commercial antibodies, including secondary only controls. Brain tissue from 13.5dpc embryos n=3 PCDH19$^{HA\text{-}FLAG}$, n=3 wild type.

This page is left intentionally blank.

Chapter 5:
Discussion and future directions

5.1 Paroxysmal phenotypes in *Prrt2* mouse models

In Chapter 2 we described a series of phenotypes in the *Prrt2* KO mouse that reflect diseases observed in patients with mutations in *PRRT2*. Most notably, we observed spontaneous paroxysms and unexplained death in *Prrt2* KO animals, with deaths also occurring in a small number of HET animals. The vast majority of patients with *PRRT2*-related disorders have heterozygous mutations in *PRRT2*, with only a few reported homozygous patients, who experience increased severity of the same episodic phenotypes. While the HET mice in our study did not display spontaneous paroxysms, our stronger KO phenotype reflects the increased severity seen in homozygous patients. Recent studies on mice lacking *Prrt2* support this, with Michetti et al. observing audiogenic paroxysms in KO animals only, and Tan et al. describing induced dyskinesia in all homozygous *Prrt2*^{STOP} animals and in fewer heterozygotes[27, 33].

Our study extends the series of behavioural abnormalities in *Prrt2* mice that reflect the spectrum of diseases observed in patients. Table 5.1.1 summarises the tests that displayed differences in HET or KO animals, comparing the outcomes of these across studies. Whilst our behavioural testing results remain relatively consistent with the Michetti et al. investigation of the *Prrt2* KO mouse, Tan et al. describe deficits in rotarod performance and shorter latency to PTZ-induced seizure that were not observed in the other two studies. The models are maintained on different C57Bl/6 substrains that are known to have genetic and phenotypic differences that may affect performance in these tests[34].

Table 5.1.1: Behavioural phenotypes described in three studies on *Prrt2* mice.

Phenotype in homozygous mice	**Michetti et al.**	**Tan et al.**	**Robertson et al. manuscript**
Mutation type	Transgenic knockout	CRISPR c649-Stop	Transgenic knockout
Spontaneous seizures/paroxysms	Developing pups: Increased - Bouncing - Loss of balance - Backward locomotion - Increased grooming Adults: Increased - Loss of balance - Backward locomotion	Spontaneous motor paroxysms in small percentage	Spontaneous motor paroxysms and seizures in small percentage Unexpected death
Induced seizures	PTZ: - No difference in seizure latency but longer duration of seizure in KO Audiogenic: - KO animals, wild running after sound with no EEG trace during events	PTZ - Shorter latency to seizure Kindling: - KO and HET, dyskinetic attacks after seizure Hyperthermia: - KO and HET, dyskinetic attacks after seizure	PTZ - No difference in seizure latency
Motor/coordination phenotype	Rotarod: - No differences Gait - Footprint normal	Rotarod: - KO shorter latency to fall Gait - Footprint normal Challenging beam - KO impaired performance	Rotarod: - No differences Gait: - Digigait- extensive differences at slower speeds (15,20cm/s)
Intellectual Disability			Morris Water Maze: - KO learning deficits

Each of the three papers study spontaneous paroxysmal events in the mice, all of which have a movement disorder component. There are inherent difficulties in defining/investigating a spontaneous phenotype in mice, particularly one that is so infrequent. One of the clinical differences of PKD over an epileptic seizure is the lack of unusual EEG activity during the event, however the rarity of the events makes EEG analysis impractical. We instead rely on events observed/captured by chance and subsequent interpretations of the behavioural manifestation. A similar problem comes to the fore with the unexpected death phenotype we observed in *Prrt2* KO mice. The age of death was variable, and deaths occurred in less than 10% of mice, making monitoring for events difficult. The C57BL/6N genetic background that our mice are maintained on is known to be relatively seizure resistant[93]. To obtain a stronger paroxysmal phenotype in mice, the *Prrt2* KO allele could be backcrossed to a more seizure-susceptible background such as DBA/2. A *Prrt2* loss of function allele could also be produced using CRISPR/Cas9 mutagenesis with wild type DBA/2 zygotes to save the time and expense of backcrossing for numerous generations. If the phenotype in DBA/2 *Prrt2* KO mice is more pronounced, it will be more straightforward to capture deaths and paroxysmal events, allowing better characterisation of these phenotypes.

While sudden unexpected death in epilepsy (SUDEP) is one explanation for the premature death phenotype, it is also possible that *Prrt2* mice experience periods of status epilepticus (seizure episodes lasting longer than 30 minutes) and die as a result. In humans, status epilepticus is considered a cause of death, excluding these instances from being classed as SUDEP [94]. An example of another genetic mouse model that displays SUDEP is that for Dravet syndrome (*Scn1a* loss of function), in which continuous ECG recordings show altered heart rates preceding death[95]. *Scn1a* encodes a sodium channel that is expressed in the central

nervous system and the heart. Given that PRRT2 was recently shown not to interact with Scn1a sodium channel $Na_v1.1$ and that *Prrt2* expression levels are very low/absent in the heart, it is likely that the cause of death differs between the two models. Neither SUDEP nor status epilepticus are common in patients with *Prrt2* mutations, however one patient with probable SUDEP has been described[96].

