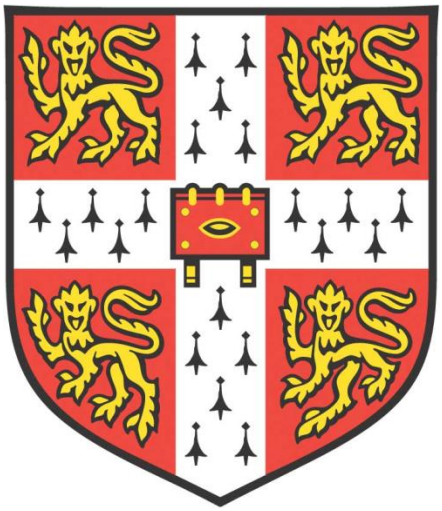
APPLYING ANIMAL MODELLING TO UNDERSTAND RARE NOVEL NEURODEVELOPMENTAL DISORDERS ASSOCIATED WITH INTELLECTUAL DISABILITY



Maria Levitin

St. John's College

University of Cambridge

Wellcome Sanger Institute

This dissertation is submitted for the degree of *Doctor of Philosophy.*

September 2017

Applying animal modelling to understand rare novel neurodevelopmental disorders associated with intellectual disability

Maria Levitin

Intellectual disability (ID) is categorised by a significant reduction in cognitive function and adaptive abilities that begin in childhood. ID is part of a heterogeneous group of neurodevelopmental conditions associated with impairment in developmental domains and a cause of particularly adverse socioeconomic impact worldwide. There have been many recent advances in identifying causative genetic mutations in previously unexplained ID cases. With these advances comes an increasing demand for understanding mechanisms underpinning these pathogenic pathways. In this PhD thesis, I have studied rare monogenic novel neurodevelopmental disorders associated with ID. The objective of the thesis was to model a subset of mutations associated with novel neurodevelopmental disorders in mice to demonstrate a causal link between mutation and phenotype and to further understand the mechanisms by which these mutations result in human neurodevelopmental disorders. In order to achieve this, I adopted a multi-phase approach. Firstly, I designed a phenotyping platform, by combining behavioural and cognitive tests with morphometric brain analysis and genome-wide transcriptional analysis. I then used this approach to study *KPTN*-related syndrome, a novel developmental disorder that to date has not been characterised in mice, successfully recapitulating the main phenotypes described in the patients. Moreover, I gained further insight into the underlying pathogenic mechanisms associated with the disorder, opening the possibility of a therapy that could treat some aspects of cognitive and morphological impairments identified in the patients with *KPTN*-related syndrome. Lastly, I determined whether such an approach could be scaled-up to study multiple novel neurodevelopmental disorders, each with a mutation associated with a haploinsufficient novel neurodevelopmental disorder. I identified specific phenotypes for each of the four mouse lines under investigation, providing a platform for comparison between several developmental disorders. These refinements contributed to a larger five-year project starting at the Sanger Institute, aimed at characterising a wider diversity of human neurodevelopmental disorders.

DECLARATION

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the contributions section within each chapter and/or specified in the text. It is not being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University of similar institution. It does not exceed the prescribed word limit for the Faculty of Biology.

Maria Levitin

September 2017

ACKNOWLEDGEMENTS

The PhD is a transformative journey with many ups and downs. Words cannot do justice to how thankful I am to all the incredible people who were a part of this journey and made it a memorable positive experience.

First of all, I want to thank my supervisor, Darren Logan. Darren guided and supported me from the very first day I decided to try *real* science as an Undergraduate Placement student in his lab, to the last days of thesis writing when he was just a phone call and train journey away despite the physical distance, and beyond the PhD submission. I am grateful for having such a wonderful role model – both on a scientific and a human being level.

I want to thank Matt Hurles for adopting me to his team and giving me all the resources I needed for the smooth running of my PhD project. I want to also say thank you to all the members of the Hurles group for welcoming me and being there whenever I needed anything.

I want to give a big thank you to Sebastian Gerety for taking me under his wing when all the changes happened, countless hours of productive discussions, always encouraging me to do my very best and, leading by example, striving to achieve the highest quality of science possible. It has been very inspiring to have a mentor who is so passionate about what he does.

I would also like to thank my behavioural guru Gabi Gurria for all her mentorship and help over the years. Gabi not only taught me so much about behavioural neuroscience, but she is also one of the nicest people I know; working with her has been truly a pleasure.

I want to thank all the other members of the team 156 – Elizabeth Wynn, Luis Saraiva, Ximena Ibarra-Soria, Cristina Dias, Laura Huckins, Sophia Liang – for all the help and insightful discussions they have provided and all the fun we had in and out of the lab. I also want to thank the cognition sub-team, including Mark Sanderson and Holly Ironfield, for a huge amount of hands-on help with behavioural experiments (Mark) and molecular work (Holly and Elizabeth).

My PhD experience would not be the same without the fantastic friends I made in Cambridge, St. John's College, and at the Sanger Institute.

The Sanger PhD program was an incredible support system throughout the whole PhD experience, and for this, I am thankful to the Graduate Office (Carl Anderson, Annabel Smith, and Christina Hedberg-Delouka) and the PhD community. Shout out goes to PhD13, my PhD year group, and the other wonderful Sanger PhDs. From the daily breakfast, lunches, and courses together, to long gossip sessions and crazy night outs on the choo-choo train – we have been through a lot with these humans and I am very grateful for such a lovely year group.

I want to put a particular emphasis on two groups of PhD friends who also became my elected family. You all know how much I love you so I will try (and fail) to keep it brief. Firstly my Sanger family – Velislava Petrova, Fernando Riveros Mckay Aguilera, Liliana Antunes, Dimitris Garyfallos, Gianmarco Raddi, Christian Owusu, and Alice Mann. Veli has been my EEBmate from the first (unconventional) day we met, she is the person who knows me inside out, and without her, my PhD would have been a completely different experience, so would have my life in general. Fernando is not only one of my best friends and a person I can rely on at any point no matter the circumstance, but he was also my go-to RNA sequencing analysis expert and helped greatly towards the end of my PhD with all the RNA sequencing data that was flooding in. I will be forever in debt to Liliana, for being an incredible friend who I could always rely on, the mode of transport for all weekend trips to the lab (with picnics), a complete life-saver during my thesis submission, and the responsible yet flowery presence in our group. The two BBs, Dimitri and Gianmarco, became my besties pretty much straight away and each in their own way became an important component of my PhD journey and without whom I could not imagine Sanger and Cambridge in general. Christian, my horror movie buddy, has always provided me with fascinating discussions and lots of positivity (even though he tried to mask it with grumpiness). Alice and I shared a love for organising and chats over hot beverages and made a great team throughout the PhD as a consequence.

Secondly, I want to say a big thank you to my St. John's ladies – Rishika Kundra and Molly Krasnodebska. These women were one of the very first people I met as I moved to Cambridge and they made me feel like I was home immediately. We have gone through our PhD journeys hand in hand and having them in my life, and for a long time as my housemates, has truly made a big difference to my time in Cambridge.

I also want to thank all my friends out of Cambridge, near and far, who had to deal with me in PhD mode and were nevertheless always very understanding and supportive. Yes, Trajche you may have the spotlight too buddy, because of your impressive persistence with which you demanded to be mentioned in my thesis. I also want to thank Molly Cranston, for looking over my thesis with unbiased and trained eyes, which was very useful.

I want to give a big and probably not enough times articulated thank you to one very loving and very patient man, Petros Sideris, who was there in the front row seat for all that PhD threw my way and was able to keep me sane throughout the whole process. He also survived the final month of living with a thesis-writing zombie and helped a tremendous amount in the most panic-heavy moments.

And finally, I want to thank the people who are the reason I got to do a PhD in the first place and have made me who I am today. More than words can ever portray – I am forever grateful for having such amazing parents who have always inspired and encouraged me to follow my dreams and believe in myself, and gave me the necessary skills and resources to make this a reality. I also want to thank my grandparents, who have since a young age doomed me to become a scientist by evoking a sense of curiosity for the unknown and the unexplored.

Abbreviations and acronyms

4E-BPs	eIF4E-binding proteins
ADHD	Attention deficit hyperactivity disorder
ASD	Autism spectrum disorder
BCND	Behavioural and Cognitive Neurodevelopmental Disorder
CANTAB	Cambridge neuropsychological test automated battery
CANTAB	Cambridge Neuropsychological Test Automated Battery
CGH	Comparative genomic hybridisation
CNS	Central nervous system
CNVs	Copy number variants
CRF	Chromatin remodelling factors
CS+	Conditioned stimulus
DD	Developmental disorders
DDD	Deciphering Developmental Disorders
DDs	Developmental disorders
DE	Differently expressed
DEPTOR	DEP domain containing mTOR interacting protein
DNM	De novo mutations
DSM-5	Diagnostic and Statistical Manual of Mental Disorders Fifth Edition
F	Inbreeding coefficient
FISH	Fluorescence in situ hybridisation
FMR1	X mental retardation 1
GATOR	GTPase activating protein for RAGA
GDD	Global developmental delay
GWAS	Genome-wide association study
H3K4	Histone H3 at Lysine 4
hIPSC	Human induced pluripotent stem cell
ICD10	International Classification of Diseases
ID	Intellectual disability
KPTN, ITFG2, C12orf66 and SZT2 mTORC1 complex	KICSTOR protein complex
MGP	Mouse Genetics Project
mLST8	Mammalian lethal with Sec13 protein 8
MRI	Magnetic resonance imaging
mSin1	Mammalian stress-activated protein kinase interacting protein 1
mTOR	Mechanistic target of rapamycin
mTORC	mTOR Complex
NGS	Next-generation sequencing

NOR	Novel object recognition
NS-ID	Nonsyndromic Intellectual disability
OD	Object displacement
OFC	Occipital frontal circumference
OFC	Orbitofrontal
P0	Postnatal day 0
PAGE	Prenatal Assessment of Genomes and Exomes
PD	Pairwise visual discrimination
PFC	Prefrontal cortex
Protor	Protein observed with rictor-1 or 2
Raptor	Regulatory protein associated with mTOR
Rheb	Ras-homolog expressed in brain
Rictor	Rapamycin insensitive companion of mTOR
S6K	S6 kinase
SD	Standard deviations
SET domain	Su(var)3-9, Enhancer of Zeste, Trithorax domain
SETD1A	SET Domain Containing 1A
SETD5	SET Domain-Containing Protein 5
S-ID	Syndromic Intellectual disability
SNP	Single-nucleotide-polymorphism
SOR	Social recognition
TSC	Tuberous sclerosis complex 1 and 2
VBM	Voxel-based morphometry
WOH	Windows of Hope
ZMYND11	Zinc finger MYND domain-containing protein 11 ZMYND11

Table of Contents

Table of Figures

Chapter 1. Introduction to intellectual disability

1.1 Preface

In my PhD thesis, I have studied rare monogenic novel neurodevelopmental disorders associated with intellectual disability (ID). There have been many recent advances in identifying causative genetic mutations in previously unexplained intellectual disability cases. However, there is still much to learn about the affected genes and the associated pathogenic pathways. In the context of this thesis, I consider as novel disorders those that have been identified after the start of my PhD, from 2014 onwards.

In this introductory Chapter, I provide a brief background of ID, focusing on genetic causes. I then provide an account of the advancements in genetic diagnosis in recent years and how they have shaped our understanding of the causes of ID. Finally, I outline some similarities and differences in brain ontology between rodents and humans, which are relevant for modelling human disorders with affected brain development in rodents.

I introduce mouse modelling of ID in Chapter 3, describe the use of mouse models to study rare recessive inherited mutations in Chapter 4 and *de novo* dominant mutations in Chapter 5. Chapter 6 contains a concluding discussion on the main findings of the thesis and outlines future areas of research.

1.2 Neurodevelopmental disorders and intellectual disability

Neurodevelopmental disorders (DD) are a heterogeneous group of conditions characterised by impairment in developmental domains, such as cognitive, language, social, and motor, and are associated with a complex set of endophenotypes (Levitt et al., 2004). The shared onset of pathologies in these disorders occurs during the period of maturation and development, which includes both prenatal and postnatal stages (Zoghbi, 2003). The most commonly diagnosed neurodevelopmental disorders, based on the Diagnostic and Statistical Manual of Mental Disorders Fifth Edition (DSM-5), include autism spectrum disorder (ASD),

attention deficit hyperactivity disorder (ADHD), and intellectual disability (ID) (American Psychiatric Association, 2013). ID, or developmental cognitive impairment, is a heterogeneous group of neurodevelopmental disorders categorised by a significant reduction in cognitive function and adaptive abilities that begin in childhood (Ropers, 2010; Salvador-Carulla & Bertelli, 2008). The term *adaptive ability* implies the capacity to carry out daily activities appropriate to the individual's age group.

ID is commonly assessed using IQ tests, such as the Wechsler Intelligence Scale for Children (WISC) or Adults (WAIS), adjusted for age and cultural background. Individuals with IQ scores of 70 or less (2 or more standard deviations below the general population mean score of 100) are diagnosed with ID (Ropers, 2010; Vissers et al., 2016). Based on DSM-5, ID was classified into four severity categories depending on the IQ score of the individual: mild, moderate, severe, and profound. However, studies often use a simplified classification of mild (IQ 50-70) and severe (IQ<50) ID (Chelly et al., 2001; Kaufman et al., 2010). Moreover, the updated 5th DSM edition has moved away from the specific IQ score categorisation, keeping the classifications but placing more emphasis on adaptive function (American Psychiatric Association, 2013). The reliable diagnosis of ID in children younger than six years old is more difficult. Therefore the term *global developmental delay* (GDD) is used instead. GDD is diagnosed when children fail to meet expected developmental milestones. Importantly, while many children diagnosed with GDD later meet ID criteria, they are diagnostically distinct (Numis and Sankar, 2016).

ID patients have an increased risk of developing comorbidities, with an estimated 30% of ID patients also having behavioural and psychiatric conditions (Cooper et al., 2007; Einfeld and Tonge, 1996; White et al., 2005). Due to the presence of other clinical features or comorbidities, ID can be further divided into syndromic (S-ID) and nonsyndromic (NS-ID) (American Psychiatric Association, 2013). The purist definition of NS-ID is when ID is the sole clinical feature. However, it is often difficult to identify other features in ID patients due to masking by the cognitive impairment or the subtlety of the other features. This often blurs the distinction between the NS- and S-ID. Moreover, the causes of these two subgroups often overlap, for example, certain genes associated with NS-ID have also been linked to S-ID cases (Kaufman et al., 2010). Since NS-ID has ID as its only manifestation, many studies

have focused on studying NS-ID as means of understanding processes involving learning and memory and cognition in general (Kaufman et al., 2010).

1.3 ID demographics

ID has a worldwide occurrence of 1-2%, with a higher prevalence of ID reported in lower socioeconomic groups and developing countries, and is diagnosed more frequently in males (Boyle et al., 2011; Van Naarden Braun et al., 2015; Maulik et al., 2011; Durkin, 2012). The former discrepancies are mainly explained by environmental factors, while the latter sex bias may in part be explained by the X-linked causes of ID (Boyle et al., 2011; Emerson, 2007). Due to its worldwide prevalence and chronic nature, ID is a significant socioeconomic burden and is listed by the International Classification of Diseases (ICD10) as the most costly of all diagnoses (Honeycutt et al., 2003; Polder et al. 2002). The average lifetime cost per person with ID is $1-2 million in Europe and the United States, with an estimated total lifetime cost of $51.2 billion in the US for people born in 2000 (Honeycutt et al., 2003).

The occurrence of severe ID is relatively stable in the population (0.3–0.5%) worldwide, while the reported prevalence of mild ID is higher but more variable, as it is influenced by external factors such as access to education and healthcare and is less clear-cut to identify (Chelly et al., 2001; Kaufman et al., 2010). Severe ID is often associated with dysmorphic features and a higher rate of behavioural abnormalities, which is why it is typically identified earlier in childhood than mild ID, which may not be diagnosed until school age (Kaufman et al., 2010). ID can either be non-progressive or worsen with age, due to comorbidities and challenges of transitioning into adulthood, or improve due to early intervention (Jeste, 2015).

1.4 Common causes of ID

ID and other neurodevelopmental disorders are caused by various environmental and genetic factors that affect the development of the nervous system (pre-, peri-, and postnatally). ID is the most frequent reason for genetic service referral, but it is important to consider both genetic and nongenetic etiologies when performing clinical evaluations

(Ehninger et al., 2008; Numis and Sankar, 2016). The environmental factors that cause ID include traumatic brain injury, prenatal and postnatal complications such as hypoxemia, maternal and childhood infections, exposure to toxic substances (pre or postnatally), nutritional deficits, and radiation (Kaufman et al., 2010; Modabbernia et al., 2016). The genetic causes are also varied and account for 25-50% of ID cases, proportionally increasing with severity (Ehninger et al., 2008). ID has a variable effect on reproductive fitness depending on the severity level, which in turn affects the genetic architecture of different ID forms. For this reason, severe genetic forms of ID are mainly sporadic (Ehninger et al., 2008).

Genetic forms of ID involve chromosomal aneusomies and structural abnormalities, X-chromosome linked defects and monogenic diseases. Autosomal and X-chromosome aneuploidies are typically associated with ID, such as Down syndrome, or trisomy 21, which is the most frequent form of ID (Rauch et al., 2006). Overall, chromosomal abnormalities are a common cause of ID, accounting for around 15% of diagnosed cases (Michelson et al., 2011). Since the 1990s, chromosome X has been a focus of attempts to elucidate genetic defects linked with ID. This led to the identification of fragile X mental retardation 1 (*FMR1*) gene as the cause of the most common inherited form of ID, fragile X syndrome (Coffee et al., 2009; Pieretti et al., 1991; Verkerk et al., 1991). With over 100 identified ID genes linked to the X chromosome, the discovery of X-linked ID genes is now approaching a plateau, but many more autosomal ID genes remain to be discovered (Lubs et al., 2012; Vissers et al., 2016). Current genomic approaches have enabled the identification of previously overlooked but highly frequently mutated ID genes *ARID1B* and *DDX3X*, each accounting for more than 1% of ID patients (Blok et al., 2015; Santen et al., 2012).

Advancements in genomic microarray technologies enabled the discovery of pathogenic genomic microdeletions and duplications associated with ID. These copy number variants (CNVs) were previously not resolved by routine chromosome analysis (Albertson and Pinkel, 2003; Grayton et al., 2012; Wagenstaller et al., 2007). The human genome contains around 12% of CNV, which contribute to the variation in individual genomes and not all of which are pathogenic (Iafrate et al., 2004; Redon et al., 2006). However, many *de novo* autosomal variants and X-linked CNVs have now been identified as causative of many ID-associated

disorders, as well as being predisposing factors for neuropsychiatric conditions (Itsara et al., 2009; Mefford et al., 2009). While autosomal-recessive ID forms are not as prevalent in outbred populations, they are the major cause of ID in populations with high rates of consanguinity (de Ligt et al., 2012; Musante and Ropers, 2014). Consanguineous marriage is defined in a clinical genetic setting as a marriage between two closely related individuals (second cousins or closer), with an inbreeding coefficient (*F*) equal or higher than 0.0156 (Hamamy, 2012). *F* is a measure of the proportion of the loci with identical gene copies from both parents that the offspring from a consanguineous union is expected to inherit. Such families are common in countries belonging to the so-called 'consanguinity belt' which comprises regions from Morocco to India, with Pakistan having a particularly high prevalence. Indeed 62.7% of marriages in Pakistan are consanguineous, around 80% of which are between first cousins (Hussain and Bittles, 1998). Children from consanguineous unions have an increased risk of autosomal recessive disorders compared to that of the general population's risk (Hamamy et al., 2011). In non-consanguineous populations, the recessive variants associated with ID are most frequently sporadic and include cases with compound heterozygous mutations (Schuurs-Hoeijmakers et al., 2012; Ten Kate et al., 2010).

1.5 Genetic diagnosis, past and present, and rare ID cases

The advancements in our knowledge of genetic causes of ID have been driven by advancements in the tools available for genetic diagnosis. Initially, genetic diagnosis relied on karyotyping to detect gross chromosome abnormalities, beginning with the discovery of trisomy 21 as the cause of Down's syndrome in 1959 (Lejeune et al., 1959). In the 1970s conventional karyotyping became a routine test and allowed for conclusive genetic diagnosis in up to 6% of ID cases (Vissers et al., 2016). The advancements in cytogenetic chromosome banding technologies and the identification of X chromosome markers increased the efficiency of detection of chromosomal abnormalities, leading to the identification of genetic causes in multiple ID-associated syndromes such as Prader–Willi syndrome (deletion of 15q11-q13) (Butler et al., 1986; Lubs and Ruddle, 1970). The next

step forward in diagnostic yield came with the introduction of Sanger sequencing in the 1970s and the development of targeted fluorescence in situ hybridisation (FISH) in the 1980s, which together increased the diagnostic yield to 6-10% (Vissers et al., 2016).

In the 1990s, the shift from radiolabelling to chromogenic Sanger sequencing increased the identification of monogenic ID cases. The research at the time largely focused on identifying ID causing genes on the X-chromosome due to sex ratio bias and ease of pedigree analysis (Tarpey et al., 2009). An example of two success stories of the period included the identification of *FMR1* in fragile X syndrome and *MECP2* in Rett's syndrome (Amir et al., 1999; Pieretti et al., 1991; Verkerk et al., 1991). Collectively, X-linked ID genes explain 10% of ID in males (Lubs et al., 2012).

Until the introduction of genomic microarrays at the beginning of this century, the research into autosomal causes of ID was lagging behind X-linked causes. Genomic microarrays enabled a better resolution and higher diagnostic yield, replacing the former strategies in the clinic (Miller et al., 2010). Two routinely used chromosome microarrays were comparative genomic hybridisation (CGH) and single-nucleotide-polymorphism (SNP) genotyping arrays, which increased the diagnostic yield to 12-23% (Gilissen et al. 2014). For recessive forms of ID, homozygosity mapping and high-density SNP microarrays, with follow-up Sanger sequencing of candidate genes, facilitated accurate and rapid detection of further recessive autosomal causes of ID (Schuurs-Hoeijmakers et al., 2011).