It is still not understood how mutations in *PRRT2* result in the spectrum of episodic disorders seen in patients, in particular, how epilepsy can occur in infancy and the distinct movement disorder in adulthood. As there is no apparent genotype-phenotype correlation, it has been suggested that the age-dependency of the disorders is due to differing levels of *Prrt2* expression over time and in different brain regions. Protein and transcript analysis have shown differing expression across time in the mouse brain (expression peaks at P14) and western blot data from Tan et al. suggests that PRRT2 may persist at higher levels into adulthood in the cerebellum than in whole brain[16, 27]. A more extensive quantitative analysis of protein expression across time and in specific brain regions may reveal expression changes that could explain why certain phenotypes predominate at different ages. This can be paired with phenotypic analysis of *Prrt2* HET and KO mice at the same time points to determine whether paroxysmal phenotypes, cognitive deficits or gait abnormalities are also age-dependent.

The *Prrt2*^{STOP} mouse genetically mimics the c.649dupC frameshift mutation observed in more than 80% of patients, however numerous other nonsense and missense mutations have been linked to BFIE, PKD and ICCA. Evidence from *in vitro* studies on selected frameshift and missense mutations suggest that failure to be transported to the membrane may be a feature of some frameshift mutations, but not missense mutant R308C[28]. This mutant did however

show significantly reduced protein levels suggesting this may be the cause for disease. Future studies could investigate the functional significance of selected missense mutations in mice. This could be achieved through CRISPR/Cas9 mediated knock-in of point mutations in mice and comparison of the resulting behavioural phenotypes to *Prrt2* KO or *Prrt2*^{STOP} mice. Analysis of protein levels for different mutants, as well as assessment of sub-cellular localisation in brain slices or primary neurons, would determine if mislocalisation or lowered expression is responsible for the pathogenicity of the mutation. It is also possible that mutant protein acts through a dominant negative mechanism, by binding and interfering with wild type PRRT2, however such interaction has not been demonstrated. Electrophysiological recordings from the cerebellum of missense mutant mice would reveal whether the phenotypes resulting from the *Prrt2*^{STOP} mutation in mice are common to other *Prrt2* mutations. These phenotypes included higher frequency miniature excitatory postsynaptic currents (mEPSCs) and increased synaptic facilitation.

Recent findings around the molecular function of PRRT2 have begun to fill the gaps in understanding how mutations may lead to disease. Studies in PC12 cells indicate that PRRT2 may be acting to inhibit SNARE proteins at the synapse and that its absence or mutation results in increases in vesicle fusion/exocytosis[31]. This increased release of neurotransmitters leads to higher neuronal excitability- a key feature of epilepsy. The inhibitory role of PRRT2 on Na^+ channels $Na_v1.2$ and $Na_v1.6$ also points to how mutations in *PRRT2* can lead to over-excitability in neurons[32]. If PRRT2 is mutated or absent and unable to limit Na^+ channel availability on the membrane, an increase in neuronal excitability could result.

The role of PRRT2 as a calcium sensor at the synapse- with loss of PRRT2 resulting in a lowered neurotransmitter release - may be having a separate and opposite effect on excitability. However, if release probability is reduced specifically in inhibitory neurons it is likely that the overall effect could be overexcitability. Interestingly, PRRT2 did not modulate $Na_V1.1$ which predominates in inhibitory neurons, suggesting that the overexcitability may not occur in these neuron types, possibly compounding the imbalance of excitation/inhibition[32, 97]. A complex interplay between these functions likely contributes to the variability in paroxysmal phenotypes observed in patients.

5.2 Increased flexibility for mouse genome editing

In Chapter 3, we tested new CRISPR variants with different PAM recognition for mouse genome editing. Whilst we showed that all variants (WT SpCas9, SpCas9 VQR, SpCas9 VRER, AsCpf1 and SaCas9 KKH) were capable of editing the mouse genome, some variants showed low editing efficiency. All endonucleases were targeted to an overlapping region in the *Ngn3* intron with the aim of reducing potential locus-specific differences in cleavage efficiency. However, variants will need to be tested at more genomic loci to gain a better understanding of mouse genome editing efficiency (especially those with lowest efficiency; VQR and VRER). These variants cleave efficiently in human cell lines, suggesting that the mutation rate we observed in mice can be improved. Lower activity may be due to sgRNA sequence-specific differences, or in the case of VQR, may be due to PAM recognition differences as a result of the flexibility in the PAM sites (recognises a NGAN PAM). The VQR variant is most effective at targeting NGAG PAM sites in mammalian cells, while the NGAT site (used in our screen) is the second preferred PAM[58]. Consistent with this, in *Caenorhabditis elegans* the VQR variant favoured NGAG PAM sites over NGAT or NGAA, with NGAC PAM sites being the least effective for eliciting mutations[98]. Trialling the VQR variant with its optimal PAM sequence will be useful to understand the overall targeting efficiency in the mouse genome.

We have recently become aware of another form of the VQR variant that has a quadruple mutation (D1135V/G11218R/R1335Q/T1337R), known as the VRQR variant[76]. This variant has improved activity on NGAA, NGAC and NGAT PAM sites, when compared to VQR, making its PAM recognition site more accurately NGAN. This will need to be trialled for mouse genome editing but may prove a more flexible option for targeting these types of PAM sites.

The selection of endonucleases to trial in this screen was based on the PAM recognition sequence and its ability to expand targetable regions of the mouse genome. Validation of the variants that were not trialled will still be important as they may prove to have higher targeting efficiency and therefore be useful additions to the mouse genome editing 'toolkit'. Some of these variants (eg. St1Cas9; PAM site NNAGAAW) have long PAM recognition sequences which have been hypothesised to reduce nuclease off-targets, another important factor to consider when selecting an endonuclease for genome editing[99, 100].

Unlike *in vitro* genome editing, the impact of off-target mutations in mice is reduced by the fact that mutations can be segregated away through breeding (provided they are not closely linked to the on-target site). However, this process can be time consuming and expensive, so minimisation of these events is desirable. Whilst testing variants for off-target activity was outside of the scope of this manuscript, it will be interesting to compare the specificity of these endonucleases on the mouse genome. It is conceivable that each of the available PAM variants can be engineered in a similar way to high fidelity SpCas9 variants (eg HF1, eSpCas9 1.1, HypaCas9) to reduce their off-target activity. Indeed, the VQR and VRQR variant mutations discussed above have been combined with the HF1 mutated residues to create high-fidelity PAM variants that show comparable on-target activities, though these still remain to be tested for genome-wide specificity[76].

When looking to introduce a point mutation or stop codon, another way to reduce off-target activity is using the base editor (BE3), consisting of SpCas9 nickase fused to rat cytidine deaminase APOBEC1 and uracil glycosylase inhibitor (UGI), which facilitates the substitution of a C-T or G-A[101]. Indel formation at off-target sites are minimised, though off-target base editing

does occur at similar rates. This has been shown to be highly efficient in mice, making it an attractive choice over HDR experiments (which often have low efficiency) if the appropriate base change is required[102]. A drawback of this technique is the requirement of a PAM sequence within ~3-5 bases of the desired base substitution. The engineered and naturally occurring CRISPR variants that are available can be used in a similar way to expand the PAM recognition options for precision base editing. Recently, the same group have described updated versions of the base editor (BE4) with higher on-target and lower off-target activity, as well as a new editor than can convert A-G or T-C, adding to the ever-increasing flexibility of the CRISPR toolkit[103, 104]. Future studies will need to focus on extending editing options to allow all combinations of nucleotide substitution and determining whether new technologies will be viable options for application in the production of genetically modified mice.

An intriguing finding of our CRISPR variant study, was the propensity of SaCas9 KKH to edit only one allele. We propose that this feature will be most useful when performing HDR experiments to insert a loss of function point mutation into a gene that causes nullizygous embryonic lethality. This phenomenon is likely due to a 'sweet spot' in the efficiency of the nuclease; high enough to induce mutations in sufficient offspring but low enough to prevent production of multiple mutations. Given this point of difference to WT SpCas9, it will be useful to compare the activity of WT SaCas9 under these circumstances. WT SaCas9 has been shown to edit the mouse genome efficiently, though the screening methods used made it unclear as to whether a single or both alleles were cleaved most frequently[65]. As the PAM sites for SaCas9 KKH and WT SaCas9 are similar, many SaCas9 KKH targets will have a PAM compatible with WT SaCas9, allowing a direct comparison of the two variants. For proof of principle, it will also be interesting to perform a comparison of WT SpCas9, SaCas9 KKH and WT SaCas9 for their

activity on genes in which loss of function causes embryonic lethality. Experiments on this type of gene were not chosen for our study as we risked missing events if lethality occurred before tissue collection. It would be expected that SaCas9 KKH (and perhaps also WT SaCas9) will allow the production of a deleterious point mutation allele in the absence of an indel on the other allele, meaning live founders are produced for establishing a colony.

5.3 Epitope tagging endogenous proteins for visualisation

In Chapter 4 we compared a series of commercial antibodies for their effectiveness in detecting endogenously tagged PCDH19. We demonstrated one HA and one FLAG antibody that specifically stained PCDH19$^{HA\text{-}FLAG}$ in mouse brain tissue. The remaining six antibodies displayed background staining on wild type tissue that may confound localisation studies. This study will have broad utility in assisting researchers in their experimental design as CRISPR/Cas9 technology has enabled simple and affordable production of mice with endogenous epitope tagged proteins.

Small epitope tags will also make *in vivo* immunoprecipitation experiments more straightforward. This technique is heavily reliant on the availability of a specific antibody for the endogenous protein of interest. We have used this technique effectively on the PCDH19$^{HA\text{-}FLAG}$ mouse to show that PCDH19 interacts with PCDH10 and PCDH17 using a FLAG antibody in place of a PCDH19 antibody[105]. For DNA binding proteins, ChIP assays could be performed using commercially available antibodies against an epitope tag rather than relying on an antibody for the chosen protein.