Sanger sequencing has been used extensively in the last three decades and was used to produce the first complete human genome sequence (Collins and McKusick, 2001; Kelavkar, 2001; Venter et al., 2001). However, the technology has relatively low throughput. The introduction of next-generation sequencing (NGS) technologies in the mid-2000s, revolutionised the field by providing powerful high-throughput tools for successful detection of causative autosomal *de novo* and recessive variants in unexplained ID cases. This increased diagnostic yield to over 30%, with 55–70% in severe ID cases (de Ligt et al, 2012; Gilissen et al., 2014; Rauch et al., 2012; Vissers et al., 2010; Worthey et al., 2011; Yang et al., 2013). Whole-exome and whole-genome sequencing has been used to sequence patient–parent trios, successfully identifying causal variants in patients with extreme

genetic and phenotypic heterogeneity (Fitzgerald et al., 2015; Gilissen et al., 2014; McRae et al., 2017; Rauch et al., 2012; Wright et al., 2015; Vissers et al., 2010).

Due to the genome-wide approach, NGS technologies can now be applied as a diagnostic tool in the absence of clinical phenotyping. Therefore these technologies are well suited for discovery of causative single nucleotide *de novo* variants in individual patients where parental DNA is available for comparison, accelerating the discovery of rare as well as hypomorphic and less-penetrant variants. Moreover, with the reduction in sequencing time and cost, NGS has now been successfully implemented in a clinical diagnostic setting (Bick et al., 2017; de Ligt et al., 2012; Monroe et al., 2016; Rauch et al., 2012; Worthey et al., 2011).

Overall, there has been a drastic increase in gene discovery in ID-associated disorders over time (Fig.1.1). Around 700 genes have now been linked to disorders where ID is the major or only feature, with a total number of identified ID genes predicted to exceed 1,000 in the next decade (Lubs et al., 2012; Vissers et al., 2016).

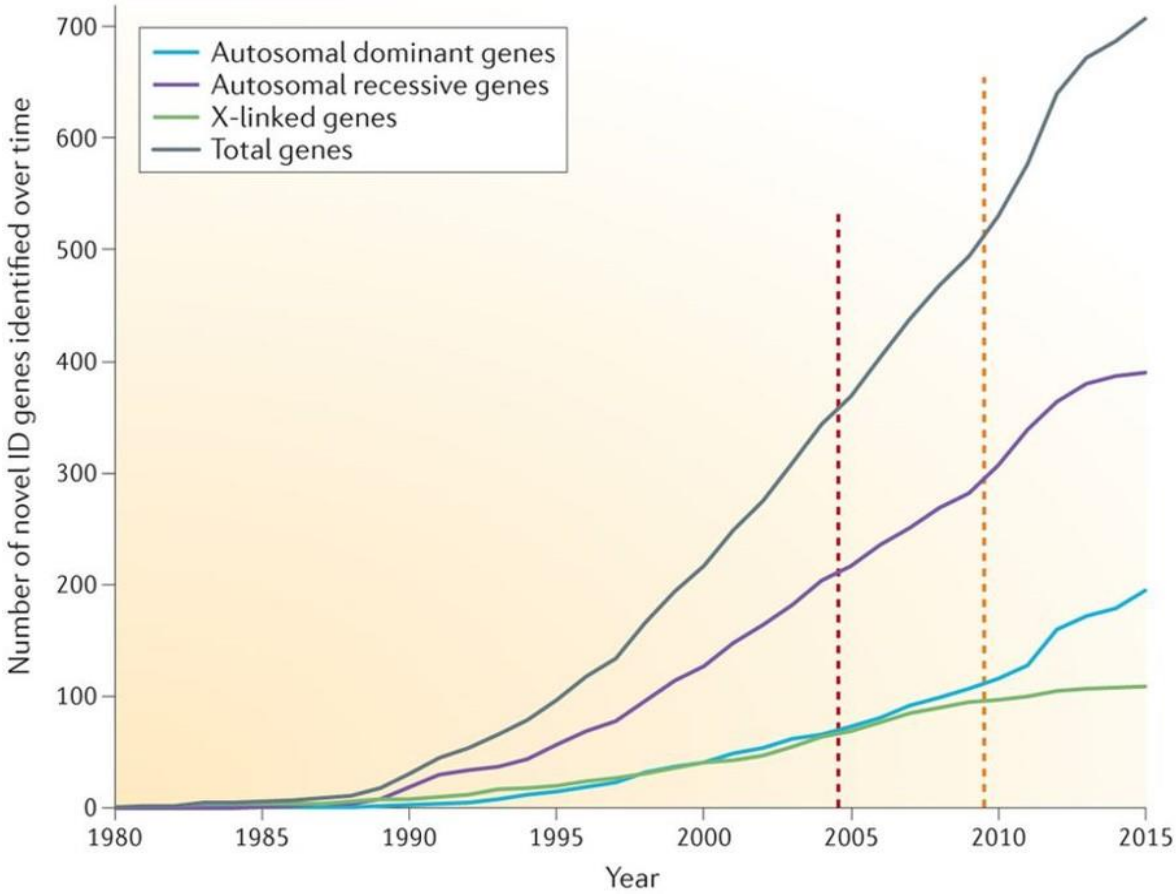


Figure 1.1 Overview of the gene discovery for intellectual disability over time, separated by types of inheritance. Red dashed line represents the introduction of genomic microarrays, while the orange dashed line represents the introduction of next-generation sequencing technologies. Figure reproduced from Vissers et al., 2016.

1.6 Brain ontology

The development of the human brain requires a very intricate and tightly regulated set of processes, with a plethora of cells proliferating, differentiating, migrating and integrating into cohesive circuitry, giving rise to a complex structure with around 85 billion neurons (Azevedo et al., 2009). The human brain total volume reaches 1,700mL in adulthood and is composed of 80% parenchyma, which is predominantly neurons and glial cells, 10% blood, and 10% cerebrospinal fluid (Williams et al., 2008). Glial cells, including oligodendrocytes, astrocytes, ependymal cells, and microglia, are 10-15 times more abundant in the brain than neurons (Williams and Herrup, 1988). It is therefore not surprising that the complex

processes of brain development are highly sensitive to errors and that neurodevelopmental disorders have been identified that are associated with each stage of brain development (Walsh and Engle, 2010).

An increased focus on the genetics of neurodevelopment disorders has facilitated an understanding of the underlying genes and pathways critical for normal brain development (Ehninger et al., 2008; Hu et al., 2014). The consequences of aberrant development processes have a wide range of associated phenotypes. For example, aberrant neuronal migration is associated with lissencephaly (a brain malformation with a thickening of the cortex), defects in neuronal progenitor proliferation is associated with microcephaly (significant reduction in head circumference), and aberrant connectivity in the brain has been linked to autism (Courchesne, 1997; Forman et al., 2005; Yang et al., 2012). Moreover, identical mutations in the same gene can cause different phenotypes due to the hypomorphic or less penetrant nature of the variant, as is the case with mutation in *POMGNT1,* which causes muscle-eye-brain disease, but of variable phenotypic severity even in related individuals (Diesen et al., 2004; Teber et al., 2008).

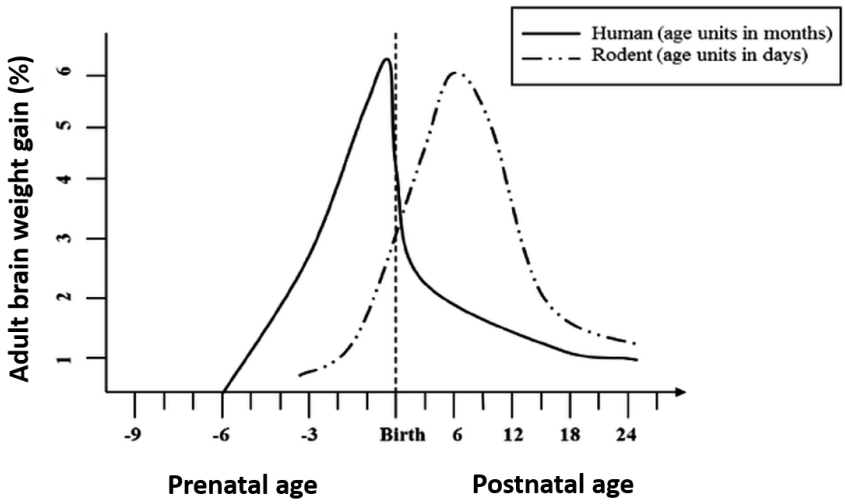
The milestones of brain development have been shown to be conserved between humans and rodents, but the processes occur along different timelines: humans have a longer period of brain development, which is thought to be associated with the development of a larger cortex and a longer postnatal period for fine-tuning and shaping of the brain circuitry (Clancy et al., 2000; Clancy et al., 2007; Rice and Barone, 2000; Semple et al., 2013) (Fig.1.2). The association between specific behaviours and the development of brain structures and circuits are comparable between humans and rodents (Rice and Barone, 2000). Because many disorders arise due to defects in the ontogeny of developmental processes and brain structures, it is critical when modelling the human developmental conditions in mice to be cognizant of differences and similarities in the brain ontogeny between the two species.

The peak of brain growth and gliogenesis occurs at 36-40 weeks of gestation in humans, around the time of birth, and in postnatal days P7-10 in rodents (Bockhorst et al., 2008; Catalani et al., 2002; Kriegstein and Alvarez-Buylla, 2009) (Fig.1.2). Cortical neurogenesis, the process by which new cortex neurons are made, starts during gestation in both rodents

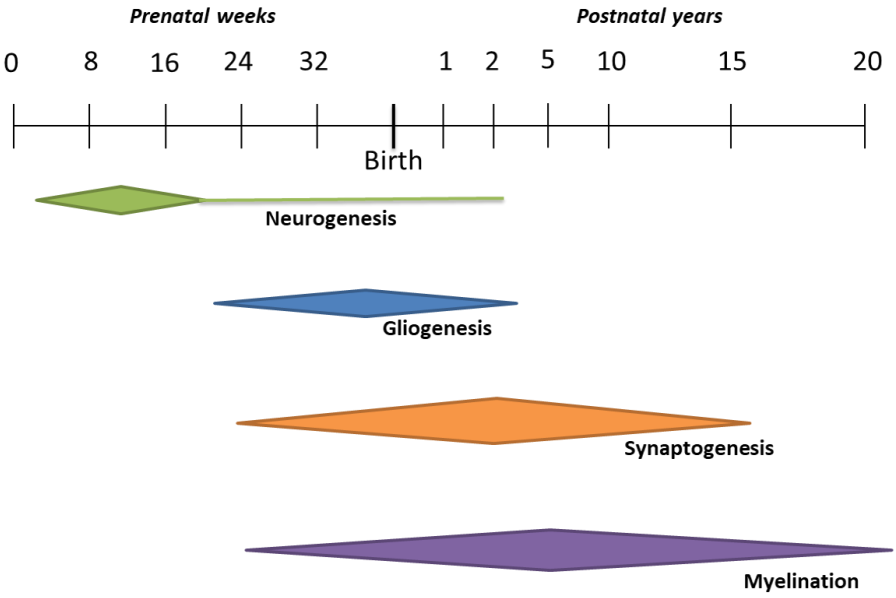
and humans, and can continue up to 2.5 years in humans (although mainly happens during gestation) and up to postnatal day (P) 15 in rodents (Babikian et al., 2010; Prins and Hovda, 2003). Neurogenesis in the hippocampus in both species continues into adulthood, but at a much lower rate than during development (Hill et al., 2015; Kitamura and Inokuchi, 2014). The brain reaches 90-95% of adult weight by 2-3 years old in children and P20-21 in rodents, which is also the time for the peak of myelination (Keshavan et al., 2002). Grey matter, consisting of neuronal cell bodies and dendrites and glial cells, initially increase from birth and then begins to reduce, with synaptic density reaching a plateau at 12-18 years in humans and P34-49 in rodents (Huttenlocher, 1979; Lidow et al., 1991). At this stage, the activity-dependent circuitry is being refined. Of note, the timing of changes in grey matter volumes tends to be region-specific. For example, grey matter in the frontal lobe reaches maximum volume in humans at 11-12 years of age, while the temporal lobe reaches maximum size later, at 16-17 years of age (Bansal et al., 2008; Sowell et al., 1999). White matter, composed primarily of myelinated axons that connect various grey matter regions, follows a different growth trajectory to grey matter, with ongoing myelination and increase in white matter volume happening beyond 20 years of age in humans and P60 in mice (Lebel and Beaulieu, 2011).

Due to the cross-species alignment in key brain developmental milestones, with comparable brain growth trajectories, rodents have been used extensively as model organisms for understanding these developmental processes. In Chapter 2, I will describe the use of rodents in the modelling of intellectual disability.

A


Human (age units in months)
Rodent (age units in days)
Adult brain weight gain (%)
1
2
3
4
5
6
-9
-6
-3
Birth
6
12
18
24
Prenatal age
Postnatal age

B


Prenatal weeks
Postnatal years
0
8
16
24
32
1
2
5
10
15
20
Birth
Neurogenesis
Gliogenesis
Synaptogenesis
Myelination

C

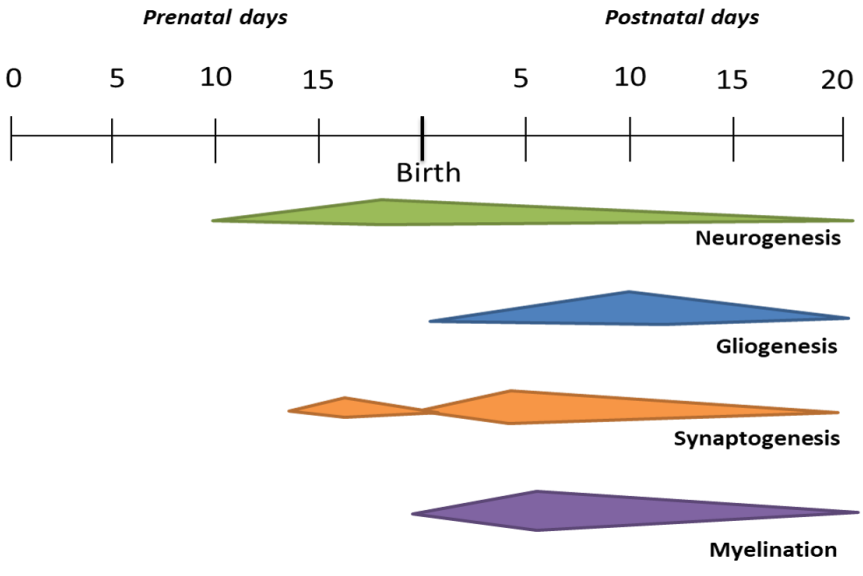

Prenatal days
Postnatal days
0
5
10
15
5
10
15
20
Birth
Neurogenesis
Gliogenesis
Synaptogenesis
Myelination

Figure 1.2 Brain ontology in humans and mice, including developmental milestones and growth of the brain over time. **A.** Brain growth in humans (solid line) and mice (dotted line), represented by the percentage of adult brain weight gain over time (in month units for humans and day units for mice). The peak brain growth spurt for humans is around the time of birth (36-40 prenatal weeks), while for mice the peak is around postnatal day 7-10. **B.** Human brain ontology over time with key conserved developmental processes: neurogenesis (green) occurs predominantly during prenatal stages but may continue up to 2.5 years (postnatal), gliogenesis (blue) peaks around 36-40 prenatal weeks, synaptogenesis (orange) peaks around 2 years of age, myelination (purple) is an on-going process that happens beyond 20 years of age. **C.** Equivalent developmental processes in the mouse: neurogenesis (green) starts around prenatal day 9.5 and the majority of neurogenesis is completed by postnatal day 15, gliogenesis (blue) peaks around postnatal days 7-10, critical period of synaptogenesis (orange) occurs during the first three postnatal weeks, peaking during second postnatal week, myelination (purple) is an on-going process that happens beyond postnatal day 20. Of note, the timings outlined above are not exact, as they vary between different brain regions. Figure adapted from Klintsova et al., 2013 and Semple et al., 2013.

1.7 PhD objective and aims

The primary objective of my PhD was to model a subset of loss-of-function mutations in mice, associated with novel neurodevelopmental disorders, identified in patients from large collaborative genetic projects, Deciphering Developmental Disorders (DDD) and Windows of Hope (WOH). This was done in order to firstly, demonstrate a causal link between mutation and phenotype, and secondly, to further understand the mechanisms by which these mutations result in human neurodevelopmental disorders. The mutations I modelled include homozygous and compound heterozygous recessive mutations (WOH dataset) and *de novo* dominant heterozygous mutations (DDD dataset).

I adopted a multi-phase approach to study these novel developmental disorders associated with ID. To successfully model human disorders in mice, it is necessary to develop a robust approach to testing cognitive and behavioural deficits in the mouse models. I therefore first designed a cognitive and behavioural phenotyping strategy for modelling DDs and ID in mice (described in Chapter 2). I then employed these behavioural and cognitive strategies, as well as morphometric and molecular paradigms, to study one such novel disorder, *KPTN*-related syndrome, in-depth (described in Chapter 3). After that, I determined whether such an approach could be scaled up to study multiple novel neurodevelopmental disorders. For this, I employed the most robust techniques from the *Kptn* work to design a behavioural and cognitive screening paradigm and employed this to test four further mouse models, each with a loss-of-function mutation in a candidate DDD gene (described in Chapter 5). Finally, in the concluding discussion of my findings, in Chapter 5, I outline potential future avenues of research opened up by this work and the broader implications of the state of the field of mouse modelling

Chapter 2. Materials and methods

2.1 Mutant alleles

All the mouse models were generated at the Wellcome Trust Sanger Institute. The *tm1a* and *tm1b* alleles contain a splice acceptor (SA), IRES:LacZ:Puro promoter-driven reporter cassette, followed by a polyadenylation site (pA) inserted upstream of the critical exon. The 'knockout-first' allele (*tm1a*) generated by insertion of an IRES:*lacZ* trapping cassette and a floxed promoter-driven *neo* cassette into an intron of the gene of interest, disrupting gene function at the mRNA level by interfering with transcription downstream of the cassette site (Skarnes et al., 2011; Testa et al., 2004; White *et al.,* 2013). *Tm1a* also has the potential to be converted to a conditional allele (*tm1c*) restoring gene activity (Skarnes et al., 2011). *Tm1b* allele was generated by flanking of the critical exon by LoxP sites. In the presence of Cre recombinase, the LoxP sites recombine leading to the excision of the critical exon. Deletion of this critical exon results in a frameshift and predicted absence of the protein product. *Zmynd11*$^{em1(IMPC)Wtsi}$ model harbouring frame shifting exon-deletion null allele was generated using the CRISPR/Cas9 system (Fig.2.1). Mouse models were kept on C57BL/6NTac background, apart from *Kptn*$^{tm1a(EUCOMM)Wtsi}$ which was on a mixed C57BL/6B background (C57BL/6Brd-Tyr<c-Brd>;C57BL/6Dnk;C57BL/6N;C57BL/6NTac).

Zmynd11$^{em1(IMPC)Wtsi}$ deletion allele flanking sequence of:

AAAAAAGGTCAAAAGAATGCTTTCCCCACACAGGGCACTGGCCATCACCTCTGTAAGCCACACCAGG
AAGCAAGCATCAGTTTAAATTCCAACATCATTATGGCCAGTGTGTCTACTTCCCACAGTGGACTGCAC
AGCTGCCATGCCTCAACTGCAGCTGTGGGGGCAGCTGCTGCTGAGCTTTGCCAATAGGAAATGAT

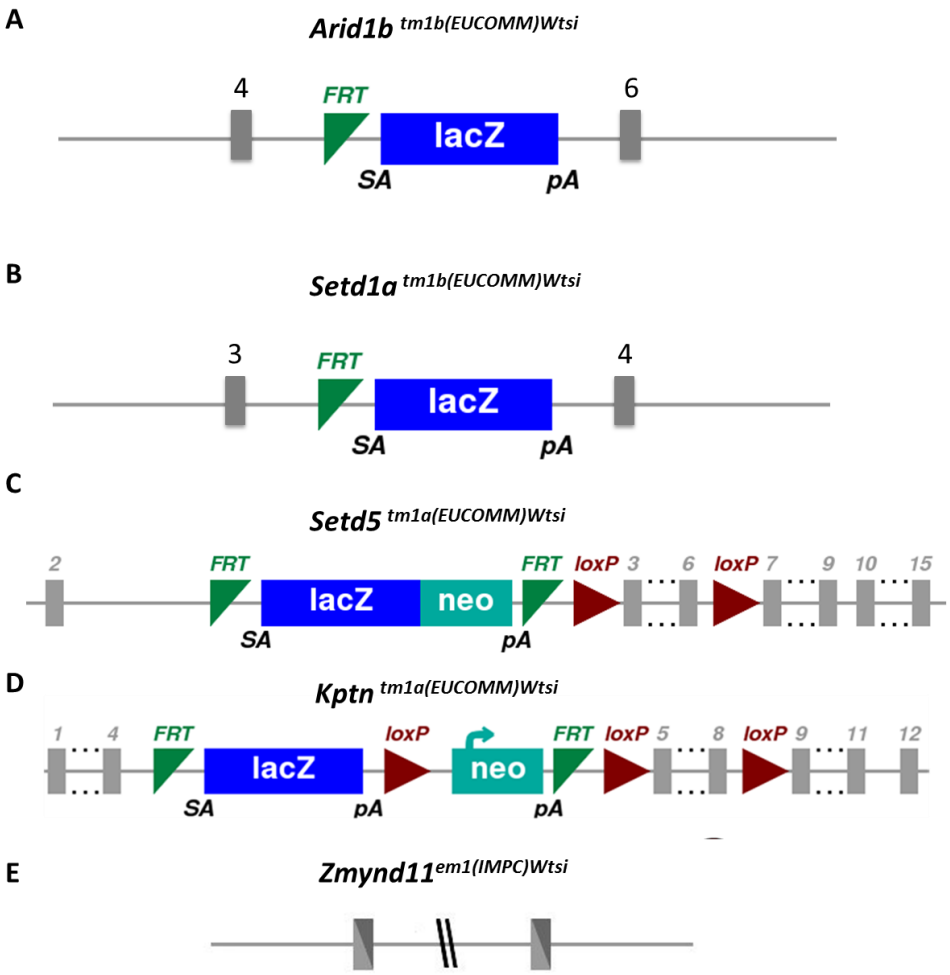


Figure 2.1 Schematic representation of the mutant alleles. (A) *Arid1b*$^{tm1b(EUCOMM)Wtsi}$ and (B) *Setd1a*$^{tm1b(EUCOMM)Wtsi}$ are *tm1b* alleles; (C) *Setd5* $^{tm1a(EUCOMM)Wtsi}$ and (D) *Kptn*$^{tm1a(EUCOMM)Wtsi}$ are *tm1a*; (E) *Zmynd11*$^{em1(IMPC)Wtsi}$ is CRISPR/Cas9 frame shifting exon-deletion

2.2 Animal husbandry

Animal husbandry was greatly facilitated by the collaboration of the WTSI Research Support Facility staff. The colony managers at the Sanger Institute provided support on colony maintenance. Housing and breeding of mice and experimental procedures were carried out under the authority of UK Home Office project and personal licenses. All procedures were carried out in accordance with the Animal Welfare and Ethical Review Body of the Wellcome

Trust Sanger Institute and the UK Home Office Animals (Scientific Procedures) Act of 1986, under UK Home Office PPL 80/2472.

Standard rodent chow and tap water were available *ad libitum*, with the exception of the touchscreen experiment that required mild food restriction. All test animals were housed 2-5 mice per cage in mixed-genotype cages, except where stated otherwise and were maintained on a 12:12 light-dark cycle with lights on at 7:00 AM, and at ~20°C and 55% humidity. Testing was done during light-hours.

2.3 Mouse behavioural paradigms

In all assays, the mice were tracked by detection of the mouse's centre and nose point using overhead infrared video cameras and a user-independent automated video tracking software Ethovision XT 8.5 (Noldus Information Technology, Wageningen, The Netherlands).

All experiments were carried out on mice aged 10-22 weeks. The mice were handled for 2-3 days before the onset of testing. When assays were piloted, wildtype mice of C57BL/6NTac genetic background were tested, unless otherwise stated. All mice were habituated to the behavioural room in their home cages for ≥ 30 minutes under the same light condition as the test. Males were tested unless stated otherwise.

Determining how many animals should be used in a study is an important factor in experimental study design. Sample size calculations reduce the probability of not detecting a statistically significant difference between groups even if it truly exists, while minimising the unnecessary use of animals (Chow et al., 2008; Kadam and Bhalerao, 2010). The factors that influence the estimated sample size are the power, the significance level, the effect size of scientific interest and inter-individual variability (Chow et al., 2008; Kadam and Bhalerao, 2010). The power, which is the probability of finding an effect the study aimed to find, was kept at 80% in all calculations. The significance level, the probability that the observed differences are likely to be due to chance, was always set at type I error of 5% (P = 0.05). The mean and the standard deviation (the variation within a group) were taken from the wildtype pilot experiments wherever possible or representative wildtype data from Chapter

5 when there was no pilot data. Estimated sample size was calculated by either comparing the mean (μ) to a reference value that represented 'no preference' and thus 'no learning'; or comparing two means from two groups (μ and μ_B), where the effect size, which is the minimum difference between two groups that is clinically significant, was decided based on previously published studies. Two-tailed Student's t-test was used. Power analysis was done using online software (Chow et al., 2008).

I received help in doing some of the assays by members of Team 29 and 156 at the Sanger Institute. Genotyping was performed by a specialist team (led by Dr. Ed Ryder), Mouse Genetics Programme, using a standard qPCR

2.3.1 Open field

Mice have an innate tendency to explore novel environments (Seibenhener and Wooten, 2015). When placed in an open field, measurements such as distance covered, time spent moving, and velocity travelled can be used to assess the overall activity and locomotion of the mice (Seibenhener and Wooten, 2015). A period of movement was defined when the mouse reached a velocity of 2 cm/s over two frames; a period of non-movement was defined when velocity was lower than 1.75 cm/s over two frames. Open field can also be used to test anxiety-like behaviour, by exploiting the conflicting tendencies of mice to spontaneously explore a novel environment and to avoid brightly lit open spaces (Kulesskaya and Voikar, 2014).

In all open field trials, mice were placed in the left corner of the open field arena. Two animals were tested in parallel in two separate open fields (74 cm x 74 cm). A *centre zone* was designated with equidistant borders to the open field walls (8 cm) (Fig.2.2). The arenas were cleaned thoroughly with ethanol free wipes to remove odour traces.

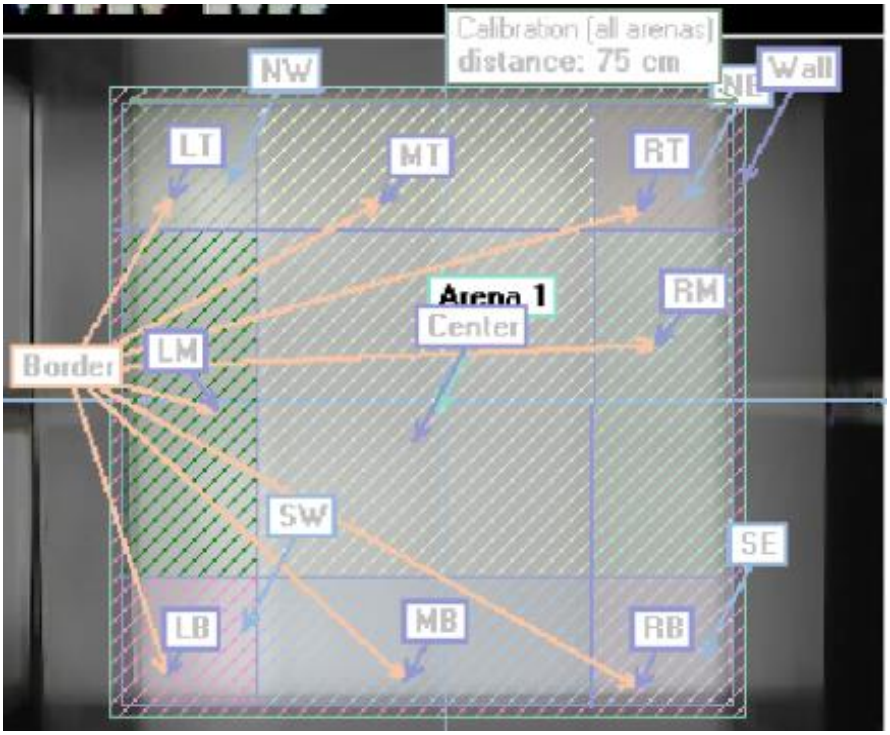


Figure 2.2 Open field arena with the centre and border areas outlined. The arena was divided using Ethovision tracking software into north west (NW), north east (NE), south west (SW), and south east (SE) and further subdivided into centre vs. border areas (comprising of the following quadrants - left top (LT), middle top (MT), right top (RT), left middle (LM), right middle (RM), left bottom (LB), middle bottom (MB), right bottom (RB)). The time spent by a mouse in these areas was calculated by the automated tracking system and the border quadrants are summed to get the final value for time spent in the borders. The arena size was calibrated in Ethovision to match the dimensions of the open field (75cm x 75cm).

2.3.1.1 Mice

In Chapter 4, I tested the mice in the open field under dim (1-3lux), in order to capture the general activity and locomotion of mice in a less stressful environment (dim light). The duration of each trial was 10 minutes. n=12 male *Kptn*$^{-/-}$ and n=13 wildtype littermate controls were tested in one day, within light-hours. In Chapter 5, the mice were tested in an open field under bright-light (200-300lux) for 15 minutes and the time spent at the borders of the arena close to the wall was also analysed in relation to the amount of time spent in the centre, as a measure of anxiety. Four mutant lines were tested in this manner - *Setd1a*

(wildtype: n=15, mutant: n=15), *Arid1b* (wildtype: n=12, mutant: n=11), *Setd5* (wildtype: n=24, mutant: n=12), *Zmynd11* (wildtype: n=14, mutant: n=12).

2.3.2 Light/dark box

Light/dark box assay was used to test for anxiety-related behaviours. The light/dark assay exploit the conflicting tendencies of mice to spontaneously explore a novel environment and to avoid brightly lit open spaces (Hascoet and Bourin, 2003; Crawley and Goodwin, 1980). The light/dark box comprises of a brightly-lit *light zone* (two-thirds of the total area) and a *dark zone* (one-third of the total area).

The assay was conducted in the dark. After 30 minutes of habituation to the room in their home cages, mice were separated into new cages (with clean bedding, food pellets, fun tunnel, and some dirty bedding to reduce fighting between the males), two mice per cage. Light/dark box was placed into an open field arena, positioned with the help of acrylic lid and empty cage to ensure the light/dark box was in the same location in the open field arena between experiments (Fig.2.3). The *light zone* had two spotlights facing down directly into the compartment, illuminating the whole of the 'light' compartment as evenly as possible (lux 300-400). The *dark zone* was sheltered from the light by a top lid and a door at the opening between the two compartments. The mouse was placed in the centre of the dark compartment. The door was released at the start of the trial. Trial duration was 10 minutes, and two mice were tested in parallel. The time spent in each of the compartments, as well as the frequency of the transition between zones, were recorded. Preference score was calculated as the time spent in one of the compartments out of the total time of the assay.

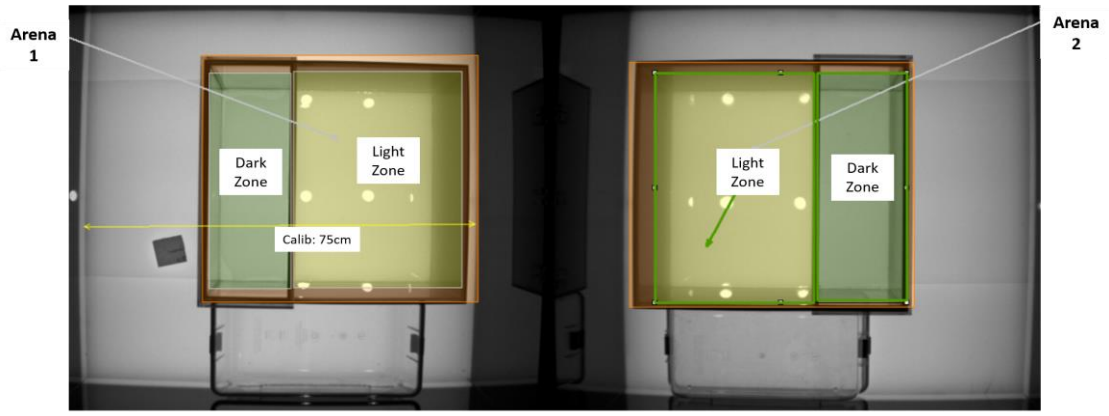


Figure 2.3 Light/dark box set-up. Image of two open fields with light/dark boxes inside. Each light/dark arenas contains a dark zone and light zone. The arena size was calibrated in Ethovision to match the dimensions of the open field (75cm x 75cm).

2.3.2.1 Mice

Light/dark assay exploits the tendency of mice to prefer the dark zone over the light zone, with mice experiencing higher levels of anxiety exploring the light zone significantly less, spending longer in the dark zone and making fewer transitions into the light zone (Blundell et al., 2009; Shum et al., 2005). To calculate the estimated sample size needed to detect a difference in mean time spent in the dark zone, the following was used: mean time spent in the dark μ = 67.96 (%; wildtype data from Chapter 5), SD= 15.77, effect size = 18-24 (%; chosen based on published studies (Blundell et al., 2009; Shum et al., 2005)). The estimated sample size was n=7-12 (accounting for an effect size of 18% or 25%: with lower n calculated from higher effect size).

Mice that did not transition into the light zone at all during the duration of the assay thus not exploring both areas, were excluded due to the possibility that the preference is due to lack of exploration.

In Chapter 4, n=14 *Kptn*$^{-/-}$ mice and n=13 wildtype littermate controls were tested, and one mutant was excluded from the final analysis because it did not transition to the light zone

and remained the dark zone the whole time. In Chapter 5, only two out of the four mouse lines were tested: *Setd1a* mice (n=14 wildtypes and n=11 mutants; none were excluded from the final analysis) and *Zmynd11* mice (n=14 wildtypes and n=12 mutants; n=2 WT and n=1 mutant were excluded due to 0 transitions to light zone).

2.3.3 Object discrimination

Novel object recognition (NOR) assay relies on rodent's innate preference for investigating the novel over the familiar object (Ennaceur and Delacour, 1988). Longer investigation of the unfamiliar object by the animal is indicative of an acquired memory for the familiar object. The NOR pilots (described in Chapter 3) were done under a modified protocol from Ennaceur and Delacour (1988) and Tuscher et al. (2015). A modified version of the NOR, object displacement (OD), is a spatial memory task that exploits the natural tendency of rodents to explore their environment and show a preference to investigate a moved object when compared to a static object (Anderson and O'Mara, 2004; Ricceri et al., 2000; Tuscher et al., 2015). OD pilot described in Chapter 3 was based on a protocol from personal correspondence with Dr Lukas von Ziegler at the Laboratory of Neuroepigenetics, Brain Research Institute, University of Zurich.

In order to test mice for spatial and non–spatial memory in a time-efficient manner, I designed a three-day paradigm combining OD and NOR assays (Fig.2.4B). In this paradigm, the OD assay was used as both an independent test for spatial memory and as the training phase for NOR (Fig.2.4Bi) (Fernandez and Garner, 2007; Murai et al., 2006).

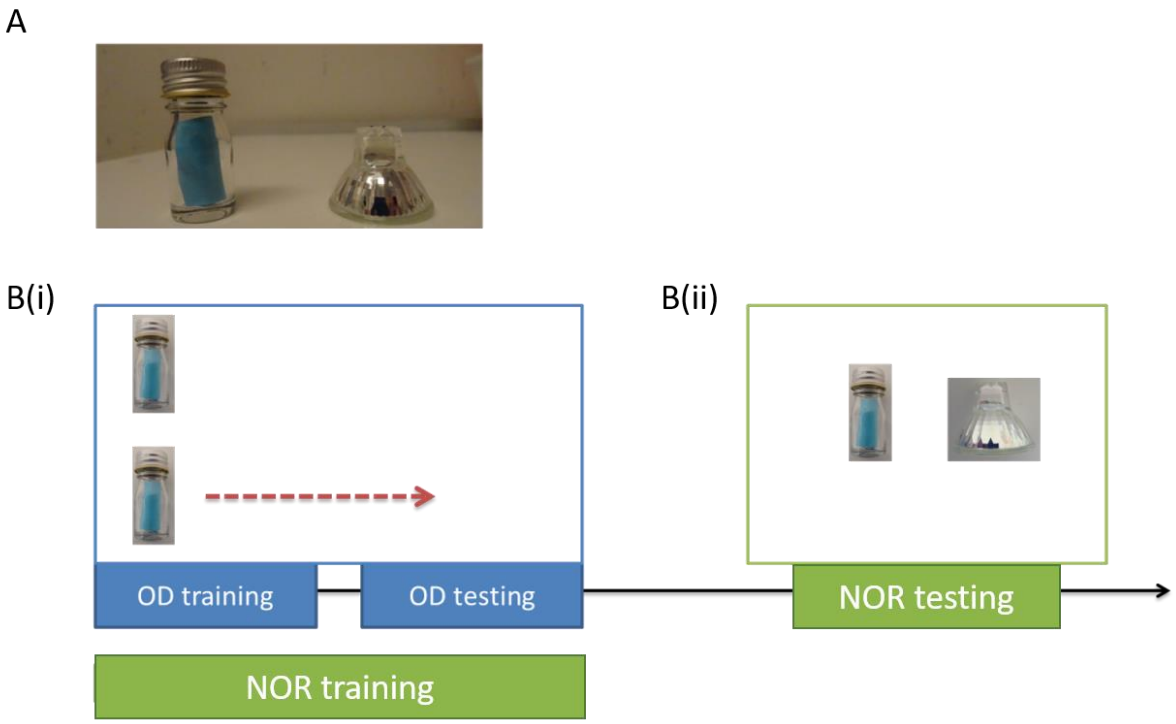


Figure 2.4 Object discrimination paradigm set-up. **A**. An image of the objects used for the paradigm – vial (left on the image) and light bulb (right on the image). **B.** The object discrimination schematics: **(i)** Object displacement (OD) acts as both a spatial test, with training phase (two identical objects placed in the north and south side of the open field) and testing phase (one of the object is moved from north west to north east), and as a novel object recognition (NOR) training, followed by **(ii)** NOR testing 24 hours later, with two objects (vial and light bulb) displayed at the same time on west and east side of the open field.

The mice were tested under dimmed lights (40 lux). Each open field was divided into four smaller arenas. A spatial queue, horizontal stripes (white tape) on the central wall (running north to south), was provided in each arena. Eight mice were tested in parallel. Each trial lasted 10 min. During Day 2 and Day 3, the time the mouse spent investigating the objects was recorded, by tracking the nose point (rather than centre point) as it came into contact with the zone around the objects.

2.3.3.1 Day 1 (habituation to the arena)

Mice were randomly assigned an arena, put into the centre of the arenas facing the striped wall, and allowed to habituate to the empty arena for 10 minutes. At the end of the 10 minutes, mice were put back into their home cages and returned to their holding room.

2.3.3.2 Day 2 (OD /NOR training)

2.3.3.2.1 OD training

Mice were placed into the same arenas as Day 1 and presented with either two identical small vials or two identical small light bulbs on the west side (NW and SW quadrants of the arena) (Fig.2.4A,Bi). The objects were counterbalanced within each run and between genotypes. After 10 minutes, mice were placed back in their home cage and left in the experimental room undisturbed for 1-2 hours.

2.3.3.2.2 OD testing

Mice were placed back to the same arenas containing the same object set as in training, but SW object was moved to SE (Fig.2.4Bi). At the end of the trial, the mice were put back into their home cages and returned to their holding room for 24 hours.

2.3.3.2.3 Day 3 (novel object testing phase)

After 24 hours the mice were brought back to the testing room, and after 30 min of habituation (in their cages) to the room, were placed in their previously allocated arenas. The arenas contained a vial and a light bulb, placed near the middle of the west and east walls of the arena (Fig.2.4Bii). For each object set, one object was the familiar one from Day 2, and the other was unfamiliar. The objects were counterbalanced for novelty (based on Day 2).

2.3.3.3 Mice

Object displacement paradigm exploits the natural tendency of rodent to prefer novelty in their environment, by investigating moved or novel objects more (Anderson and O'Mara, 2004; Ennaceur and Delacour, 1988; Ricceri et al., 2000; Tuscher et al., 2015). In order to calculate the estimated sample size needed to detect a preference for a moved object, the following was used: mean μ= 61.01 (% preference of investigating the moved object rather than the stationary object (wildtype pilot; Chapter 3)), SD=13.53, reference value = 50 (%; 'no preference' for either of the objects). The calculated sample size was n=12.

Mice that did not meet the pre-determined exclusion criterion of <10 seconds total investigation of objects in a 10-minute assay were removed from the analysis and subsequent phases of the assay (Arqué et al., 2008; Leger et al., 2013).

For the Kptn cohort (Chapter 4), out of initial n=12 for both genotypes, one wildtype and one *Kptn*$^{-/-}$ were excluded from the Day 1 (OD acquisition phase), four wildtypes and two mutants were excluded from the Day 2 (OD testing phase) due to their low overall investigation (<10sec), therefore the total number of mice in the OD analysis was wildtype n=7 and *Kptn*$^{-/-}$ n=9. One wildtype was excluded from Day 3 (NOR testing) due to low overall investigation time (total NOR analysis: wildtype n=6, *Kptn*$^{-/-}$ n=9).

2.3.4 Social recognition

Social recognition assay exploits the innate preference of mice to investigate novel conspecifics and tests olfactory-mediated memory (Ferguson et al., 2001; Guan et al., 1993; Kogan et al., 2000; Winslow and Insel, 2004). It has previously been shown that performance in this assay is dependent on neuroanatomical structures such as the olfactory bulb, amygdala, and hippocampus (Ferguson et al., 2001; Guan et al., 1993; Kogan et al., 2000; Winslow and Insel, 2004). I used the social recognition protocol used in the lab previously (Dias et al., 2016; Sanchez-Andrade et al., 2005). The assay comprises two days: Day 1 is memory acquisition and Day 2 is 24 h memory retention test (Fig.2.5).

Testing was conducted under red light. All trials were recorded (but not tracked) by Ethovision and manually scored. Testing arenas comprised of six empty cages (removed of bedding and enrichment materials), placed in the open field and separated with dividers so that the test animals could not see each other. Twelve animals were tested in parallel in two open fields and twelve arenas in total. Test animals were habituated to the room (and red light) for half an hour in their home cages and were then habituated to the arenas for 10 minutes before the start of the testing.

Due to technical difficulties, females from the *Kptn* cohort were tested in Chapter 4. In Chapter 5, only males were tested for all the mouse lines under the investigation.

2.3.4.1 Stimuli

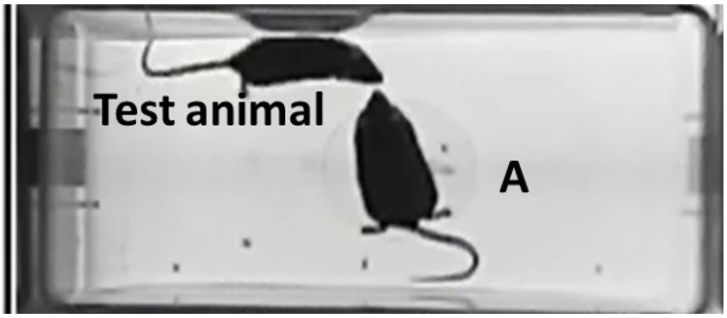
Mice used as stimuli were matched to test animals according to their weight, size, and gender, and will be henceforth referred to as *stimuli*. Stimuli for Day 1 trial 1-4 were of different genetic background to those used for trial 5. For trial 5, in Chapter 4 (*Kptn* cohort), C57BL/6NTac stimuli were used. For all other lines (Chapter 5) stimuli with 129 background were used in trial 5.

To ensure familiar and unfamiliar stimuli pair for each test animals could be distinguished cage mates were never used. For *Kptn* cohort familiar-unfamiliar stimuli pair were from different colonies, but of the same genetic background (129P2/OlaHsdWtsi + C57BL/6J). For

all other lines (except *Arid1b*) C57BL/6J and C57BL/6J(50.0%)CBA/Ca(50.0%) mice were used as stimuli. The stimuli were counterbalanced for novelty.

For each trial, the same stimulus was used for 3 test animals to reduce the number of animals needing to be anaesthetised. Stimulus animals were subject to non-terminal anaesthesia with ketamine/xylazine (intraperitoneal injection 1 g/0.1 g per kg of body weight) and were recovered with atipamezole (subcutaneous injection 0.5 mg) in a home cage with a warm water bottle.