An interesting application for endogenous epitope tagging that has recently arisen is the SLENDR technique (single-cell labelling of endogenous proteins by CRISPR/Cas9-mediated homology-directed repair). CRISPR/Cas9 reagents and HDR templates containing epitope tags for insertion are introduced by *in utero* electroporation into the developing mouse brain. This allows visualisation of protein localisation via light microscopy and even electron microscopy for high resolution mapping of protein localisation, without the time-consuming process of producing mouse colonies with endogenous epitope tags. The tags are also inserted into fewer

cells, allowing sub-cellular mapping in sparse neurons without the issue of dense tissue and cell overlap interrupting visualisation. The antibodies validated in Chapter 4 will be particularly useful for this technique, where labelling of a specific pattern is required in mouse brain tissue.

A point of concern when tagging a protein with an epitope tag is that the epitope may affect the normal function of the protein, confounding interpretations of endogenous gene function or abolishing activity all together. Studies have revealed altered protein structure and interactions as a result of small epitope tags[106, 107]. The HA tag also has a caspase cleavage site and is cleaved during apoptosis[108]. To overcome this issue, a recent study by Georgieva et al. described new epitope tags called 'inntags' (innocuous protein tags)[109]. Inntags are longer (40-96 amino acids) than FLAG or HA epitopes (8 and 9 amino acids respectively) but, unlike FLAG or HA, are self-structured protein domains, overcoming the high propensity for disorder. Researchers showed that inntags did not cause incorrect folding or mislocalisation of proteins and had less effect on function than commonly used strings of six FLAG or HA tags. Specific mouse monoclonal antibodies were developed for these tags allowing immunofluorescence and immunoprecipitation. Whilst promising, the use of these tags across a more diverse range of experiments will determine whether they consistently show less functional impact than FLAG or HA tags.

Future experiments could test these *in vivo* by using CRISPR technology to insert an inntag at the terminus of a protein of interest. Inntag length is likely to make insertion more challenging, however the PAM variants validated in Chapter 3 will give flexibility in where a DSB can be induced, allowing a cut site at or close to the insertion site for most efficient insertion. In order

for inntags to be useful for endogenous protein localisation and investigation, the antibodies against the tags will need to become readily available.

5.4 Concluding remarks

The data presented in this thesis highlights the utility and flexibility of mouse models to investigate gene function and disease. We describe phenotypes in the *Prrt2* KO mouse that reflect the hallmarks observed in patients, further establishing it as a model for *Prrt2*-related diseases. Chapter 3 examines new approaches for production of genetically modified mice using CRISPR technology. We test a series of endonuclease variants with different PAM recognition for their ability to create targeted mutations in mice and also uncover a new feature of the SaCas9 KKH variant to cleave a single allele. The final study in this thesis assesses the efficacy of antibodies for immunofluorescence on mice with a PCDH19 endogenous epitope tag, showing that not all commercially available antibodies are suitable for this application.

Much of the future for mouse genome modification lies with CRISPR/Cas9 genome editing. As more PAM variants and altered forms become available, virtually any mutation will be possible at any loci, including specialised insertions such as endogenous epitope tags. Importantly, with the ease of production improving, a bottleneck in productivity will now move downstream to the phenotypic characterisation of genetically modified mouse strains; a process that can be slow and expensive. This, however, remains essential to fully uncover a gene's function and role in disease.

References

1. Valtorta F, Benfenati F, Zara F et al. PRRT2: from Paroxysmal Disorders to Regulation of Synaptic Function. *Trends Neurosci.* 2016;39:668-679. DOI: 10.1016/j.tins.2016.08.005

2. Komor AC, Badran AH, Liu DR. CRISPR-Based Technologies for the Manipulation of Eukaryotic Genomes. *Cell.* 2017;169:559. DOI: 10.1016/j.cell.2017.04.005

3. Cohen J. Mice made easy. *Science.* 2016;354:538-542. DOI: 10.1126/science.354.6312.538

4. Agrotis A, Ketteler R. A new age in functional genomics using CRISPR/Cas9 in arrayed library screening. *Front Genet.* 2015;6:300. DOI: 10.3389/fgene.2015.00300

5. Hoban MD, Bauer DE. A genome editing primer for the hematologist. *Blood.* 2016;127:2525-2535. DOI: 10.1182/blood-2016-01-678151

6. Jasin M, Rothstein R. Repair of strand breaks by homologous recombination. *Cold Spring Harb Perspect Biol.* 2013;5:a012740. DOI: 10.1101/cshperspect.a012740

7. Hesdorffer DC, Logroscino G, Benn EK et al. Estimating risk for developing epilepsy: a population-based study in Rochester, Minnesota. *Neurology.* 2011;76:23-27. DOI: 10.1212/WNL.0b013e318204a36a

8. Chen T, Giri M, Xia Z et al. Genetic and epigenetic mechanisms of epilepsy: a review. *Neuropsychiatr Dis Treat.* 2017;13:1841-1859. DOI: 10.2147/NDT.S142032

9. Thomas RH, Berkovic SF. The hidden genetics of epilepsy-a clinically important new paradigm. *Nat Rev Neurol.* 2014;10:283-292. DOI: 10.1038/nrneurol.2014.62

10. Focus on epilepsy. *Nat Neurosci.* 2015;18:317. DOI: 10.1038/nn.3964

11. Wen D, Noh KM, Goldberg AD et al. Genome editing a mouse locus encoding a variant histone, H3.3B, to report on its expression in live animals. *Genesis.* 2014;52:959-966. DOI: 10.1002/dvg.22827