A



B

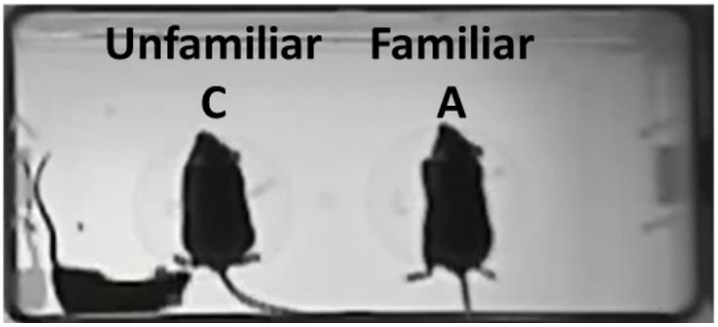


Figure 2.5 Images depicting test animal and stimuli on Day 1 and Day 2 of social recognition. A. Day 1 - the test animal investigates one stimulus (A) for four trials. **B**. Day 2 - the test animal is presented with a simultaneous choice of a familiar stimulus (A) and a novel unfamiliar stimulus (C).

2.3.4.2 Day 1 (memory acquisition)

A conspecific anaesthetised mouse (stimulus A) was placed on top of a clean petri dish and placed in the centre of the test arena for 1 minute (Fig.2.5A). This was repeated four times, with inter-trial intervals of 10 minutes. On the fifth trial a novel stimulus (B) was presented instead of stimulus A. Time investigating the stimulus (when the test mouse actively sniffed and interacted with the stimulus) was recorded manually, blind to genotype. Mice were then returned to their home cages and returned to their holding room for 24 hours.

2.3.4.3 Day 2 (memory retention test)

After a 24 hours, the test mice were put back in their allocated testing arenas and re-exposed to the familiar stimulus animal (A) used on Day 1 (Trial 1-4) trials at the same time as an unfamiliar stimulus animal (C). The stimuli were placed on Petri dishes and placed on opposite sides of the arena for 2 minutes. The time the test animal spent with stimuli was recorded, and the difference in investigation time between familiar and unfamiliar stimuli was compared per genotype using multiple t-tests with multiple comparison corrections.

2.3.4.4 Mice

Social recognition assay exploits the natural tendency of mice to investigate an unfamiliar vs a familiar conspecific (stimuli) (Ferguson et al., 2001; Guan et al., 1993; Kogan et al., 2000; Winslow and Insel, 2004). To calculate the estimated sample size needed to detect a preference for the unfamiliar stimuli, the following was used: mean μ=63.48 (% preference for unfamiliar; wildtype pilot data not shown), SD= 14.15, reference value = 50 (%; 'no preference' for either stimulus). The estimated sample size was n=9.

Mice were excluded if their overall investigation time was <10 seconds on trial 1 or if they did not investigate one of the stimuli during Day 2. The mice used in Chapter 4 (*Kptn* cohort; female only): started with n=12 per genotype. One wildtype and four mutants were excluded due to the low investigation on trial 1 on Day 1 (final: wildtype n=11, *Kptn*$^{-/-}$ n=8).

The mice used in Chapter 5 (male only): *Setd1a* cohort - wildtype n=9 and *Setd1a*$^{+/-}$ n=10; *Setd5* cohort – on Day 1 wildtype n=15, *Setd5*$^{+/-}$ n=10, excluded 1 wildtype on Day 2; *Zmynd11* cohort – wildtype n=14, *Zmynd11*$^{+/-}$ n=10, excluded 1 mutant on Day 2; *Arid1b* cohort - wildtype n=6, *Arid1b*$^{+/-}$ n=7.

2.3.5 Sociability

Sociability is an olfactory-mediated social assay reliant on the natural affinity of wildtype mice to prefer interaction with a novel conspecific over a novel inanimate object (Moy et al., 2004; Silverman et al., 2010). The three chamber paradigm, which is widely used as a test for ASD-like phenotypes in rodent models, consists of one chamber with a holder containing a stimulus mouse, the central second chamber that is empty, and the third chamber with a holder containing an object (Moy et al., 2004; Silverman et al., 2010). The sociability protocol used in Chapter 5 has been previously established in our laboratory (Dias et al., 2016), and was modified here to account for the preference of mice for novel objects. Therefore, instead of habituating the mice to empty chambers first, I habituated the mice to chambers containing empty holders.

The test was conducted under dimmed lights (40lux) in an open field divided into four quadrants. One of the quadrants was sealed off, the remaining three chambers (L shape configuration) had openings through which mice could transition between chambers (Fig.2.6). The *mouse zone* was defined as the area around the holder containing a mouse stimulus, while the *object zone* was defined as the area around the holder containing an object (Fig.2.6). Awake stimuli (size and sex-matched) were of 129 background due to their reported relative hypo-activity predisposition when compared to some other inbred background strains (and care was taken to use novel mice since 129 mice were also used as stimuli in social recognition) (O'Leary et al., 2011). Stimuli could freely move inside the holder. The holes in the utensil holder allowed for the test animals to see and sniff the stimuli.

The time the test mouse spent in object and mouse zones, as well as in the mouse and object chambers, were recorded by tracking the nose and centre points respectively, using

Ethovision (Fig.2.6). The mice were initially placed in the central chamber and allowed to explore all the chambers freely for 10 minutes (habituation phase). The mice were then ushered back to the centre and the doors were closed. A stimulus mouse was placed in one utensil holder while an object, a small plastic bottle, was placed in the other utensil holder. The doors were then released and the test mouse was allowed to freely explore all chambers for another 10 minutes. Preference for mouse zone was calculated as a proportion of overall investigation time giving a score from 0-1, with 0.5 indicating no preference.

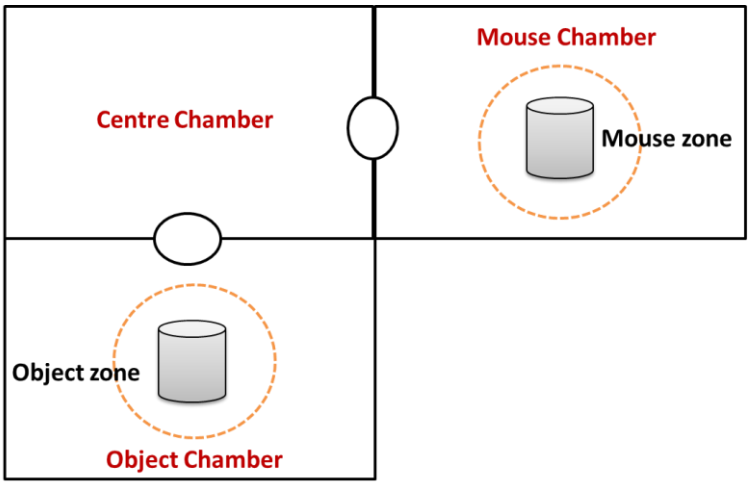


Figure 2.6 Schematic representation of the three-chamber sociability set up. The set up contains a centre, mouse, and object chambers with defined mouse and object zones around the holders containing a mouse and an object respectively.

2.3.5.1 Mice

Sociability assay is reliant on the natural affinity of wildtype mice to prefer a conspecific (stimuli) over an inanimate object, therefore spending longer in the mouse zone (Moy et al., 2004; Silverman et al., 2010). To calculate the estimated cohort size needed to detect preference for the mouse zone, the following was used: mean μ = 0.131 (preference score

for mouse zone; wildtype data from Chapter 5), SD= 0.1375, reference value =0.5 (chance), 5% type I error rate, 80% power. The calculated cohort size was n=9.

Mice were excluded if their overall investigation time was <10 seconds or if they did not investigate one of the stimuli at all. In Chapter 5 the following number of mice were used. *Arid1b* cohort: wildtype n=12, *Arid1b* mutant n=10; *Setd5* cohort: wildtype n=11, *Setd5* mutant n=10; *Setd1a* cohort: wildtype n=14, *Setd1a* mutant n=15.

2.3.6 Pairwise discrimination

Automated touchscreen technology is a platform for assessing cognitive function in rodents (Bussey *et al.,* 2012; Morton *et al*., 2006). Pairwise visual discrimination (PD) assay is one of the cognitive tasks that can be assessed using this platform. PD is an operant conditioning task that tests reward-based associative perceptual memory and learning (Bubser et al., 2014; Horner *et al*., 2013). The protocol used in Chapter 4 was adapted from previous studies and Campden instruction manual for PD task (for mouse touch screen systems and ABET II) (Brigman et al., 2007; Horner et al., 2013; Morton et al., 2006). The piloting of this assay is described in Chapter 3.

Because the assay is dependent on appetitive reward, the mice were food restricted to achieve a gradual reduction of 10-15% of the initial body weight, and this weight was maintained throughout training and testing (Table 1). The weight of the mice was measured daily, and the food was adjusted accordingly.

Table 1 Food restriction guidelines.

Weight loss (%)	Amount of food (g)
<10	1.5
10-12.0	2
12-13.0	2.5
13-15	3
15.0-17.9	3.5
>18	4
20	cull as per Home Office regulations

Before testing was carried out, the mice underwent several pre-testing procedural training phases, also known as *schedules*, where they learned how to select an image by nose-poking the screen, initiate a new trial and collect the reward (Horner et al., 2013) (Fig.2.7). Each phase had its own set of predefined criteria that dictated whether the mouse was ready to pass to the next phase of training (Fig.2.7). The training phase required the mice to learn the correct instrumental responses. Only once the mouse passed all these stages by satisfying the relevant criteria, could it progress to the PD task.

The training started with habituation of the mouse to the chambers (habituation 1) and learning how to collect the reward (habituation 2). The mouse was then presented with one image, appearing one at a time on one of the screens in a pseudo-randomised order ('initial touch' phase). After the image was displayed for 30 seconds, the mouse received a reward independent of whether it nose-poked the image or not. The mouse was encouraged to nose-poke the image (stimulus) on the screen by receiving three times more reward without the 30 seconds delay if it did. A new trial would start when the reward was collected. In the next training phase, the mouse had to nose-poking the image in order to receive a reward, as well as collect this reward to progress to the new trial ('must touch' training phase). In 'must initiate' training phase, the mouse had to do all the requirements of the previous phases, as well as initiate the start of that day's training session by receiving a free delivery of food which it had to collect from the reward tray, after which the first trial would start. For all the above training phases (from initial touch to must initiate) the mice had to complete 30 trials in 60 minutes to progress. The final training phase ensured the mouse had learned to associate *correct* nose-poking with a reward, by 'punishing' the *wrong* nose-pokes (poking the part of the screen without an image). If the mouse nose-poked wrongly, it received *a time-out* period of 5 seconds and the lights were inverted (all testing occurred in the dark). To complete the 'Punishment' training phase the mice had to complete 30 trials in 60 minutes and get 70% of the trials correct.

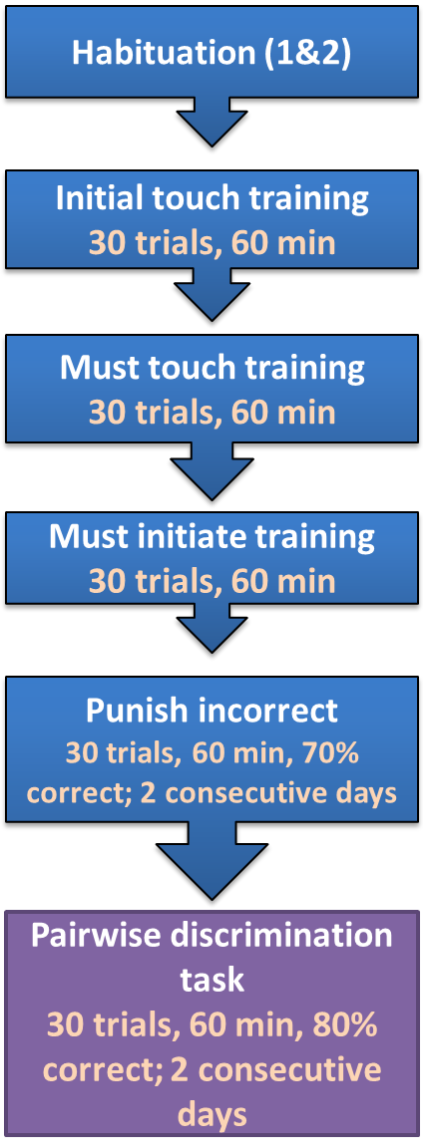


Figure 2.7 Pairwise discrimination flowchart. Blue boxes depict training phases (habituation, initial touch, must touch, must initiate, punish incorrect) and the purple box depicts pairwise discrimination testing phase. The criteria that needed to be achieved for the mice to progress to the next phase is shown in orange.

In the PD test phase, the *pairwise discrimination task*, the mice were presented with a choice of two novel simultaneously appearing images: the conditioned stimulus (CS+) associated with a reward and CS- that was not rewarded (Fig.2.8A). The mouse had to complete 30 trials in 60 minutes and get 80% of the trials correct (by nose-poking CS+) over

two consecutive days in order to complete the PD task. The performance in this task was assessed by the number of sessions it took for the mouse to meet the criteria. If the mouse nose-poked CS- image instead of CS+, it was punished with a time-out, followed by a *correction trial*, where the images kept appearing in the same position on the screens until CS+ was selected. The correction trials did not count toward the total session trial number or the final score (the percentage of trials selected correctly). The correction trials were designed to counteract potential side and stimulus biases, aid with learning, and ensure that mice got a certain number of rewards per session regardless of their performance in the non-correction trials.

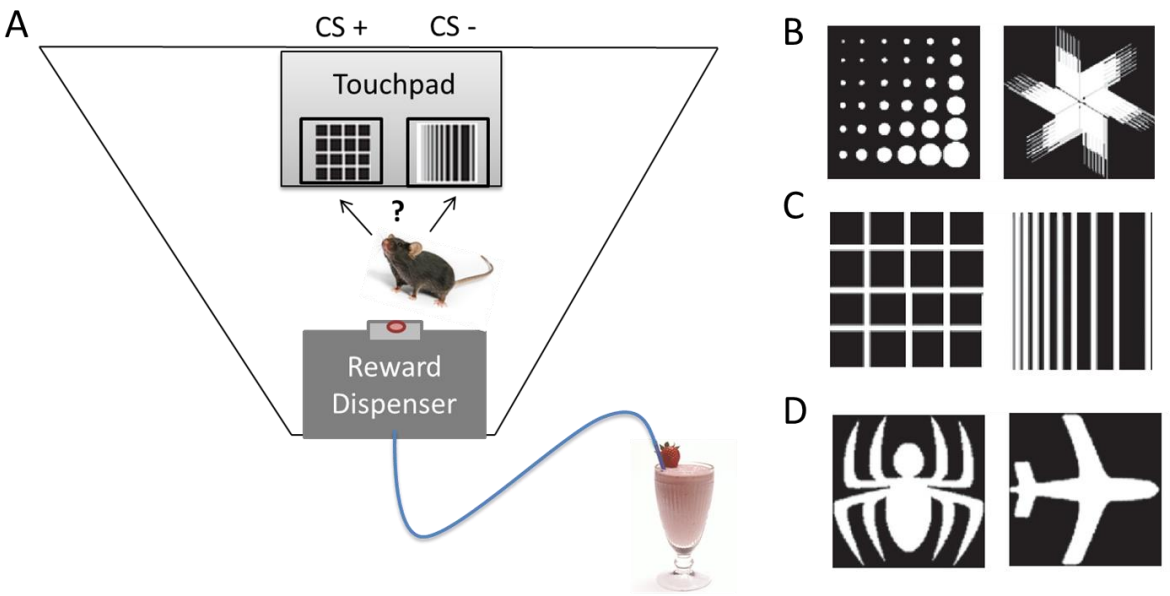


Figure 2.8 Pairwise discrimination set-up. A. A schematic representation of the pairwise discrimination task, in which the mouse is exposed to two novel images instantaneously and needs to pick the CS+ (the image that is pre-assigned to be linked to a reward) to get a strawberry milkshake reward. **B-D**. Image sets that have been used in the literature for pairwise discrimination. Image set B was discontinued due to reported image bias (correspondence with Campden Instruments Product Specialist). Image C was used in Pilot 1, while image D was used in Pilot 2 (Chapter 3, section 3.2.4) and Chapter 4 (section 4.2.3).

There were an additional set of pre-determined rules that were set during the piloting of this task (discussed in Chapter 3):

1. Because of the weekend break in testing, no new schedules were tested on Monday; all schedules from Friday were repeated. This was to take into account potential memory deficits in the mutant mice and have all animals re-baselined on Monday before continuing onto the next phase.

2. If however, the mouse did not pass criteria again on Monday – they repeated the schedule until they passed the criteria before proceeding to a new training phase.

3. In phases where the criteria had to be met two days in a row, if the mouse did Day 1 of criteria on the Friday but did not pass the criteria on Monday – they would need to pass the criteria two days in a row again to proceed.

2.3.6.1 Mice

All the mice were progressed through the assay on an individual basis. Mice were tested in the same testing box each time throughout all the sessions. There was a total of four boxes, allowing for four mice to be tested in parallel. Once each mouse reached the criteria in the PD task phase they were no longer tested. In order to represent the percentage of correct trials over sessions in PD task, for those mice that completed the PD task earlier, the criteria mean (correct trial percentage over the last two days) was calculated, and this number was included into the overall mean for the subsequent sessions. Therefore, when the percentage of trials was plotted against sessions, each session had an equal number of mice, even though some mice finished the assay faster than other. The number of days taken to reach criteria was plotted separately. There was a pre-defined 25 days cut-off point (described in Chapter 3, section 3.2.4), after which if mice did not complete the criteria, they were excluded.

The days needed to reach the PD test criteria are indicative of learning and memory abilities (Copping et al., 2017; Yang et al., 2015). To calculate the estimated sample size needed to

detect difference in days to reach criteria between two groups, the following was used: mean of days necessary to reach criteria μ = 14.6 (days; wildtype pilot data, Chapter 3), SD=7.09 and effect size = 6-8 (days; chosen based on published studies (Copping et al., 2017; Yang et al., 2015)). The estimated cohort size that would be necessary to detect a significant difference in the means with a smaller effect size of 6 days was n=22. A larger effect size of 8, would require n=13 animals.

In Chapter 4, due to breeding constraints, a cohort of n=12 per genotype was put through training, however only n=10 wildtype and n=8 *Kptn*$^{-/-}$ were progressed onto the pairwise discrimination task (two wildtypes and four mutants were excluded because they did not reach criteria within 25 day cut-off).

2.3.7 Barnes maze

Barnes maze is a spatial test that is a dry alternative of the Morris water maze and relies on the inherent tendency of mice to want to escape an aversive environment (Harrison et al., 2006; Harrison et al., 2009; Koopmans et al., 2003). Mice placed on a brightly lit open table surface with 20 holes around the periphery must repeatedly locate an escape box beneath one of the holes, the *target*, with the aid of spatial cues (Harrison et al., 2006; Harrison et al., 2009; Koopmans et al., 2003). Barnes maze had the dimensions of 120cm diameter, 20x5cm holes and 4 goal box locations (Fig.2.9). External cues of different shapes and sizes were placed around the walls of the room, to help the mouse navigate. The mice were tested in the dark with overhead lights facing the Barnes maze tables, under maximum illumination.

Three days before the first session, cages were cleaned and an escape box was placed in the centre of the clean home cage. Mice were placed into the box and allowed to climb out of it. If the mouse did not leave the box after a minute, it was gently encouraged back out into the cage.

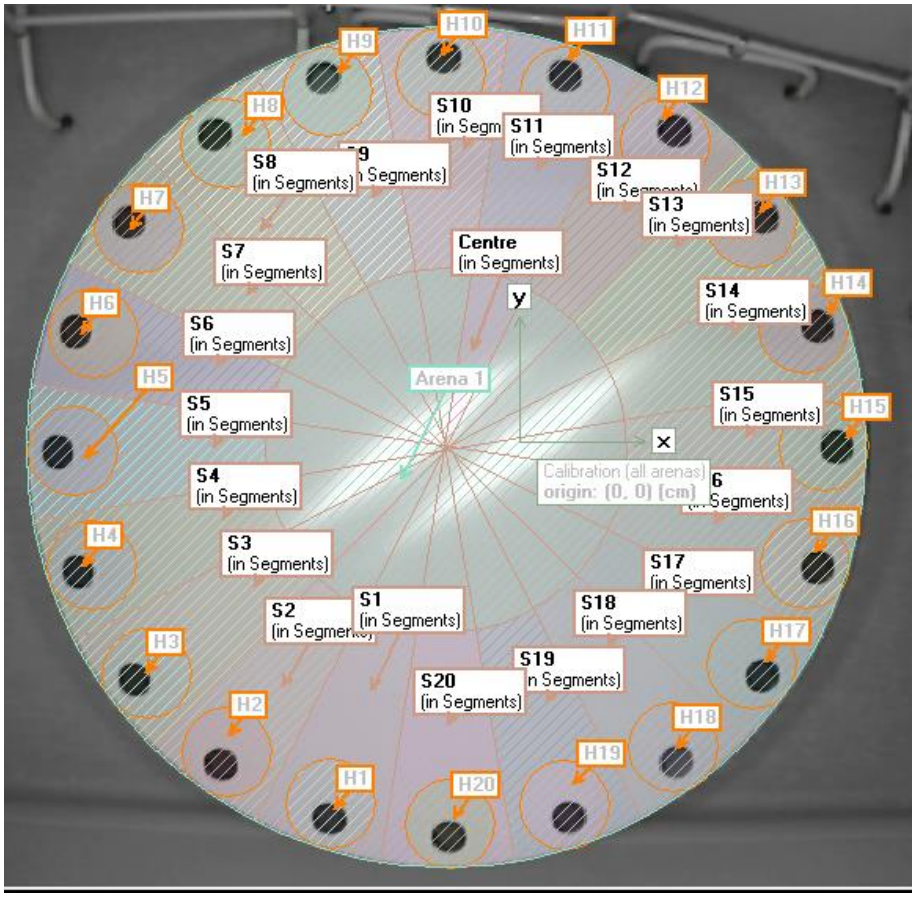


Figure 2.9 Schematic of the Barnes maze table set-up. Predefined areas around the 20 holes (H1-20), triangular segments from the hole towards the centre (S1-20), and the central circular area (Centre) are outlined.

The mice were tested sequentially and in the same order on every day, with two mice tested in parallel on two tables. Fifteen minutes before testing, mice were singly housed in new cages with bedding, food, and fun tunnel, and a handful of old bedding from the home cage of the mouse.

Before the first trial on Day 1, the mice were habituated to the escape box in the *introduction trial* (Table 2). The mouse was placed into the escape box and the opening of the box was blocked for 20 seconds, after which the mouse was free to climb out of it and explore for the remainder of the one minute. After one minute, if the mouse was not in the box, it was shepherded towards the entrance of the target hole and into the box. Mice were

trained for two days to locate the target hole, four trials per day (Table 2; Fig.2.10A,B). Trials ended either after four minutes or when the mouse went inside the escape box. Learning was assessed by measuring primary latency to reach the target and a total number of errors recorded before the mice went inside the escape box. After two days of training, the escape box was removed and the percentage of time spent around the target hole relative to other holes was calculated during the *24h probe trial* (Table 2; Fig.2.10A). After this trial, the location of the target hole was changed (Fig.2.10C) and the mice were re-trained to locate the new target (for three days) (Table 2; Fig.2.10A). Training N2 started directly after the 24h probe trial and consisted of only two training trials on Day 3, followed by four trials per for two more days (Day 4-5) (Table 2; Fig.2.10A). On Day 8, 72 hours after the last training, the escape box was removed and the mice were tested for memory retention of the target location in the *72h probe trial* (Table 2; Fig.2.10A-C). It was important to train the mice after the initial 24h probe and to re-baseline before the subsequent 72h probe, to counteract the potential extinction caused by a lack of escape hole on the 24h probe day (Chudasama et al., 2001).

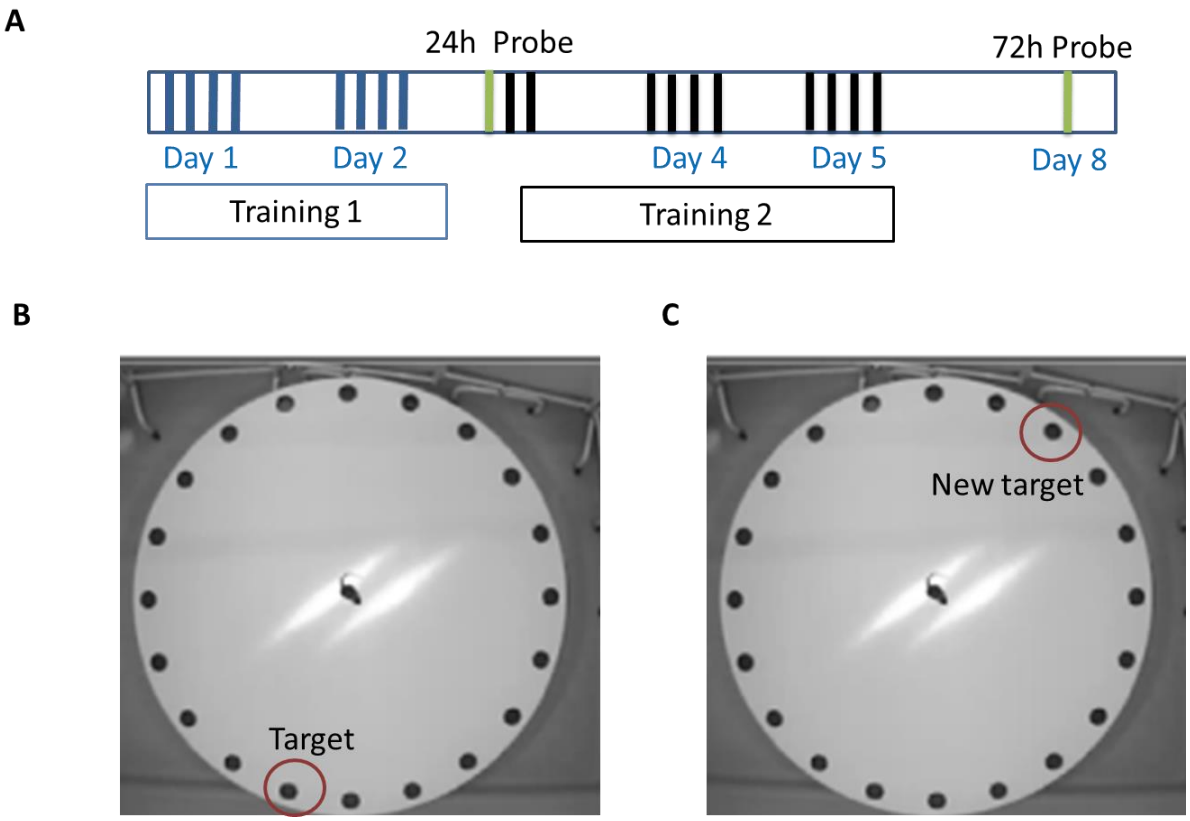


Figure 2.10 Barnes maze set up. A. Experimental design consisting of training with one target location (training 1) for two days (Day 1-2; four trials per day; blue bars), followed by a 24h probe trial (green bar) and then training with a different target location (training 2) for two days (Day 3-5; four trials; black bars) followed by a 72h probe trial (green bar); **B.** Target hole locations during training N1; **C.** Target hole locations during training N2.

Table 2 Barnes maze schedule over eight days.

Day	Trial	Type	Duration	Notes
1	Pre-trial	Introduction	1 minutes	
1	1	Training 1	4 minutes	
1	2	Training 1	4 minutes	
1	3	Training 1	4 minutes	
1	4	Training 1	4 minutes	
2	1	Training 1	4 minutes	
2	2	Training 1	4 minutes	
2	3	Training 1	4 minutes	
2	4	Training 1	4 minutes	
3	1	24h Probe	4 minutes	No Goal Box
3	2	Training 2	4	Location of target

			minutes	changed
3	3	Training 2	4 minutes	
4	1	Training 2	4 minutes	
4	2	Training 2	4 minutes	
4	3	Training 2	4 minutes	
4	4	Training 2	4 minutes	
5	1	Training 2	4 minutes	
5	2	Training 2	4 minutes	
5	3	Training 2	4 minutes	
5	4	Training 2	4 minutes	
8	1	72h Probe	4 minutes	No Goal Box

2.3.7.1 Mice

The percentage of time mice spent at the target hole comparative to other holes when the escape box is removed during probe trial is indicative of memory retention (Harrison et al., 2006; Harrison et al., 2009; Koopmans et al., 2003). The estimated cohort size needed to detect a preference in time spent at the target hole was calculated using the following: mean μ =12.845 (% time at the target; wildtype data Chapter), reference value = 5 (% time; 'no preference' for any hole), SD=8.11. The estimated cohort size n=9.

Kptn (n=16 wildtype, n=11 *Kptn* mutants), *Setd1a* (n=15 wildtype, n=12 *Setd1a* mutants), and *Arid1b* (n=16 wildtype, n=12 *Arid1b* mutants) cohorts underwent both Training 1 and 2. *Setd5* (n=12 both genotypes) cohort underwent only Training N1 and 24h probe. The inclusion Training N2 and the 72h probe came after *Setd5* mouse line had already been tested, and due to breeding issues with this line, it was not possible to repeat the assay to test for the 72h probe. *Zmynd11* mice did not undergo Training N2, but rather were re-trained briefly (two training trials) with the same target location immediately after 24h probe and then tested for 72h memory retention during the 72h probe. This made *Zmynd11* (n=12 both genotypes) data not directly comparable to the other lines tested, but allowed to test a shorter Barnes maze protocol for future studies.

2.3.8 Statistics

The statistical analyses, including Student's t-test, Welch's t-test, two-way ANOVA, and one-way ANOVA were done using GraphPad Prism 6 software. The error bars on all the graphs represent the standard error of the mean (SEM). Where the assumption of normality was rejected by the D'Agostino-Pearson test the non-parametric Mann-Whitney *U* test was used instead. Matched and paired analysis was done wherever appropriate.

2.4 MRI voxel-based morphometry of mouse brains

For brain morphometric analysis (Chapter 4), 23-week old adult male and female mice (n=8 per gender) were terminally anaesthetised by intraperitoneal injection of 0.1mL pentobarbitone sodium. After the mice were no longer responsive, 20 mL of cold PBS was injected into the left ventricle of the exposed heart at low speed, followed by 20 ml of 4% PFA, either by hand (20-50 ml syringe) or using a perfusion pump. After the transcardial perfusion, skulls were collected and stored in formalin solution at 4^0C until they were ready to be imaged, at which point the skulls were transferred to PBS.

MRI (Magnetic Resonance Imaging) imaging and analysis were performed by Dr. Stephen J. Sawiak of the Wolfson Brain Imaging Centre, University of Cambridge, using previously published methodology (Sawiak *et al.,* 2013). Brains were scanned using a Bruker PharmaScan 47/16 system at 4.7T with a manufacturer-provided birdcage transmit-receive coil. The imaging protocol was fast spin echo (scan parameters: repetition time 2000 ms, effective echo time 16 ms, echo train length 4, bandwidth 32 kHz, matrix 256×192×128, field of view 1.79×1.34×0.90 cm3, resolution 70 µm isotropic with two averages). MRI Tensor-based morphometry brains were segmented into tissue types: grey matter, white matter, cerebrospinal fluid and 'other'. This was followed by voxel-based quantification of brain volume. Voxel-based morphometry is an automated process of analysing morphological differences between images of brains by performing voxel-wise statistics (Ashburner and Friston, 2000). The images are registered into the same stereotactic space and segmented into images of usually four tissue types: grey matter, white matter, cerebrospinal fluid and 'other'. This technique has been previously shown to successfully detect even subtle differences in Huntington's disease mouse model brain (Sawiak *et al.,* 2013). To control the type I error rate due to multiple comparisons, an adjusted p-value was used for a false-discovery rate at $p < 0.05$.

2.5 X-ray cephalometrics

Skull X-ray imaging was performed on live 14-week old (± 4 days) age-matched littermates and *Kptn*$^{-/-}$ mutants (n=7 per genotype) (Chapter 4). Digital X-ray images were acquired using the Faxitron system MX20 scanner (Faxitron X-ray Corporation) in collaboration with the Mouse Genetics Programme (MGP) staff. Mice were anaesthetised for the procedure (intraperitoneal injection of weight adjusted ketamine/xylazine anaesthetic solution) and recovered with atipamezole at experiment terminus.

2.6 Histomorphometric analysis (P0 and adult brains)

For the histomorphological study (Chapter 4), adult mice aged 16 weeks (n=8 per genotype and sex) were perfused in the same way as described in section 2.4, the brain was removed from the skull and post-fixed in formalin for 48 hours, after which it was kept in PBS. For the P0 brains, male mice were culled at birth using schedule 1, brains were removed from skull and post-fixed for 12 days before being transferred into PBS.

The morphometric analyses were performed by Dr. Binnaz Yalcin and lab members at the Institute of Genetics and Molecular and Cellular Biology, France using a previously published method (Mikhaleva et al., 2016). Brains were paraffin embedded and 5µm sectioned were collected and analysed. Seventy-eight brain parameters across 20 distinct brain regions were analysed for neuroanatomical defects in mutant mice relative to littermate controls (Table 3). This consisted of a systematic quantification of the same three coronal brain regions, namely the Section 1 (Bregma +0.98mm), Section 2 (Bregma -1.34mm) and Section 3 (Bregma -5.80mm), down to cell level resolution and blind to the genotype (Table 3). To minimize environmental and genetic variation, mice were separately analysed according to their gender.

Focusing on two cortical regions (motor and somatosensory) of adult male mice, cell count, regional area, cell size, solidity and circularity were quantified at each cortical layer (layer I to layer VI) as well as the corresponding layer area at position Bregma -1.34mm.

Proliferation was measured in 16 weeks old mice by staining for a cell proliferation marker Ki67 using 3 sections throughout the rostro-caudal extent of the dentate gyrus. The sections were probed using 1:1000 rabbit polyclonal anti-Ki67 (Vector labs, VPK451). For Nissl staining, slides were incubated with 0.1% cresyl violet acetate solution in a 56°C water bath.

For brain morphology analysis of P0 mouse brains, a quantification approach, equivalent to that used for adults, of 53 parameters of size and surface at Bregma +2.19mm (Section 1, Table 4) and +3.51mm (Section 2) (equivalent to adult section 1 and 2) was used for male P0 mice (n=8 wildtype littermate controls, n=9 *Kptn*$^{-/-}$) (Table 4).

Sections were scanned using Slide scanner (Hamamatsu, NanoZoomer 2.0HT, C9600 series) and accessories (racks and NanoZoomer digital pathology, version 2.5.64 software) at 20x magnification and analysed using ImageJ (78 measurements) or manually quantified (Ki67 positive cell count).

Table 3 List of parameters for adult coronal analysis.

Section analysed	Parameter	Units
Section 1 (Bregma +0.98mm)	Total brain area	cm^2
	Lateral ventricle, left hemisphere	cm^2
	Lateral ventricle, right hemisphere	cm^2
	Cingulate cortex, left hemisphere	cm^2
	Cingulate cortex, right hemisphere	cm^2
	Genu of the corpus callosum	cm^2
	Caudate putamen, left hemisphere	cm^2
	Caudate putamen, right hemisphere	cm^2

	Anterior commissure, left hemisphere	cm^2
	Anterior commissure, right hemisphere	cm^2
	Piriform cortex, left hemisphere	cm^2
	Piriform cortex, right hemisphere	cm^2
	Cingulate cortex, width, left hemisphere	cm
	Cingulate cortex, width, right hemisphere	cm
	Cingulate cortex, height	cm
	Genu of the corpus callosum, width, top	cm
	Genu of the corpus callosum, width, bottom	cm
	Genu of the corpus callosum, height	cm
	Primary motor cortex, left hemisphere	cm
	Primary motor cortex, right hemisphere	cm
	Secondary somatosensory cortex, left hemisphere	cm
	Secondary somatosensory cortex, right hemisphere	cm
Section 2 (Bregma -1.34mm)	Total brain area	cm^2
	Lateral ventricle, left hemisphere	cm^2
	Lateral ventricle, right hemisphere	cm^2
	Dorsal third ventricle	cm^2
	Retrosplenial granular cortex, c region, left	cm^2

	hemisphere	
	Retrosplenial granular cortex, c region, right hemisphere	cm^2
	Corpus callosum	cm^2
	Dorsal hippocampal commissure	cm^2
	Hippocampus	cm^2
	Amygdaloid nucleus, left hemisphere	cm^2
	Amygdaloid nucleus, right hemisphere	cm^2
	Piriform cortex, left hemisphere	cm^2
	Piriform cortex, right hemisphere	cm^2
	Mammillothalamic tract, left hemisphere	cm^2
	Mammillothalamic tract, right hemisphere	cm^2
	Internal capsule, left hemisphere	cm^2
	Internal capsule, right hemisphere	cm^2
	Optic tract, left hemisphere	cm^2
	Optic tract, right hemisphere	cm^2
	Fimbria of the hippocampus, left hemisphere	cm^2
	Fimbria of the hippocampus, right hemisphere	cm^2
	Habenular nucleus, left hemisphere	cm^2
	Habenular nucleus, right hemisphere	cm^2
	Third ventricle	cm^2

	Hypothalamus, left hemisphere	cm^2
	Hypothalamus, right hemisphere	cm^2
	Retrosplenial granular cortex, c region, width, left hemisphere	cm
	Retrosplenial granular cortex, c region, width, rigtht hemisphere	cm
	Retrosplenial granular cortex, c region, height	cm
	Corpus callosum, width	cm
	Corpus callosum, height	cm
	Total internal length of pyramidal cells	cm
	Dentate gyrus, left hemisphere	cm
	Dentate gyrus, right hemisphere	cm
	Molecular layer of the hippocampus, left hemisphere	cm
	Molecular layer of the hippocampus, right hemisphere	cm
	Radiatum layer of the hippocampus, left hemisphere	cm
	Radiatum layer of the hippocampus, right hemisphere	cm
	Oriens layer of the hippocampus, left hemisphere	cm
	Oriens layer of the hippocampus, right	cm

	hemisphere	
	Primary motor cortex, left hemisphere	cm
	Primary motor cortex, right hemisphere	cm
	Secondary somatosensory cortex, left hemisphere	cm
	Secondary somatosensory cortex, right hemisphere	cm
Section 3 (Bregma -5.80mm)	Number of folia	
	Total brain area	cm^2
	Fourth ventricle	cm^2
	Pons	cm^2
	Pyramidal tract, left	cm^2
	Pyramidal tract, right	cm^2
	Genu of the facial nerve, left	cm^2
	Genu of the facial nerve, right	cm^2
	Cochlear nucleus, left	cm^2
	Cochlear nucleus, right	cm^2
	Lateral cerebellar nucleus, left	cm^2
	Lateral cerebellar nucleus, right	cm^2
	Interposed cerebellar nucleus, anterior part, left	cm^2

	Interposed cerebellar nucleus, anterior part, right	cm^2
	Internal granular layer	cm^2

Table 4 List of parameters for P0 coronal analysis.

Section analysed	Parameter	Units
Section 1 (Bregma 2.19mm)	Total Brain Area	cm^2
	Area of Lateral Ventricle_Left	cm^2
	Area of Lateral Ventricle_Right	cm^2
	Area of Cingulate Cortex_Left	cm^2
	Area of Cingulate Cortex_Right	cm^2
	Width of Cingulate Cortex_Left	cm
	Width of Cingulate Cortex_Right	cm
	Height of Cingulate Cortex	cm
	Area of Genu of Corpus Callosum	cm^2
	Width of genu of Corpus Callosum	cm
	Height of genu of Corpus Callosum	cm
	Area of Caudate Putamen_Left	cm^2
	Area of Caudate Putamen_Right	cm^2
	Area of anterior commissure_Left	cm^2
	Area of anterior commissure_Right	cm^2
	Height of Motor cortex_Left	cm
	Height of Motor cortex_Right	cm

	Height of Somatosensory cortex_Left	cm
	Height of Somatosensory cortex_Right	cm
Section 2 (Bregma 3.51mm)	Total Brain Area	cm^2
	Area of Lateral ventricule_Left	cm^2
	Area of Lateral ventricule_Right	cm^2
	Area of Dorsal 3rd ventricule	cm^2
	Area of 3rd ventricle	cm^2
	Area of Retrosplenial Granular cortex_Left	cm^2
	Area of Retrosplenial Granular cortex_Right	cm^2
	Width of Retrosplenial Granular cortex_Left	cm
	Width of Retrosplenial Granular cortex_Right	cm
	Height of Retrosplenial Granular cortex	cm
	Area of the splenium of the corpus callosum	cm^2
	Width of the splenium of the corpus callosum	cm
	Height of the splenium of the corpus callosum	cm
	Area of Hippocampus_Left	cm^2
	Area of Hippocampus_Right	cm^2
	Total internal length of Pyramidal layer_Left	cm
	Total internal length of Pyramidal layer_Right	cm
	Area of the internal capsule_Left	cm^2

	Area of the internal capsule_Right	cm^2
	Area of the fimbria_Left	cm^2
	Area of the fimbria_Right	cm^2
	Area of the Amygdala_Left	cm^2
	Area of the Amygdala_Right	cm^2
	Area of the Hypothalamic nucleus_Left	cm^2
	Area of the Hypothalamic nucleus_Right	cm^2
	Area of the Habenula_Left	cm^2
	Area of the Habenula_Right	cm^2
	Area of the Thalamus	cm^2
	Height of Motor cortex_Left	cm
	Height of Motor cortex_Right	cm
	Height of Primary Somatosensory cortex_Left	cm
	Height of Primary Somatosensory cortex_Right	cm
	Height of Secondary Somatosensory cortex_Left	cm
	Height of Secondary Somatosensory cortex_Right	cm

2.7 RNA sequencing analysis

2.7.1 RNA tissue extraction

Tissue was homogenised in buffer RLT plus (Qiagen, 1053393) with β-mercaptoethanol (Sigma, M3148; 10µl/ml) using Qiagen TissueLyser LT, with sterile RNAse-ZAP treated steel

beads and operated at 50Hz for 2 minutes. Samples (n=6 per genotype) were pre-treated on gDNA eliminator columns and then extracted on RNeasy Plus columns as per manufacturer's protocol (Qiagen, Venlo, Netherlands), and were immediately snap frozen on dry ice and stored at -80C. An aliquot of each sample was quantified using 2100 Bioanalyzer (Agilent Technologies).

2.7.2 Library preparation

Library preparation and sequencing was performed by the WTSI DNA Pipelines Illumina Low-Throughput Team. Multiplexed libraries were prepared for sequencing using Illumina RNA Library Preparation Kits as per manufacturer's protocol. Paired end sequencing was performed on the Illumina HiSeq 2000 or V4 generating 75bp reads.

2.7.3 Alignment, mapping and differential gene expression

STAR version 2.5.2b was used to align sequenced reads to the altered version of the mouse reference genome and map the reads.

The following parameters were used:

fastq_convert = 1, star = 1, mem_cram = 2000, queue_cram = normal, mem_fastq = 18000, queue_fastq = long, mem_star = 36000, queue_star = long, mismatch = 4

The count data was used as input for differential gene expression analysis using R (version 3.2.2) DESeq2 package (versions 1.17.16). DESeq2 is a conservative model for differential expression analysis that employs a Benjamini-Hochberg procedure to control for multiple testing, returning an adjusted p-value (padj) for the differential gene expression. DESeq2 returns an estimation of Log2 fold change, a regularised log transformation which reduces the false positive rate for genes with low counts and high dispersion. MA plots were used to visualize the expression differences between genotypes, by transforming the data onto M (log ratio) and A (mean average) scales and plotting these values. The Log2 fold changes presented in the MA plots throughout this thesis refer to shrinkage estimation values (not

fold change of normalised read counts). DESeq2 cut-off of BHadjusted p-value<0.05 was used for all analyses. For differential gene expression analysis in DESeq2, transcript counts mapped to the mitochondrial genome were excluded

For the identification of functionally enriched terms in the differentially expressed genes, I performed Gene ontology (GO) and KEGG pathway enrichment analyses using the grpofiler online suite (http://biit.cs.ut.ee/gprofiler/index.cgi). A threshold of 5% FDR and an enrichment significance threshold of P<0.05 (hypergeometric test with Benjamini-Hochberg False Discovery rate correction for multiple testing) was used. In all analyses, a background comprised of only the expressed genes in the tissue studied (genes where the adjusted p-value yielded a numerical value, different to NA).

Chapter 3. Cognitive assay development

3.1 Introduction

3.1.1 Mouse modelling of intellectual disability

The mammalian central nervous system (CNS) contains a large number of cell types, with a unique combination of gene expression repertoires, which are necessary for cognitive function (Lein et al., 2007). As outlined in Chapter 1, due to the complexity and intricacy of the system, minimal changes to the genetic programming of the brain can lead to a diverse set of conditions ranging from neurodevelopmental to neurodegenerative disorders (Hu et al., 2014).