12. Heron SE, Dibbens LM. Role of PRRT2 in common paroxysmal neurological disorders: a gene with remarkable pleiotropy. *J Med Genet.* 2013;50:133-139. DOI: 10.1136/jmedgenet-2012-101406

13. Becker F, Schubert J, Striano P et al. PRRT2-related disorders: further PKD and ICCA cases and review of the literature. *J Neurol.* 2013;260:1234-1244. DOI: 10.1007/s00415-012-6777-y

14. Ebrahimi-Fakhari D, Saffari A, Westenberger A et al. The evolving spectrum of PRRT2-associated paroxysmal diseases. *Brain.* 2015;138:3476-3495. DOI: 10.1093/brain/awv317

15. Bruno MK, Hallett M, Gwinn-Hardy K et al. Clinical evaluation of idiopathic paroxysmal kinesigenic dyskinesia: new diagnostic criteria. *Neurology.* 2004;63:2280-2287.

16. Chen WJ, Lin Y, Xiong ZQ et al. Exome sequencing identifies truncating mutations in PRRT2 that cause paroxysmal kinesigenic dyskinesia. *Nat Genet.* 2011;43:1252-1255. DOI: 10.1038/ng.1008

17. Szepetowski P, Rochette J, Berquin P et al. Familial infantile convulsions and paroxysmal choreoathetosis: a new neurological syndrome linked to the pericentromeric region of human chromosome 16. *Am J Hum Genet.* 1997;61:889-898. DOI: 10.1086/514877

18. Caraballo R, Pavek S, Lemainque A et al. Linkage of benign familial infantile convulsions to chromosome 16p12-q12 suggests allelism to the infantile convulsions and choreoathetosis syndrome. *Am J Hum Genet.* 2001;68:788-794. DOI: 10.1086/318805

19. Heron SE, Grinton BE, Kivity S et al. PRRT2 mutations cause benign familial infantile epilepsy and infantile convulsions with choreoathetosis syndrome. *Am J Hum Genet.* 2012;90:152-160. DOI: 10.1016/j.ajhg.2011.12.003

20. Lee HY, Huang Y, Bruneau N et al. Mutations in the gene PRRT2 cause paroxysmal kinesigenic dyskinesia with infantile convulsions. *Cell Rep.* 2012;1:2-12. DOI: 10.1016/j.celrep.2011.11.001

21. Riant F, Roze E, Barbance C et al. PRRT2 mutations cause hemiplegic migraine. *Neurology.* 2012;79:2122-2124. DOI: 10.1212/WNL.0b013e3182752cb8

22. Liu Q, Qi Z, Wan XH et al. Mutations in PRRT2 result in paroxysmal dyskinesias with marked variability in clinical expression. *J Med Genet.* 2012;49:79-82. DOI: 10.1136/jmedgenet-2011-100653

23. Najmabadi H, Hu H, Garshasbi M et al. Deep sequencing reveals 50 novel genes for recessive cognitive disorders. *Nature.* 2011;478:57-63. DOI: 10.1038/nature10423

24. Labate A, Tarantino P, Viri M et al. Homozygous c.649dupC mutation in PRRT2 worsens the BFIS/PKD phenotype with mental retardation, episodic ataxia, and absences. *Epilepsia.* 2012;53:e196-199. DOI: 10.1111/epi.12009

25. Rossi P, Sterlini B, Castroflorio E et al. A Novel Topology of Proline-rich Transmembrane Protein 2 (PRRT2): hints for an intracellular function at the synapse. *J Biol Chem.* 2016;291:6111-6123. DOI: 10.1074/jbc.M115.683888

26. Valente P, Castroflorio E, Rossi P et al. PRRT2 Is a Key Component of the Ca(2+)-Dependent Neurotransmitter Release Machinery. *Cell Rep.* 2016;15:117-131. DOI: 10.1016/j.celrep.2016.03.005

27. Tan GH, Liu YY, Wang L et al. PRRT2 deficiency induces paroxysmal kinesigenic dyskinesia by regulating synaptic transmission in cerebellum. *Cell Res.* 2017. DOI: 10.1038/cr.2017.128

28. Liu YT, Nian FS, Chou WJ et al. PRRT2 mutations lead to neuronal dysfunction and neurodevelopmental defects. *Oncotarget.* 2016;7:39184-39196. DOI: 10.18632/oncotarget.9258

29. Stelzl U, Worm U, Lalowski M et al. A human protein-protein interaction network: a resource for annotating the proteome. *Cell.* 2005;122:957-968. DOI: 10.1016/j.cell.2005.08.029

30. Sudhof TC. Neurotransmitter release: the last millisecond in the life of a synaptic vesicle. *Neuron.* 2013;80:675-690. DOI: 10.1016/j.neuron.2013.10.022

31. Coleman J, Jouannot O, Ramakrishnan SK et al. PRRT2 Regulates Synaptic Fusion by Directly Modulating SNARE Complex Assembly. *Cell Rep.* 2018;22:820-831. DOI: 10.1016/j.celrep.2017.12.056