Mus musculus has been extensively used as a model organism to identify causal links between mutations and disease phenotypes, characterise candidate genes, provide insight into the associated disease mechanisms, and inform treatment. Mice are commonly used for modelling human disease for several reasons. Firstly, approximately 99% of mouse genes have human orthologues, and on average 85% of protein-coding regions are identical in both species (Chinwalla et al., 2002). Secondly, due to the availability of a high-quality reference genome and sophisticated genetic tools, the mouse genome can be manipulated and analysed with relative ease (Weyden et al., 2011). And lastly, the similarity of neurobiological pathways and their behavioural outputs in mice and humans, allows for parallels to be made with human conditions (Bucan and Abel, 2002).

As outlined in Chapter 1, genetic alterations are among the leading causes of intellectual disability (ID), disrupting the functionality of the nervous system and development of cognitive function (Ehninger et al., 2008; Fitzgerald et al., 2016; Wright et al., 2015). Genetically modified mice have been extensively used to study the pathologies associated with ID, which also sheds light on the pathways involved in normal cognition (Bakker et al., 1994; Dias et al., 2016; Celen et al., 2017; Mircsof et al., 2015). Testing the mice in a battery of behavioural and cognitive assays and employing multiple tasks for each behavioural domain of interest can allow for the detection of subtle cognitive and behavioural defects (Bucan and Abel, 2002; Crawley, 1997). If the mouse models parallel the phenotypes of the

human disorder, they may offer a valuable tool to study the molecular basis of, and aberrant pathways are leading to, the pathologies (Guy et al., 2001; Gogliotti et al., 2017). When characterising the mouse model, before undertaking assays of complex behaviour, it is critical to first measure general health parameters, as well as the sensory and motor functions of the mice, to avoid possible misinterpretation of the subsequent results (Spencer et al., 1995; Strekalova et al., 2005). Moreover, because mouse behaviour is sensitive to environmental changes and prior experiences, to ensure reproducibility the order in which the mice are tested should be consistent and the environmental conditions, such as external noise, housing conditions, handling regime, and time of testing must be controlled (Chesler et al., 2002; McIlwain et al., 2001; Ramos et al., 1997). Moreover, the genetic background strain of the test animals should be considered and controlled. Inbred strains contain fixed polymorphisms, some of which result in neurobehavioural differences between strains that affect their performance in behavioural and cognitive assays (Homanics et al., 1999; Keane et al., 2011). For example, 129 inbred strains have been reported to be less active than C57BL/6 strains in explorative assays, display increased anxiety-like behaviour and, depending on the sub-strain tested, there is conflicting data on their learning and memory abilities (Abramov et al., 2008). FVB/N mice are visually impaired, due to the *rd* mutation that causes retinal degeneration, making them inappropriate for assays that rely on visual cues, such as the Barnes maze (O'Leary et al., 2011). Therefore, to maximise the chances of detecting a phenotype associated with a mutated candidate gene, it is important to avoid modelling in strains with extreme traits, due to ceiling or floor effects that may mask the phenotypic consequences of the mutation. C57BL/6 strains have average performance in many behavioural paradigms. Moreover, testing mice of a common genetic background, such as C57BL/6 strains, facilitates the comparison of results across different laboratories (Bothe et al., 2005; Crawley et al., 2008). In this thesis work, all the mouse models characterised were of C57BL/6 background.

Some conditions, for example, those with complex inheritance patterns and multiple interacting genes are more challenging to replicate efficiently in mouse models (Watase and Zoghbi, 2003). In contrast, studies modelling highly penetrant monogenic disorders, such Type A and B Niemann-Pick disease in which the mutation causes the ablation of

sphingomyelinase enzyme activity, have successfully recapitulated the main endophynotes of the disorders, manifested as changes in behavioural and cognitive capacity, in the mice (Horinouchi et al., 1995; Otterbach and Stoffel, 1995). In this thesis work, I focused on modelling complex traits and monogenic disorders.

3.1.2 The approach and aims

My first objective was to established and refine a series of behavioural assays tailored for the identification of phenotypes associated with neurodevelopmental disorders. In this thesis work, I focused on assessing cognition through learning and memory paradigms, as they can be used to test multiple cognitive domains, including associative learning, spatial and non-spatial learning, short-term and long-term memory (Sweatt, 2004; Nithianantharajah and Grant, 2013). Out of the behavioural assays (outlined in Table 5) used extensively in our laboratory at the start of my PhD, only two were cognitive tests - Barnes maze (testing spatial memory and learning) and social recognition (testing olfactory memory) (Dias et al., 2016; Huckins et al., 2013; Sánchez-Andrade et al., 2005). I, therefore, set out to expand the cognitive assay repertoire to include novel object recognition (NOR) - assessing non-spatial memory, object displacement (OD) - assessing spatial memory, and pairwise visual discrimination (PD) - assessing association-based learning, to enable the detection of a dynamic range of deficits, associated with neurodevelopmental disorders, in a reliable and reproducible manner.

Table 5 Behavioural and cognitive assay usage status in our team at the start of my PhD.

Assay name	Testing	Existing status in the lab
Open field	Locomotion and activity	Used extensively
Light/dark	Anxiety-like behaviour	Recently optimised
Social recognition	Olfactory memory	Used extensively
Novel object recognition	**Recognition memory**	**Not established**
Object displacement	**Spatial memory**	**Not established**
Barnes maze	Spatial memory	Recently optimised
Pairwise discrimination	**Associative learning**	**Not established**
Sociability	Social interaction	Recently optimised

In this Chapter, I will discuss the optimisation of the assays highlighted in bold (Table 5), outlining the series of undertaken pilots, which aided in the development of a testing strategy I subsequently used in Chapter 4 and 5.

3.1.3 Novel object recognition and object displacement assays

Recognition memory relies on the ability to identify a previously encountered item as familiar and is dependent on the functioning of the medial temporal lobe of the brain, which consists of neuroanatomical structures, including the hippocampus and the adjacent parahippocampal regions, including the perirhinal and entorhinal cortices (Squire et al., 2007). Novel object recognition (NOR) assay, initially outlined in 1988, was designed to test recognition memory similarly to visual paired-comparison task used in clinical neuropsychology (Ennaceur and Delacour, 1988). NOR assay relies on rodent's innate

preference for investigating novel over the familiar object (Ennaceur and Delacour, 1988). Longer investigation of the unfamiliar object by the test animal is indicative of an acquired memory for the familiar object.

The parahippocampal regions of the temporal lobe (in particular the perirhinal cortex) and the prefrontal cortex (PFC) have been implicated to have a critical role in the NOR task (Banks et al., 2012; Bussey et al., 1999). While the role of perirhinal cortex in NOR has been well established, the importance of the hippocampus in NOR is disputed (Broadbent et al., 2004; Forwood et al., 2005; Gilbert and Kesner, 2003; Hammond et al., 2004; Kim et al., 2014; Yi et al., 2016). The inconsistency of NOR results between different groups may be due to the variability in the testing method across studies (training procedures, arena size, and type of objects) or the differences in the type or extensiveness of hippocampal lesions. There is a body of work which has shown the involvement of the hippocampus in NOR when it is tested in a complex environment with spatial and contextual cues (Balderas et al., 2008; Forwood et al., 2005; Nemanic et al., 2004). This suggests that the hippocampus may be involved in the retrieval of contextual information, while the other regions of the temporal lobe are required for consolidation of familiarity of the objects. Furthermore, there is increasing evidence of functional overlap between various regions in the brain, implying the roles between brain regions are not as delineated as previously thought (Vann and Albasser, 2011).

A modified version of the NOR assay, object displacement (OD), was developed to assess spatial memory in rodents (Ennaceur et al., 1997). OD task exploits the natural tendency of rodents to explore their environment and the preference to investigate a moved object when compared to a static object (Tuscher et al., 2015). While the NOR assay tests the memory for the *identity* of the object, the OD task assesses the memory for the *location* of the object. It is, therefore, unlike NOR, a spatial memory task, and relies primarily on the hippocampus (Anderson and O'Mara, 2004; Ricceri et al., 2000). As with NOR, multiple testing strategies exist for this assay (Fernandez and Garner, 2007; Dere et al., 2005; Larkin et al., 2008).

3.1.4 Pairwise visual discrimination task

Automated touchscreen technology is becoming an increasingly popular platform for assessing cognitive function in rodents (Bussey *et al.,* 2012; Morton *et al*., 2006). Although relatively less ethological than some traditional methods, touchscreen technology exploits the natural tendency of rodents to investigate novelty and learn the consequence of exploring particular stimuli (Horner *et al*., 2013a; Horner *et al*., 2013b). The touchscreen tasks have a great level of translational potential, as they operate similarly to the Cambridge neuropsychological test automated battery (CANTAB) methods used to test patients (Bussey et al., 2008; Nithianantharajah et al., 2015; Robbins et al., 1994). Moreover, multiple cognitive tests, such as autoshaping, visual discrimination, and visuomotor conditional learning can be run using the same apparatus in a standardised and relatively high-throughput manner (Horner *et al*., 2013; Oomen *et al*., 2013).

Pairwise visual discrimination (PD) assay is one of the cognitive tasks that can be assessed using the touchscreen platform and is an operant conditioning task that tests reward-based associative perceptual memory and learning (Bubser et al., 2014; Horner *et al*., 2013). It is a hippocampus-independent task and is reliant on neuroanatomical structures such as the dorsal striatum, which mediates the formation of reinforced stimulus-response associations (Brigman et al., 2013; Chudasama and Robbins, 2003; Izquierdo et al., 2013). Of note, PD requires intact motor and visual abilities for the rodent to navigate inside the chamber, collect the reward, respond to the stimuli on the screen when necessary, and discriminate between two visual stimuli. Mouse strains have been shown to perform differently on this the task (Graybeal et al., 2014). For example, FVB/N mice cannot be used due to their visual impairment (O'Leary et al., 2011).

3.2 Results

3.2.1 NOR assay

Since NOR assay relies on the ability of mice to differentiate between objects, the choice of objects is critical to the study. There is vast variability in the literature of what objects are

used, and often this varies not only between studies but between test animals in the same study, making it difficult to reproduce results (Tuscher et al., 2015). I therefore first investigated which objects would generate the most reproducible results, aiming to use the same set of objects for all the mice. Objects were chosen based on similar size, but different textures and shapes, to allow for easier discrimination (Fig.3.1).

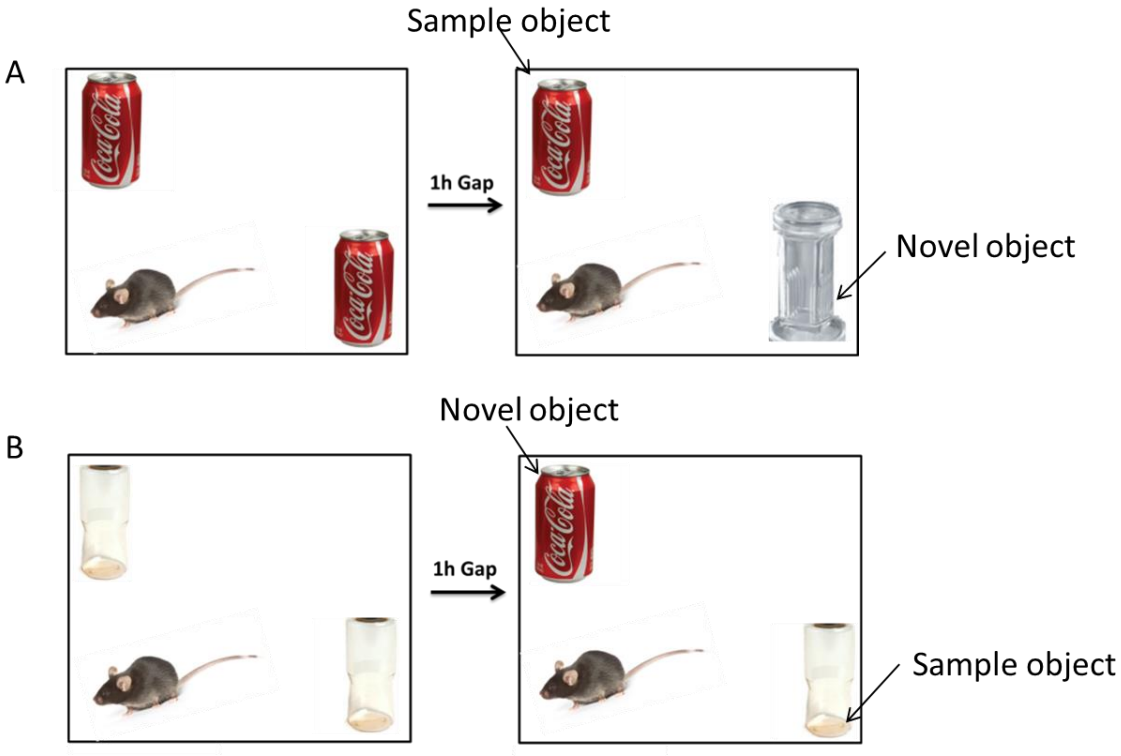


Figure 3.1 Novel object recognition set-up in the two pilot studies. In both pilot studies, the mice were habituated to two identical (sample) objects, presented with the familiar (sample) object and an unfamiliar (novel) objects after 1-hour interval. **A**. In the first pilot study, the sample object was a 300mL Coca-Cola can and the novel object was a glass slide washstand; **B.** In the second pilot, the sample object was a 300mL Coca-Cola can and the novel object was a water bottle, used in mouse caging, without the lid.

The first object pair tested was a Coca-Cola can and glass slide washstand (Fig.3.1A). Both objects can be easily purchased, are a similar size, but different textures and colours. To assess whether mice have an inherent preference for either of the objects, naïve wildtype

mice were presented with both objects and the time taken to investigate both objects were recorded. The mice investigated the glass stand on average significantly longer than the Coca-cola can (P=0.0391; Wilcoxon matched-pairs signed rank test) (Fig. 3.2A). Due to this inherent bias, the mice were tested with a different object set - water bottle and a Coca-Cola can (Fig.3.1B). There was no significant inherent preference for either of the objects (P= P=0.25, Wilcoxon matched-pairs signed rank test) and this pair was thus picked for the NOR assay (Fig.3.1B; Fig. 3.2B). The Coca-cola can data failed D'Agostino & Pearson omnibus normality test in pilot 1 (P=0.006) and pilot 2 (P=0.004).

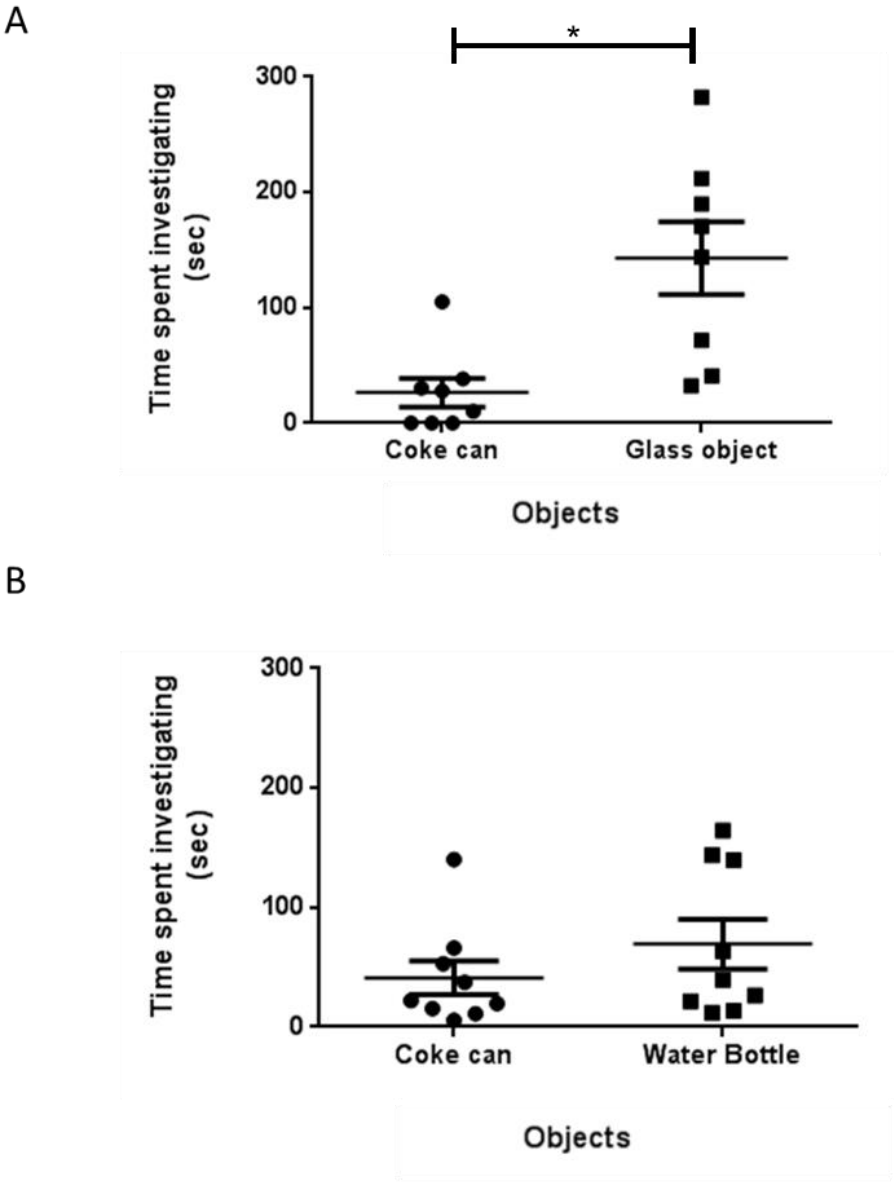


Figure 3.2 Novel object recognition inherent preference test. The time spent investigating either of the two objects from the pair was plotted. **A.** Naïve wildtype adult mice (n=8) spent significantly longer investigating the glass stand relative to the Coca-cola can (P=0.0391*). **B**. There was no inherent bias detected for either of the objects (P=0.25). Values are plotted as mean ± SEM.

Once the object set was picked, a new cohort of mice (n=9) was put through NOR training phase, during which the mice were exposed to two water bottles for 10 min (Fig.3.1B). One

hour later, in the testing phase, the mice were exposed to two objects: water bottle (familiar sample object) and Coca-cola can (novel object) for 5 minutes and the time spent investigating both objects was recorded (Fig.3.1B). The preference for the unfamiliar object was calculated as a proportion of the total investigation time. No significant difference was observed in the wildtype mice in the time mice spent near either of the objects (P=0.9328, t=0.08704 df=8, two-tailed paired Student's *t*-test), implying that the mice were unable to distinguish between the two objects (Fig.3.3).

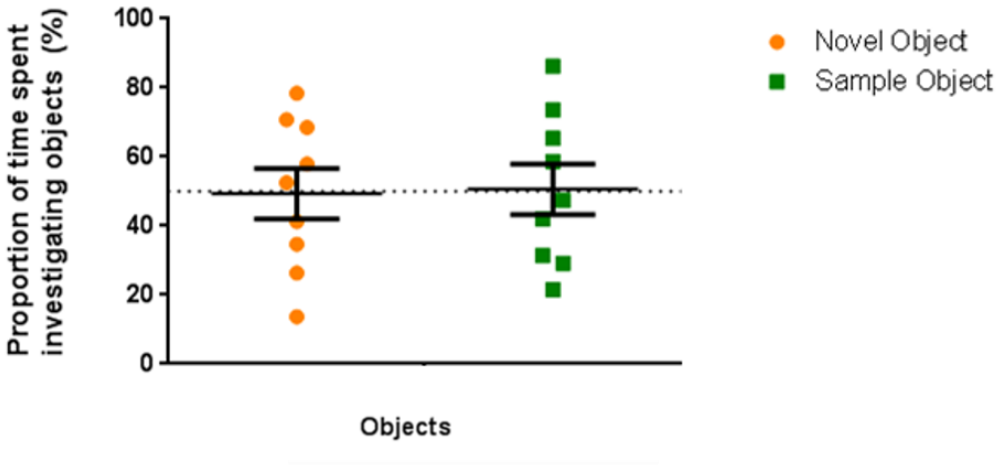


Figure 3.3 Novel object recognition pilot. The plot shows the percentage of time wildtype mice (n=9) spent investigating the familiar sample object (plotted in green) and unfamiliar novel object (plotted in orange). There was no significant increase in the percentage of time spent with a novel object (P= 0.9328). Values are plotted as mean ± SEM.

3.2.2 OD assay

Three identical Lego objects were used for the OD pilot, based on a personal correspondence with Dr Lukas von Ziegler at the Laboratory of Neuroepigenetics, Brain Research Institute, University of Zurich. In the training phase, the mice were habituated to the empty chamber for three days in a row, after which they were presented with three

identical Lego objects. In the testing phase, the mice were re-exposed to the three Lego objects 24 hours later, but one of the objects was moved to a different location in the arena (Fig.3.4). The time spent investigating both objects was recorded.

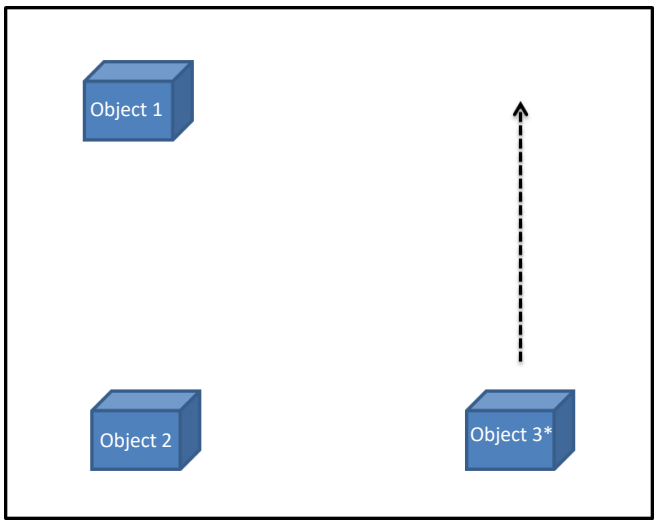


Figure 3.4 Schematic representation of the object displacement protocol. One of the identical objects (Object 3*) moved between the training and testing phases (24 hours apart).