32. Fruscione F, Valente P, Sterlini B et al. PRRT2 controls neuronal excitability by negatively modulating Na+ channel 1.2/1.6 activity. *Brain.* 2018;141:1000-1016. DOI: 10.1093/brain/awy051

33. Michetti C, Castroflorio E, Marchionni I et al. The PRRT2 knockout mouse recapitulates the neurological diseases associated with PRRT2 mutations. *Neurobiol Dis.* 2017;99:66-83. DOI: 10.1016/j.nbd.2016.12.018

34. Simon MM, Greenaway S, White JK et al. A comparative phenotypic and genomic analysis of C57BL/6J and C57BL/6N mouse strains. *Genome Biol.* 2013;14:R82. DOI: 10.1186/gb-2013-14-7-r82

35. Rosenthal N, Brown S. The mouse ascending: perspectives for human-disease models. *Nat Cell Biol.* 2007;9:993-999. DOI: 10.1038/ncb437

36. Palmer EE, Jarrett KE, Sachdev RK et al. Neuronal deficiency of ARV1 causes an autosomal recessive epileptic encephalopathy. *Hum Mol Genet.* 2016;25:3042-3054. DOI: 10.1093/hmg/ddw157

37. Penagarikano O, Abrahams BS, Herman EI et al. Absence of CNTNAP2 leads to epilepsy, neuronal migration abnormalities, and core autism-related deficits. *Cell.* 2011;147:235-246. DOI: 10.1016/j.cell.2011.08.040

38. Tan HO, Reid CA, Single FN et al. Reduced cortical inhibition in a mouse model of familial childhood absence epilepsy. *Proc Natl Acad Sci U S A.* 2007;104:17536-17541. DOI: 10.1073/pnas.0708440104

39. Capecchi MR. Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. *Nat Rev Genet.* 2005;6:507-512. DOI: 10.1038/nrg1619

40. Bradley A, Anastassiadis K, Ayadi A et al. The mammalian gene function resource: the International Knockout Mouse Consortium. *Mamm Genome.* 2012;23:580-586. DOI: 10.1007/s00335-012-9422-2

41. Rouet P, Smih F, Jasin M. Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol Cell Biol.* 1994;14:8096-8106.

42. Tsai SQ, Joung JK. Defining and improving the genome-wide specificities of CRISPR-Cas9 nucleases. *Nat Rev Genet.* 2016;17:300-312. DOI: 10.1038/nrg.2016.28

43. Carroll D. Genome engineering with zinc-finger nucleases. *Genetics.* 2011;188:773-782. DOI: 10.1534/genetics.111.131433

44. Kim YG, Cha J, Chandrasegaran S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc Natl Acad Sci U S A.* 1996;93:1156-1160.

45. Carbery ID, Ji D, Harrington A et al. Targeted genome modification in mice using zinc-finger nucleases. *Genetics.* 2010;186:451-459. DOI: 10.1534/genetics.110.117002

46. Meyer M, de Angelis MH, Wurst W et al. Gene targeting by homologous recombination in mouse zygotes mediated by zinc-finger nucleases. *Proc Natl Acad Sci U S A.* 2010;107:15022-15026. DOI: 10.1073/pnas.1009424107

47. Davies B, Davies G, Preece C et al. Site specific mutation of the Zic2 locus by microinjection of TALEN mRNA in mouse CD1, C3H and C57BL/6J oocytes. *PLoS One.* 2013;8:e60216. DOI: 10.1371/journal.pone.0060216

48. Barrangou R, Fremaux C, Deveau H et al. CRISPR provides acquired resistance against viruses in prokaryotes. *Science.* 2007;315:1709-1712. DOI: 10.1126/science.1138140

49. Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol.* 2014;32:347-355. DOI: 10.1038/nbt.2842

50. Sternberg SH, Redding S, Jinek M et al. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature.* 2014;507:62-67. DOI: 10.1038/nature13011

51. Jinek M, Chylinski K, Fonfara I et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science.* 2012;337:816-821. DOI: 10.1126/science.1225829

52. Mali P, Yang L, Esvelt KM et al. RNA-guided human genome engineering via Cas9. *Science.* 2013;339:823-826. DOI: 10.1126/science.1232033

53. Cong L, Ran FA, Cox D et al. Multiplex genome engineering using CRISPR/Cas systems. *Science.* 2013;339:819-823. DOI: 10.1126/science.1231143

54. Nishimasu H, Ran FA, Hsu PD et al. Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell.* 2014;156:935-949. DOI: 10.1016/j.cell.2014.02.001

55. Jinek M, Jiang F, Taylor DW et al. Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science.* 2014;343:1247997. DOI: 10.1126/science.1247997

56. Paquet D, Kwart D, Chen A et al. Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature.* 2016;533:125-129. DOI: 10.1038/nature17664

57. Hirano H, Gootenberg JS, Horii T et al. Structure and Engineering of Francisella novicida Cas9. *Cell.* 2016;164:950-961. DOI: 10.1016/j.cell.2016.01.039

58. Kleinstiver BP, Prew MS, Tsai SQ et al. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature.* 2015;523:481-485. DOI: 10.1038/nature14592

59. Ran FA, Cong L, Yan WX et al. In vivo genome editing using Staphylococcus aureus Cas9. *Nature.* 2015;520:186-191. DOI: 10.1038/nature14299