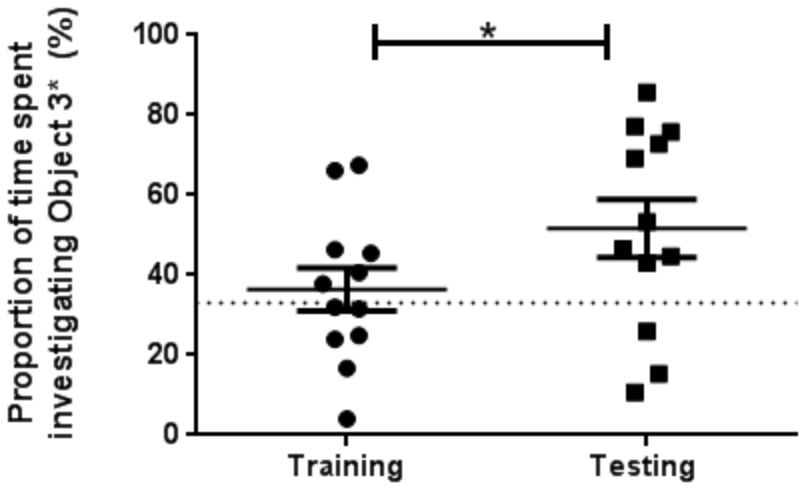


Figure 3.5 Object displacement pilot with Lego objects. The plot shows the percentage of time mice (wildtypes, n=12) spent investigating Object 3* as a proportion of overall investigation time during training (Day 1) and testing (Day 2) phase. There was a significant difference in time spent investigating the moved object (Object 3*) during testing when compared to the initial investigation during the training phase (P=0.0347*). A significant difference from chance or no preference (33%; represented by dotted line) and was detected only during testing (training: P = 0.54; testing: P= 0.026). Values are plotted as mean ± SEM.

The time spent investigating the moved object (Object 3*) was calculated as a proportion of the total investigation time. A significant difference from the reference value of 33% (one out of the three objects, and implying no preference for any of the objects) for Object 3* was detected only during testing (training: P = 0.54, t=0.6364 df=11; testing: P= 0.026, t=2.574 df=11; One sample t-test) (Fig.3.5). There was a significant increase in the investigation of Object 3* between training and testing phase (mean difference: 15.27sec, P=0.035, t=2.408 df=11, paired two-tailed Student's *t*-test), indicating that the wildtypes were able to discriminate between moved and stationary objects (Fig.3.5).

3.2.3 Object discrimination paradigm

I next investigated whether the NOR and OD assays could be used not only sequentially in order to assess both spatial and non–spatial memory, but in addition whether OD assay could be used as NOR acquisition training. For this, I modified both NOR and OD protocols (Chapter 2, section 2.3.3).

13 male naïve wildtype mice were presented with an object pair (vial and light bulb) and the time they spent investigating both of the objects was recorded (Chapter 2, section 2.3.3). No inherent preference for either of the objects was detected (P= 0.9238, t=0.09771 df=12, paired Student's t-test) (Fig.3.6A). Due to the lack of object bias, a new set of wildtype mice (n=14) were tested using the object discrimination paradigm (which involved OD assay, followed by NOR assay as outlined in Chapter 2, section 2.3.3, Fig.2.4A). The results showed an increase in time spent with the moved object at OD test phase (P= 0.0061, t=3.046 df=13, one-sample Student's *t*-test), as well as an increase in the time spent with the novel object, 24h later, during NOR testing (P<0.0001, t=6.073 df=12, one-sample Student's *t*-test) (Fig.3.6C). One mouse was eliminated from NOR analysis because it did not meet the pre-determined exclusion criteria (Chapter 2, section 2.3.3).

Overall, the wildtype mice were able to discriminate between moved vs stationary objects, as well as showing a preference for the novel object. These results suggest that using the vial-light bulb object pair, object discrimination can be used to assess both spatial memory deficits (OD phase) and recognition memory (NOR phase) in mutant mice.

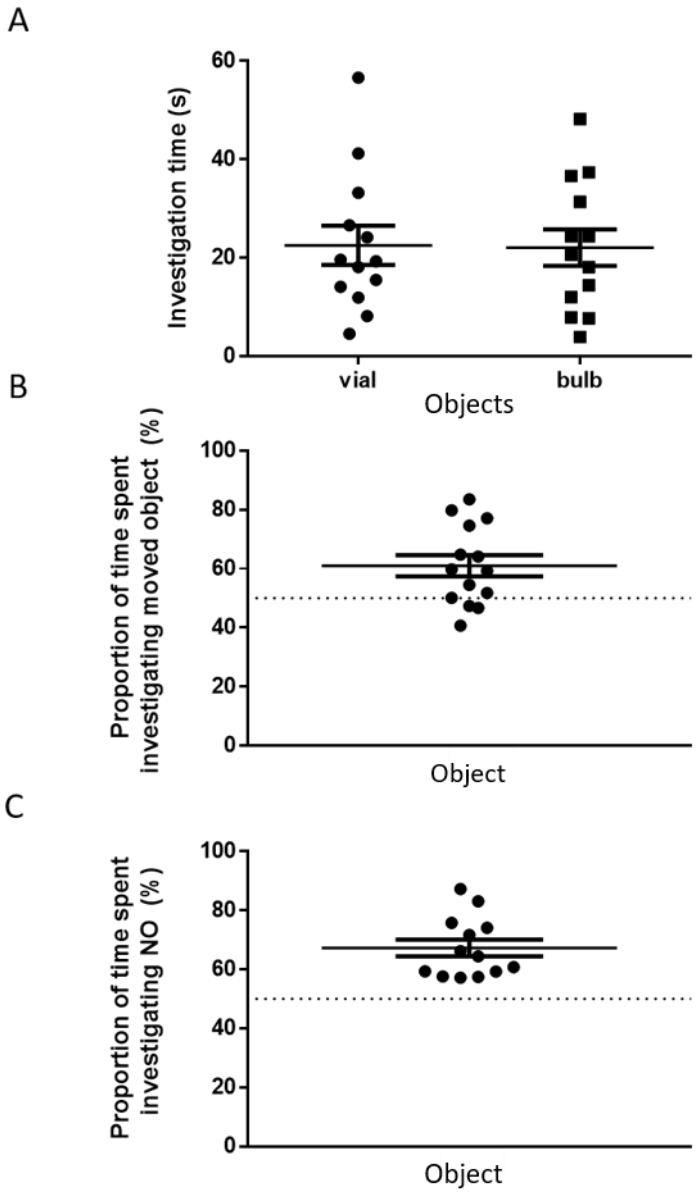


Figure 3.6 Object discrimination testing. A. The time naïve wildtype mice spent investigating two objects (vial and light bulb) is plotted. No difference in investigation time of the object was detected (P= 0.9238). **B.** The results of object displacement testing phase plotted as a proportion (%) of total time spent with the moved object. The mice (n=14) investigated the moved object significantly longer (P=0.0094) than expected by 50% chance (dotted line). **C**. The results of novel object recognition testing phase plotted as the proportion (%) of total time spent with the novel object. The mice (n=13) investigated the novel object significantly longer (P<0.0001) than expected by 50% chance (dotted line). Values are plotted as mean ± SEM.

3.2.4 PD assay

I had two aims when establishing this assay at the Sanger Institute. Firstly, I wanted to add to my battery of memory and learning tests an assay that is hippocampus-independent, to have the ability to capture a broader range of cognitive deficits (Chapter 4). Secondly, due to its translational potential and automated design, I wanted to assess whether this type of assay could be used in a cognitive screen (Chapter 5). The latter posed certain limitations to the protocol design - no testing on the weekend and time limitation for the duration of the assay. To take into account the weekend break in testing, I established a set of stringent PD rules before the start of the pilot. Due to the time constraint imposed on the assay, the aim was to identify, based on the outcomes of the piloting, a fixed time after which the experiment would be stopped even if not all the animals completed the criteria.

3.2.4.2 Pilot Number 1

The first pilot I carried out was with C57BL/6J mice (n=12). Mice were housed in groups of four and advanced through training phases as a group (Chapter 2, section 2.3.6). This entailed not moving any mice onto the next training phase until all the mice passed the set criteria, allowing for synchrony of assessment. At the start of a new training phase, there was a reduction in the number of trials completed in 60 minutes (Fig.3.7A). However, as the mice did more sessions and learned the required set of skills to complete the task, they increased the number of trials, eventually reaching the criteria for that training phase (Fig.3.7A). Two mice out of 12 did not complete criteria in the *Punishment phase* and were excluded from subsequent sessions. The rest of the mice (N=10) were moved onto PD testing.

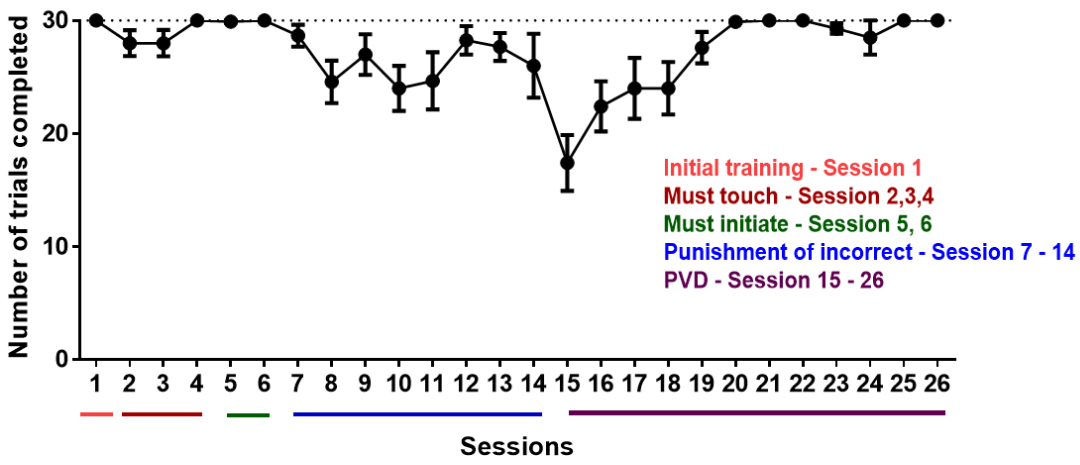
In PD testing, the mice were exposed to two images (image 1 and 2), counterbalanced for which image was assigned as CS+ (Chapter 2, Section 2.3.6, Figure 3.7). It took 12 sessions for all the mice to reach criteria of 30 trials in 60 minutes with 80% of trials done correctly over two consecutive days, with a *correct trial* defined as a trial where the CS+ image was

nose-poked (Fig.3.7B). The time taken to complete the trials decreased significantly as the mice increased their *accuracy* (measured as % of trials done correctly) (Fig.3.7C.). The number of correction trials was inversely correlated with the accuracy of performance (number of trials done correctly) (significant interaction: $F(11, 198) = 23.02$, $P<0.0001$, two-way ANOVA) (Fig.3.7D).

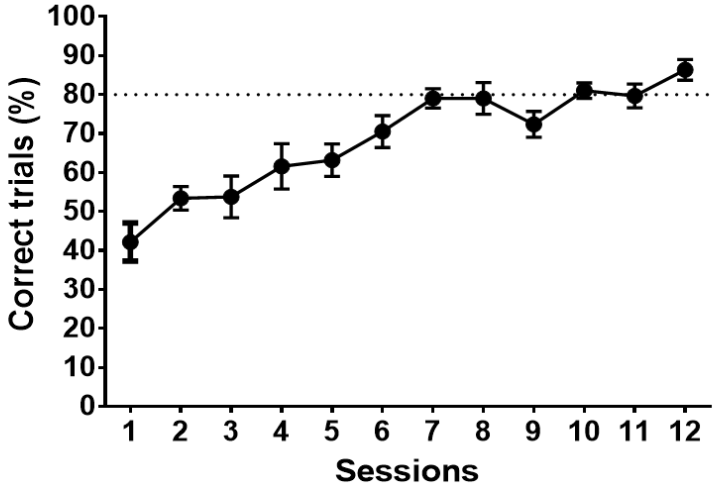
To test for the effect of weekend breaks on the performance of the mice in PD task, I analysed the number trials completed (Fig.3.7A) and speed of completion (Fig.3.7E) after a two-day break between Session 3 and 4 (labelled as Session 17 and 18 in Fig.3.7A) and after a three-day break between Session 8 and 9 (labelled as Session 22 and 23 in Fig.3.7A). There was no significant difference in the trials completed after two- or three-day break in PD task (Session 3 vs 4: $P >0.9999$, Session 8 vs 9: $P=0.9138$, post-hoc analysis after one-way ANOVA), nor did the time taken to complete the trials differ (Session 3 vs 4: $P= 0.7958$, Session 8 vs 9: 0.3464, post-hoc analysis after one-way ANOVA).

When the performance was analysed based on two groups, depending on what image was assigned as CS+ (group 1: all the mice that had image 1 as CS+, group 2: image 2 as CS+) there was an overall difference in performance between groups (significant interaction: $F(11, 84) = 3.096$, $P=0.0016$, group difference: $F(1, 84) = 19.05$, $P<0.0001$, two-way ANOVA). There was no difference in the percentage of trials done correctly per CS+ group in session 1, however, in session 3 and 4 the group with image 2 as CS+ had significantly higher percentage of trials done correctly, implying a potential image bias (session 3: $P \leq 0.001$; session 4: $P \leq 0.0001$, post hoc analysis after a two-way ANOVA) (Fig.3.7F).

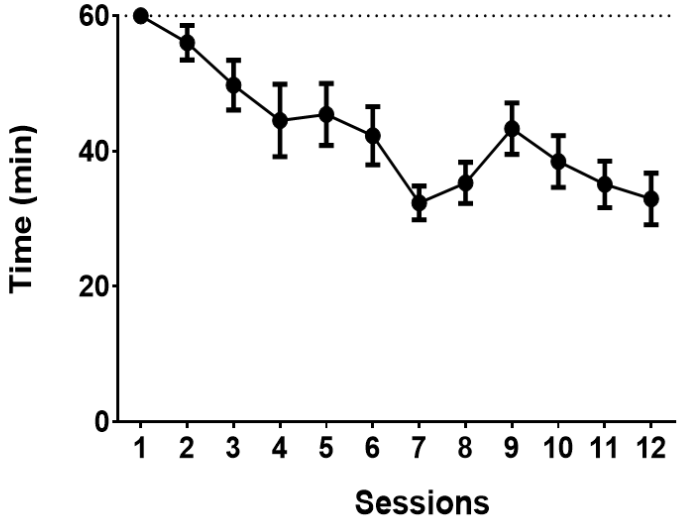
A


Number of trials completed
0
10
20
30
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
Sessions
Initial training - Session 1
Must touch - Session 2,3,4
Must initiate - Session 5, 6
Punishment of incorrect - Session 7 - 14
PVD - Session 15 - 26

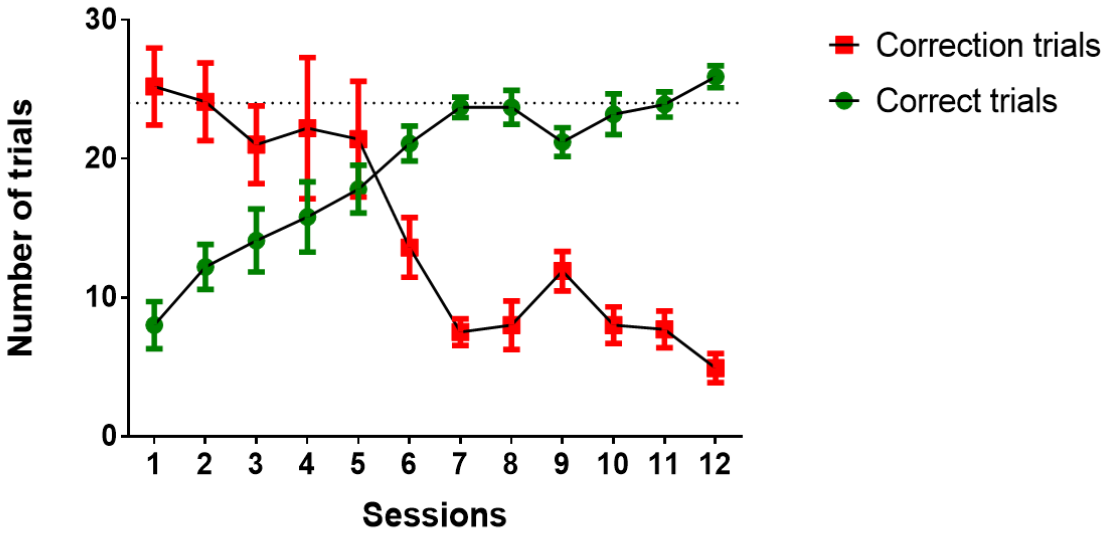
B


Correct trials (%)
0
10
20
30
40
50
60
70
80
90
100
1 2 3 4 5 6 7 8 9 10 11 12
Sessions

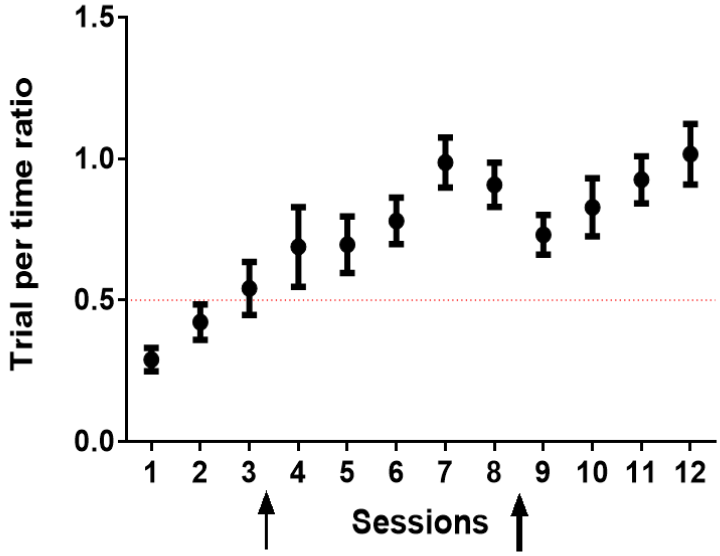
C


Time (min)
0
20
40
60
1 2 3 4 5 6 7 8 9 10 11 12
Sessions

D


Number of trials
30
20
10
0
1 2 3 4 5 6 7 8 9 10 11 12
Sessions
Correction trials
Correct trials

E


Trial per time ratio
1.5
1.0
0.5
0.0
1 2 3 4 5 6 7 8 9 10 11 12
Sessions

F

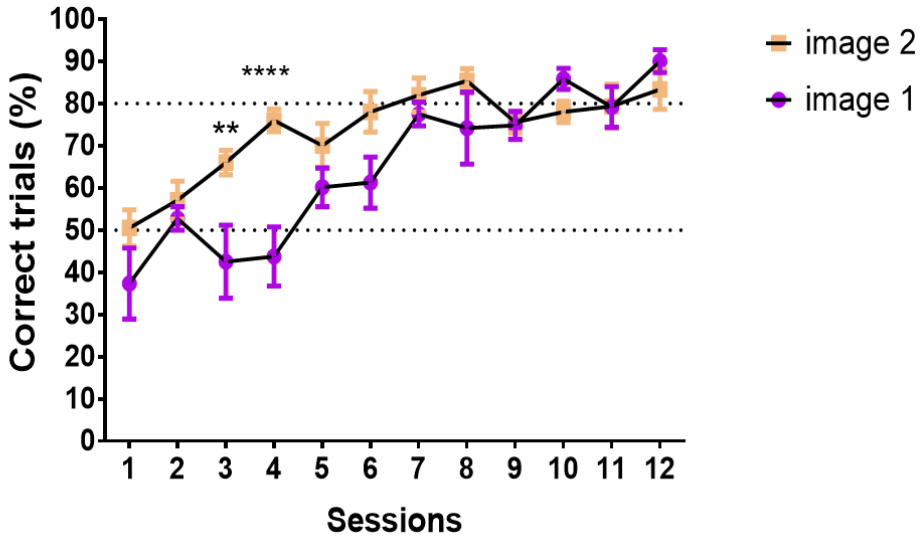

Correct trials (%)
100
90
80
70
60
50
40
30
20
10
0
****
**
1 2 3 4 5 6 7 8 9 10 11 12
Sessions
image 2
image 1

Figure 3.7 Pairwise discrimination Pilot N1 results. A. The number of trials completed by mice over all the session, throughout the whole assay (training phases and PD test phase included). Each session is equivalent to one day of training/testing, and are serially numbered. The session ended once the mouse either did 30 trials or 60 minutes elapsed (the time available). Mice were moved to the next phase once all the mice in the cohort completed the criteria for that phase. As the mice progressed through a phase (had more sessions doing the same task), their performance improved (more trials were completed in 60 minutes). As the mice started a new phase, the number of trials completed decreased in the first sessions for that phase compared to the last sessions of the previous training phase. *Initial training* phase took one session (Session N1), *Must touch* phase took three sessions (Session N2-4), *Must initiate* phase took two sessions (Session N5-6), *Punishment of incorrect* phase (Session N7-14), and pairwise discrimination (PD) testing phase (Session 15-26). In PD testing, there was no difference in trials completed between session 17 and 18 (two-day break between sessions) ($P > 0.9999$) and session 22 and 23 (three-day break in between sessions) ($P = 0.9138$). The dotted line at 30 indicates the maximum trials that can be done per session. **B**. Percentage of correct trials (when CS+ image was nose-poked) out of the total trials done per session. The dotted line represents the criteria of 80% correct trials (24 out of 30) **C.** The time (minutes) that it took for mice to complete each PD testing session. The maximum time (60 minutes) is shown with a dotted line. **D.** The number of correction trials (in red) and correct trials (in green) done in each PD session. As the number of correct trials increase, the number of correction trials (trials that follow nose-poking CS- image) decrease. The maximum number of non-correction trials is 30, whereas there is no limit on the number of correction trials. The dotted line represents the criteria of 24 correct trials out of 30 (80% correct). **E**. Trial per time ratio plotted per session, with anything below 0.5 implying the criteria of 30 trials in 60 minutes was not reached. The mice had to get above 0.5, as well as get 80% of the 30 trials correct to complete the task successfully. The 0.5 cut-off is annotated with a red line. There was difference in trials per time ratio between session 3 and 4 (two-day break between these sessions; thin arrow) ($P = 0.7958$) and session 8 and 9 (three-day break in between those sessions; thick arrow) ($P = 0.3464$). **F.** The percentage of correct trials done per session was plotted separately for the two CS+ groups (depending on which image was assigned as CS+: image 1 (in purple) or image 2 (in peach)). On the first session, there was no difference in the percentage of correct trials per CS+ group, and both groups had around 50% chance (dotted line). In session 3 and 4, the group with image 2 as CS+ had a significantly higher percentage of trials done correctly (session 3: $P \leq 0.001$***; session 4: $P \leq 0.0001$****). The second dotted line represents the criteria of 80% correct trials. Values are plotted as mean ± SEM.