60. Esvelt KM, Mali P, Braff JL et al. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat Methods.* 2013;10:1116-1121. DOI: 10.1038/nmeth.2681

61. Zetsche B, Gootenberg JS, Abudayyeh OO et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell.* 2015;163:759-771. DOI: 10.1016/j.cell.2015.09.038

62. Kleinstiver BP, Prew MS, Tsai SQ et al. Broadening the targeting range of Staphylococcus aureus CRISPR-Cas9 by modifying PAM recognition. *Nat Biotechnol.* 2015;33:1293-1298. DOI: 10.1038/nbt.3404

63. Kim Y, Cheong SA, Lee JG et al. Generation of knockout mice by Cpf1-mediated gene targeting. *Nat Biotechnol.* 2016;34:808-810. DOI: 10.1038/nbt.3614

64. Hur JK, Kim K, Been KW et al. Targeted mutagenesis in mice by electroporation of Cpf1 ribonucleoproteins. *Nat Biotechnol.* 2016;34:807-808. DOI: 10.1038/nbt.3596

65. Zhang X, Liang P, Ding C et al. Efficient Production of Gene-Modified Mice using Staphylococcus aureus Cas9. *Sci Rep.* 2016;6:32565. DOI: 10.1038/srep32565

66. Fu Y, Sander JD, Reyon D et al. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol.* 2014;32:279-284. DOI: 10.1038/nbt.2808

67. Tsai SQ, Zheng Z, Nguyen NT et al. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol.* 2015;33:187-197. DOI: 10.1038/nbt.3117

68. Kim S, Kim D, Cho SW et al. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res.* 2014;24:1012-1019. DOI: 10.1101/gr.171322.113

69. Lin S, Staahl BT, Alla RK et al. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *Elife.* 2014;3:e04766. DOI: 10.7554/eLife.04766

70. Mali P, Aach J, Stranges PB et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol.* 2013;31:833-838. DOI: 10.1038/nbt.2675

71. Ran FA, Hsu PD, Lin CY et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell.* 2013;154:1380-1389. DOI: 10.1016/j.cell.2013.08.021

72. Tsai SQ, Wyvekens N, Khayter C et al. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat Biotechnol.* 2014;32:569-576. DOI: 10.1038/nbt.2908

73. Guilinger JP, Thompson DB, Liu DR. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat Biotechnol.* 2014;32:577-582. DOI: 10.1038/nbt.2909

74. Wyvekens N, Topkar VV, Khayter C et al. Dimeric CRISPR RNA-Guided FokI-dCas9 Nucleases Directed by Truncated gRNAs for Highly Specific Genome Editing. *Hum Gene Ther.* 2015;26:425-431. DOI: 10.1089/hum.2015.084

75. Slaymaker IM, Gao L, Zetsche B et al. Rationally engineered Cas9 nucleases with improved specificity. *Science.* 2016;351:84-88. DOI: 10.1126/science.aad5227

76. Kleinstiver BP, Pattanayak V, Prew MS et al. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature.* 2016;529:490-495. DOI: 10.1038/nature16526

77. Chen JS, Dagdas YS, Kleinstiver BP et al. Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature.* 2017. DOI: 10.1038/nature24268

78. Davis AJ, Chen DJ. DNA double strand break repair via non-homologous end-joining. *Transl Cancer Res.* 2013;2:130-143. DOI: 10.3978/j.issn.2218-676X.2013.04.02

79. Deriano L, Roth DB. Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage. *Annu Rev Genet.* 2013;47:433-455. DOI: 10.1146/annurev-genet-110711-155540

80. Betermier M, Bertrand P, Lopez BS. Is non-homologous end-joining really an inherently error-prone process? *PLoS Genet.* 2014;10:e1004086. DOI: 10.1371/journal.pgen.1004086

81. Decottignies A. Alternative end-joining mechanisms: a historical perspective. *Front Genet.* 2013;4:48. DOI: 10.3389/fgene.2013.00048

82. Shen B, Zhang J, Wu H et al. Generation of gene-modified mice via Cas9/RNA-mediated gene targeting. *Cell Res.* 2013;23:720-723. DOI: 10.1038/cr.2013.46

83. Singh P, Schimenti JC, Bolcun-Filas E. A mouse geneticist's practical guide to CRISPR applications. *Genetics.* 2015;199:1-15. DOI: 10.1534/genetics.114.169771

84. Wang H, Yang H, Shivalila CS et al. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell.* 2013;153:910-918. DOI: 10.1016/j.cell.2013.04.025

85. Fujii W, Kawasaki K, Sugiura K et al. Efficient generation of large-scale genome-modified mice using gRNA and CAS9 endonuclease. *Nucleic Acids Res.* 2013;41:e187. DOI: 10.1093/nar/gkt772

86. Zhang L, Jia R, Palange NJ et al. Large genomic fragment deletions and insertions in mouse using CRISPR/Cas9. *PLoS One.* 2015;10:e0120396. DOI: 10.1371/journal.pone.0120396

87. Birling MC, Schaeffer L, Andre P et al. Efficient and rapid generation of large genomic variants in rats and mice using CRISMERE. *Sci Rep.* 2017;7:43331. DOI: 10.1038/srep43331

88. Yang H, Wang H, Shivalila CS et al. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell.* 2013;154:1370-1379. DOI: 10.1016/j.cell.2013.08.022

89. Mikuni T, Nishiyama J, Sun Y et al. High-Throughput, High-Resolution Mapping of Protein Localization in Mammalian Brain by In Vivo Genome Editing. *Cell.* 2016;165:1803-1817. DOI: 10.1016/j.cell.2016.04.044

90. Yang H, Wang H, Jaenisch R. Generating genetically modified mice using CRISPR/Cas-mediated genome engineering. *Nat Protoc.* 2014;9:1956-1968. DOI: 10.1038/nprot.2014.134

91. Quadros RM, Miura H, Harms DW et al. Easi-CRISPR: a robust method for one-step generation of mice carrying conditional and insertion alleles using long ssDNA donors and CRISPR ribonucleoproteins. *Genome Biol.* 2017;18:92. DOI: 10.1186/s13059-017-1220-4

92. Pederick DT, Richards KL, Piltz SG et al. Abnormal Cell Sorting Underlies the Unique X-Linked Inheritance of PCDH19 Epilepsy. *Neuron.* 2018;97:59-66 e55. DOI: 10.1016/j.neuron.2017.12.005

93. Deckard BS, Lieff B, Schlesinger K et al. Developmental patterns of seizure susceptibility in inbred strains of mice. *Dev Psychobiol.* 1976;9:17-24. DOI: 10.1002/dev.420090104

94. Nashef L, So EL, Ryvlin P et al. Unifying the definitions of sudden unexpected death in epilepsy. *Epilepsia.* 2012;53:227-233. DOI: 10.1111/j.1528-1167.2011.03358.x

95. Auerbach DS, Jones J, Clawson BC et al. Altered cardiac electrophysiology and SUDEP in a model of Dravet syndrome. *PLoS One.* 2013;8:e77843. DOI: 10.1371/journal.pone.0077843

96. Labate A, Tarantino P, Palamara G et al. Mutations in PRRT2 result in familial infantile seizures with heterogeneous phenotypes including febrile convulsions and probable SUDEP. *Epilepsy Res.* 2013;104:280-284. DOI: 10.1016/j.eplepsyres.2012.10.014

97. Ogiwara I, Miyamoto H, Morita N et al. Nav1.1 localizes to axons of parvalbumin-positive inhibitory interneurons: a circuit basis for epileptic seizures in mice carrying an Scn1a gene mutation. *J Neurosci.* 2007;27:5903-5914. DOI: 10.1523/jneurosci.5270-06.2007

98. Bell RT, Fu BX, Fire AZ. Cas9 Variants Expand the Target Repertoire in Caenorhabditis elegans. *Genetics.* 2016;202:381-388. DOI: 10.1534/genetics.115.185041

99. Muller M, Lee CM, Gasiunas G et al. Streptococcus thermophilus CRISPR-Cas9 Systems Enable Specific Editing of the Human Genome. *Mol Ther.* 2016;24:636-644. DOI: 10.1038/mt.2015.218

100. Kim E, Koo T, Park SW et al. In vivo genome editing with a small Cas9 orthologue derived from Campylobacter jejuni. *Nat Commun.* 2017;8:14500. DOI: 10.1038/ncomms14500

101. Komor AC, Kim YB, Packer MS et al. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature.* 2016;533:420-424. DOI: 10.1038/nature17946

102. Kim K, Ryu SM, Kim ST et al. Highly efficient RNA-guided base editing in mouse embryos. *Nat Biotechnol.* 2017;35:435-437. DOI: 10.1038/nbt.3816

103. Komor AC, Zhao KT, Packer MS et al. Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity. *Sci Adv.* 2017;3:eaao4774. DOI: 10.1126/sciadv.aao4774

104. Gaudelli NM, Komor AC, Rees HA et al. Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. *Nature.* 2017. DOI: 10.1038/nature24644

105. Pederick D, Richards K, Piltz S et al. Abnormal cell sorting underlies the unique X-linked inheritance of PCDH19 Epilepsy. *bioRxiv.* 2017. DOI: 10.1101/178822

106. Chant A, Kraemer-Pecore CM, Watkin R et al. Attachment of a histidine tag to the minimal zinc finger protein of the Aspergillus nidulans gene regulatory protein AreA

causes a conformational change at the DNA-binding site. *Protein Expr Purif.* 2005;39:152-159. DOI: 10.1016/j.pep.2004.10.017

107. Song J, Markley JL. Cautionary tail: the presence of an N-terminal tag on dynein light-chain Roadblock/LC7 affects its interaction with a functional partner. *Protein Pept Lett.* 2007;14:265-268.

108. Schembri L, Dalibart R, Tomasello F et al. The HA tag is cleaved and loses immunoreactivity during apoptosis. *Nat Methods.* 2007;4:107-108. DOI: 10.1038/nmeth0207-107

109. Georgieva MV, Yahya G, Codo L et al. Inntags: small self-structured epitopes for innocuous protein tagging. *Nat Methods.* 2015;12:955-958. DOI: 10.1038/nmeth.3556