3.2.4.3 Pilot Number 2

Based on results from Pilot 1, I further optimised the protocol to address several limitations:

1. The mice tested in Pilot N1 were C57BL/6J background, whereas all mutant mice to be tested (Chapter 4 and 5) are of C57BL/6N background. I therefore tested C57BL/6N next.

2. Even though group advancement of the mice is easier to coordinate and analyse, due to the variability in the completion of each phase, there is a risk of over-training some of the mice. One proposed way of eliminating this is resting those animals that completed criteria and running the entire cohort with a refresher session to re-baseline the mice to the same level, once the whole group has completed criteria (Horner et al., 2013). However, another possible regime of training is advancing all the mice individually. This avoids overtraining but does not allow the group to be synchronised. I applied that latter option in my second pilot.

3. After a personal communication with Dr Stacey Rizzo, Jackson Laboratory, I added two further changes to the protocol. Firstly, mice were now singly housed for the duration of the assay. This allowed for better control of their food intake, as well as reduced any confounding effects of hierarchical structures in the home cages on their behaviour and task performance. Secondly, I ensured that each mouse had an allocated touchscreen chamber, and the mouse consistently was tested in this chamber throughout all the sessions. This was to eliminate the effect of any possible differences there may be between chambers.

4. Due to a potential image bias observed in Pilot N1, I changed image set to the plane and spider image set (Chapter 2, section 2.3.6, Fig.2.8D).

The mice (n=12) advanced through the training phases in a comparable manner to the first pilot (Fig.3.8).

During PD testing, the mice (n=10) initially got 50% of trials correct on session 1 (which indicates no preference for either of the images) and increased their performance by increasing the percentage of correct trials done (Fig.3.9A,C), while decreasing the time taken to complete each session (Fig.3.9B), reaching criteria in an average of 14.6 sessions (Fig.3.8, Fig.3.9A). Based on the average number of days (one day = one session) taken for mice to reach criteria in the two pilots (Pilot 1 = 12 days, Pilot 2 = 15 days) and existing data from other studies (personal communication with Dr Stacey Rizzo, Jackson Laboratory; Copping et al., 2017; Yang et al., 2015), I set 25 days of PD task as the maximum cut-off point for future studies. Two mice had to be excluded from Pilot 2 analysis because they did not reach criteria after 25 sessions.

Because each mouse progressed individually to the next phase and there was no group synchrony, once each mouse reached PD criteria, the average percentage of correct trials of the two criteria days was calculated. This was then plotted alongside mice that have yet not completed the trials until all mice reached criteria or 25 sessions (Fig.3.9A).

To assess whether there was an image bias when using the new image pair, I analysed the two image CS+ groups separately. There was no significant difference in the performance (% correct trials) between mice that were assigned image 1 as CS+ (n=6) and image 2 as CS+ (n=4) (significant interaction: 0.0392, image difference: $F(1, 8) = 2.245$, $P= 0.1725$, two-way ANOVA; no significant difference between images post hoc analysis after a two-way ANOVA), indicating a lack of image bias when spider and plane images were used (Fig.3.9D). The spider and plane images were therefore used subsequently as a default (Chapter 2, section 2.3.6, Fig.2.8D).

To confirm the results are due to learning, after 25 sessions, the CS+ and CS- image values were swapped. Five further sessions were run in this manner (n=9), and as expected the mice dropped below 50% correct trials in the first session (session 3 in Fig.3.10), reaching the 50% correct at the fifth session (Session 7 in Fig.3.10).

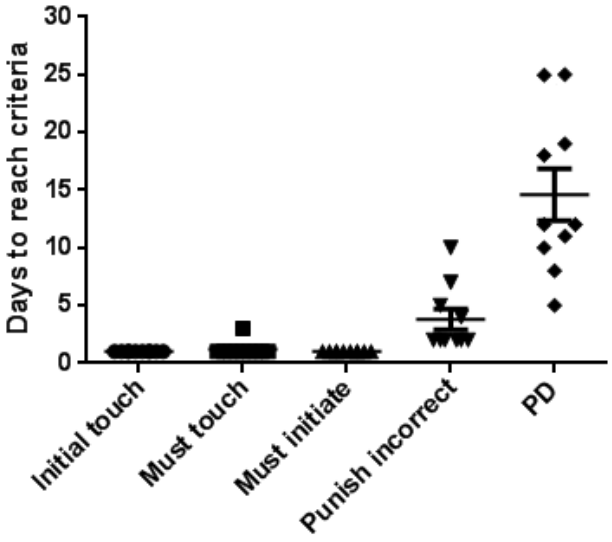
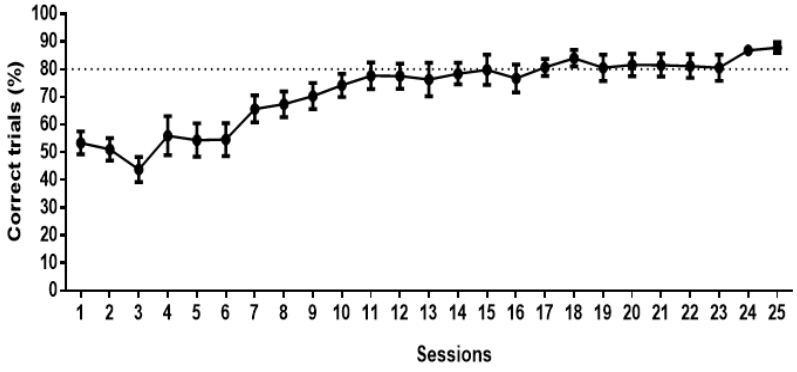
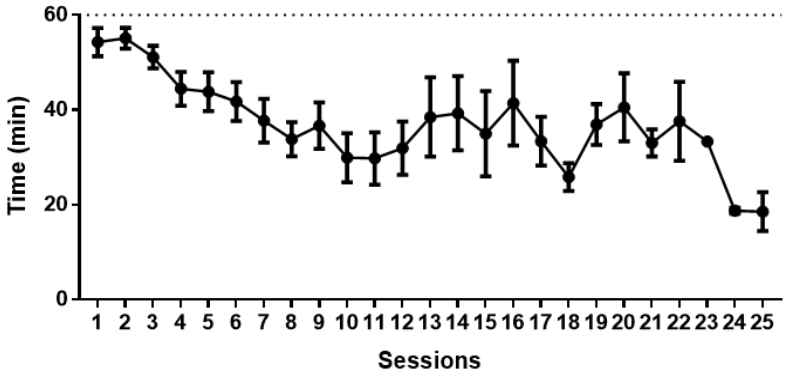


Figure 3.8 Days needed to complete each pairwise discrimination phase. Days (which were equivalent to sessions, because mice had a session a day) needed to reach criteria was plotted for all training phases (average number of days: Initial touch = 1, Must touch =1.4, Must initiate=1, Punish incorrect =4.6) and pairwise discrimination testing phase (PD = 14.6). Criteria for all training phases from Initial touch to Must initiate was 30 trials in 60 min. Criteria for Punish incorrect is 30 trials in 60 min, 24 correct, for two consecutive days. PD requires for the mouse to complete 30 trials in 60 minutes, 80% of trials correct, for two consecutive days. Values are plotted as mean ± SEM.

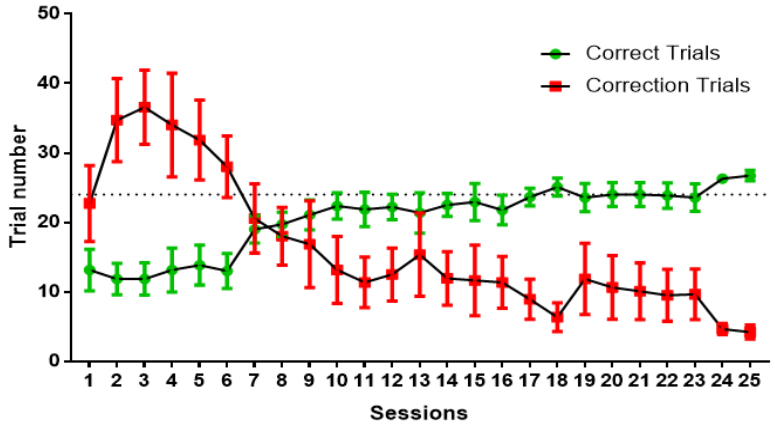
A


Correct trials (%)
Sessions

B


Time (min)
Sessions

C


Correct Trials
Correction Trials
Trial number
Sessions

D

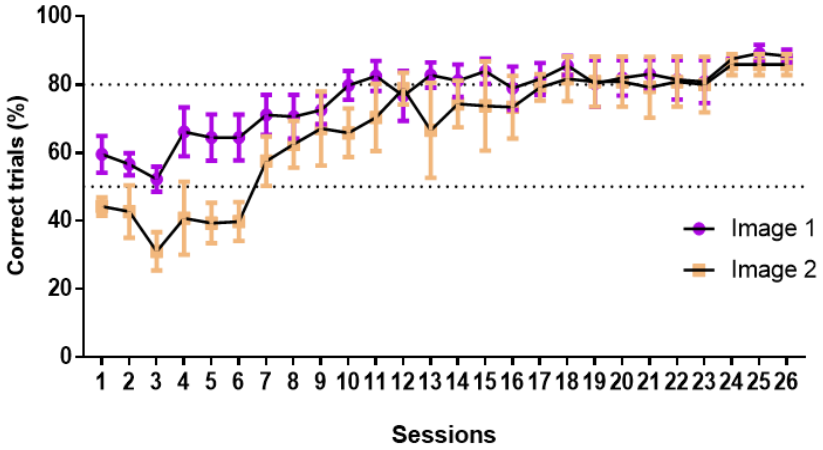


Image 1
Image 2
Correct trials (%)
Sessions

Figure 3.9 Pairwise discrimination Pilot N2 results. A. The percentage of correct trials (when CS+ image was nose-poked) out of the total trials done per session, plotted for all the sessions (1-25). The dotted line represents the criteria of 80% correct trials. **B.** The time (minutes) that it took for mice to complete each PD testing session. The maximum time (60 minutes) is shown with a dotted line. **C.** The number of correction trials (in red) and correct trials (in green) done in each PD session. As the number of correct trials increase, the number of correction trials (trials that follow nose-poking CS- image) decrease. The maximum number of non-correction trials is 30, whereas there is no limit on the number of correction trials. The dotted line represents the criteria of 24 correct trials out of 30 (80% correct). The percentage of trials done correctly, over total number of trials in 60 minutes was plotted over sessions. **D**. The percentage of trials done correctly out of total completed trials, divided into groups based on which image was CS+ (image 1 (in purple) or image 2 (in peach)). There was no significant between the performance (% correct trials done) in group 1, assigned image 1 as CS+, and group 2, assigned image 2 as CS+ (P=0.1725). 50% (chance) and 80% correct trials (criteria) are represented with dotted lined. Values are plotted as mean ± SEM.

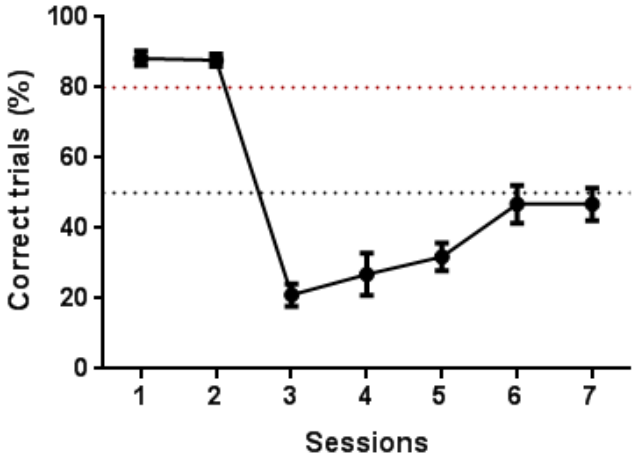


Figure 3.10 Performance when CS+ and CS- images were reversed. Performance (n=9) in the last two sessions of PD task (Session N1-2) are shown (same as the last two sessions in Fig.3.9A), followed by Session N3-7, in which the images are swapped (CS- image becomes CS+ and vice versa). The percentage of correct trials drops on Session N3 accordingly (below 80% correct trials criteria annotated with a red dotted line) and slowly increases with more sessions, reaching just under 50% (black dotted line; representing chance) on the fifth session (Session N7). Values are plotted as mean ± SEM.

3.3 Discussion

3.3.1 Object discrimination paradigm

The major challenge when optimising both novel object recognition (NOR) and object displacement (OD) assays was the choice of objects. It took several pilots to establish the best performing object set for both of the assays. Several key points emerged as the result of piloting. Firstly, mice have inherent biases towards investigating man-made objects, which is why inherent preference must be assessed before testing to avoid false positives and negatives. Dividing the cohort so that the objects are counterbalanced between animals (alternating which object is the familiar one) partially controls for these potential biases, but significant differences in innate preference can mask potential phenotypes. Secondly, the chosen objects have to be sufficiently different from one another for the mice to be able to distinguish between them. Thirdly, even if there is no detectable inherent preference for either of the objects, this does not necessarily imply the mice are able to distinguish between the objects. For example, even though I did not detect an inherent preference for either of the initial objects used (Cola cola can and water bottle) in NOR, the mice did not show a preference for the unfamiliar object after training, indicating they were unable to distinguish between the objects. The mice successfully discriminated between the final set of objects, small glass vial and light bulb piloted in the object discrimination paradigm, resulting in the subsequent use of this object set (Chapter 4).

Due to the similarity in set-up, I used OD not only as an independent spatial test, but also as the training phase for NOR. Importantly, the performance in NOR was not reliant on performance in OD, as both training and testing in OD served as memory acquisition training for NOR irrespective of whether the mice remembered the new location of the object. This combined paradigm, therefore, enables independent detection of deficits in spatial and non-spatial components of memory in an efficient manner, both in terms of utilizing only one set of objects for both tests and in terms of time efficiency. In order to use both assays using this paradigm, the OD interval time between training and testing was shortened to 1-2h, which is a suggested optimal time interval by Murai et al. (2006), while the interval time

of NOR was extended to 24h, which has been used in other studies (Fernandez and Garner, 2007). This design enabled testing shorter (1h) vs longer (24h) term memory.

The mice did not have an object bias for either of the objects used (vial and light bulb) and were able to distinguish both between moved and stationary (in OD phase), by spending an average of 61% of total investigation time with the moved object, and between novel and familiar objects (in NOR phase), by spending on average 67% of total investigation time with the novel object. Of note, the pre-determined exclusion criterion (of <10 seconds of total investigation of objects in a 10-min assay) ensures that the animals analysed have explored both objects sufficiently during the acquisition phases. Moreover, removing the mice that are not motivated to explore the objects during testing phase reduces inter-individual variability and increases the accuracy of the results (Arqué et al., 2008; Leger et al., 2013).

3.3.2 Pairwise discrimination

Touchscreen technologies are an attractive system to test cognitive function in an automated manner that parallels the way patients are tested with CANTAB technology (Bussey et al*.,* 2008; Luciana et al., 1997; Robbins et al., 2010). Moreover, the technology offers a wide range of available cognitive tasks, allowing testing specific brain regions and cognitive functions (Horner et al., 2013; Oomen et al., 2013). To date, touchscreen technology has not been used at the Sanger Institute. I, therefore, aimed to establish one of the touchscreen assays, pairwise discrimination (PD). This assay tests for associative perceptual learning and memory. Due to time constraints reversal of the CS+ and CS- images served only as a confirmation of associative learning in the PD task, and not as a reversal learning task. Reversal learning tests for cognitive and behavioural flexibility, and is associated with different brain regions (namely subregions of the PFC) than PD task (which is associated primarily with the dorsal striatum) (Chudasama et al., 2001; Chudasama and Robbins, 2005; Horner et al., 2013; Miller, 2000). It, therefore, would test different domains of cognitive function and should be considered on hypothesis-driven bases for specific mouse lines, as an additional assay.

I completed two pilots of pairwise discrimination (PD) task, using the results of the first pilot to change several parameters in the second one, further optimising the assay. In Pilot N2, the mice were individually advanced to the next phase in order to avoid overtraining. Mice with C57BL/6N background were used instead of C57BL/6J mice (Pilot N1), in order to test mice of the same background as mice under investigation in Chapter 4 and 5. This was important because of reported differences in performance in touchscreen learning based on genetic background (Graybeal et al., 2014).

In order to keep better control of the food intake, maintaining full appetitive motivation, and avoid the confounding effects of hierarchical behaviours in home cages, for the second pilot I singly housed the mice throughout the course of the test rather than group housed them. Housing environment has been shown to affect rodent behaviour, with the reported increase in locomotor activity in singly housed mice, task- and strain-specific effects on stress levels, and detectable memory impairment in some cognitive tasks (Võikar et al., 2004). Although touchscreen studies often use group housed animals, from personal correspondence with Dr Stacey Rizzo from Jackson Laboratory, who routinely uses singly housed animals, there are no observed deficits in the pairwise discrimination in singly housed wildtypes. Although a direct comparison cannot be made between the performances of the mice in the two pilots I have carried out, due to variability testing methodology, both sets of wildtypes completed training within the expected timeline, suggesting that housing conditions are not a critical factor for memory acquisition and learning in this task.

My motivations for piloting pairwise discrimination (PD) assay were two-fold. Firstly, I wanted to assess whether PD could be used as part of a cognitive phenotyping screen for characterising a diversity of novel neurodevelopmental disorders (described further in Chapter 5). This phenotyping platform was designed to serve as a proof of principle and refinement platform of a large 5-year project at the Sanger Institute, aimed at characterising a diversity of novel neurodevelopmental disorders arising from mutations identified in the Decipher Developmental Disorders and Prenatal Assessment of Genomes and Exomes projects. In order for PD task to be included as part of this phenotyping screen, it had to fit into the time frames of the screen, which posed a major limitation, and have the sensitivity

for detecting subtle cognitive impairments. In both instances, the assay did not meet the relevant criteria to be included in the phenotyping screen (Chapter 5).

The overall time taken for the mice to progress through the assay (training and testing) was over a month, which did not fit with the timelines of the phenotyping screen of four weeks of testing per mouse line (discussed in Chapter 5). The time taken to reach the learning criteria in the pilots was similar to, although on the upper end of, other reported the wildtype performance of 10-16 days in PD. Of note, the criteria for PD varies between 70-85% correct trials between studies which may explain the shorter time frames in some of them (Brigman et al, 2008; Morton et al., 2006; Romberg et al., 2013; Yang et al., 2015). Due to weekend breaks in testing, it was critical to ensure these did not introduce an additional confounding effect on the performance of mice in the PD testing phase. The set of rules for the advancement of mice to the next phase after a weekend break increased however the absolute time of the assay, even though this was not reflected in the analysis of the time taken to reach criteria. Nevertheless, the stringency of the pre-set rules was especially important if the assay was to be used as part of the larger screen testing multiple mouse lines.

In order to detect a significant difference in days (effect=6 days) to reach criteria, using Pilot 2 wildtype data (Chapter 2, section 2.3.6), with 80% power and 5% type I error rate, the sample size would need to be n=22. For a smaller cohort of n=13, we would only have enough power (80% power, 5% type I error rate) to detect a larger difference (effect=8 days); therefore lacking the sensitivity to detect subtle differences with a smaller cohort size.

The second motivation for piloting PD assay was to include a hippocampus-independent test in the repertoire of available behavioural assays in our team. This was particularly relevant in assessing the hippocampal involvement in the loss-of-function *Kptn* mouse model in Chapter 4, where larger impairment effects were expected.

3.3.3 Concluding remarks

Overall, the establishment and optimisation of the assays in this Chapter provide an important addition to the repertoire of already existing behavioural and cognitive assays in our team, to characterize a variety of mouse models (described in Chapter 4 and Chapter 5). The newly established assays provided opportunities to characterize the mice with a multi-layered approach, ensuring a range of overlapping cognitive domains were tested. For example, object displacement, a spatial hippocampus-dependent task, can be used to confirm results from the Barnes maze. Novel object recognition (NOR) task complements both pairwise discrimination and social recognition tasks for different reasons. On the one hand, NOR task assesses the memory of the mouse for the familiar object without the confounders of social interaction and olfactory learning. On the other hand, NOR is a non-spatial task, which complements pairwise discrimination (PD) that explored non-hippocampus dependent memory deficits, providing further information on brain regions associated with the neurodevelopmental condition under investigation.

Chapter 4. *KPTN*

4.1 Introduction

4.1.1 *KPTN*-related syndrome

Windows of Hope (WOH) is a large population-based medical project set up to identify genetic causes of inherited conditions, including distinct uncharacterised developmental disabilities, focusing on the Anabaptist communities in the USA. Due to the nature of these communities, most mutations identified through this project are recessive and consanguineously inherited. One finding from WOH was a group of nine patients, six males and three females, belonging to four nuclear families, with a distinct inherited undiagnosed developmental delay syndrome (Fig.4.1) (Baple *et al.,* 2014). The most consistent phenotypes found in the patients include global developmental delay of variable severity, macrocephaly with frontal bossing, anxiety, stereotypies and repetitive speech. Additional features found in a subset of the patients include hyperactivity, recurrent pneumonia, splenomegaly, childhood hypotonia, and primary seizures.

Macrocephaly refers to an abnormally large head, including scalp, cranial bone and intracranial contents and is diagnosed with occipital frontal circumference (OFC) 2 standard deviations (SD) above the mean, with severe macrocephaly defined as OFC 3+ SD above the mean (Rollins et al., 2010). Macrocephaly can be present from birth or can be progressive. Absolute macrocephaly, often referred to simply as macrocephaly, is measured without taking the overall stature into account, while relative macrocephaly refers to OFC within 2 SD, but measurements of stature above 2 SD (Rollins et al., 2010). Macrocephaly may be due to megalencephaly (enlargement of the brain parenchyma) or due to other conditions, such as cranial hyperostosis (thickening of the skull) and hydrocephalus (accumulation of fluid in the brain) (Williams et al., 2008). In this thesis, the term macrocephaly will be used to include megalencephaly.

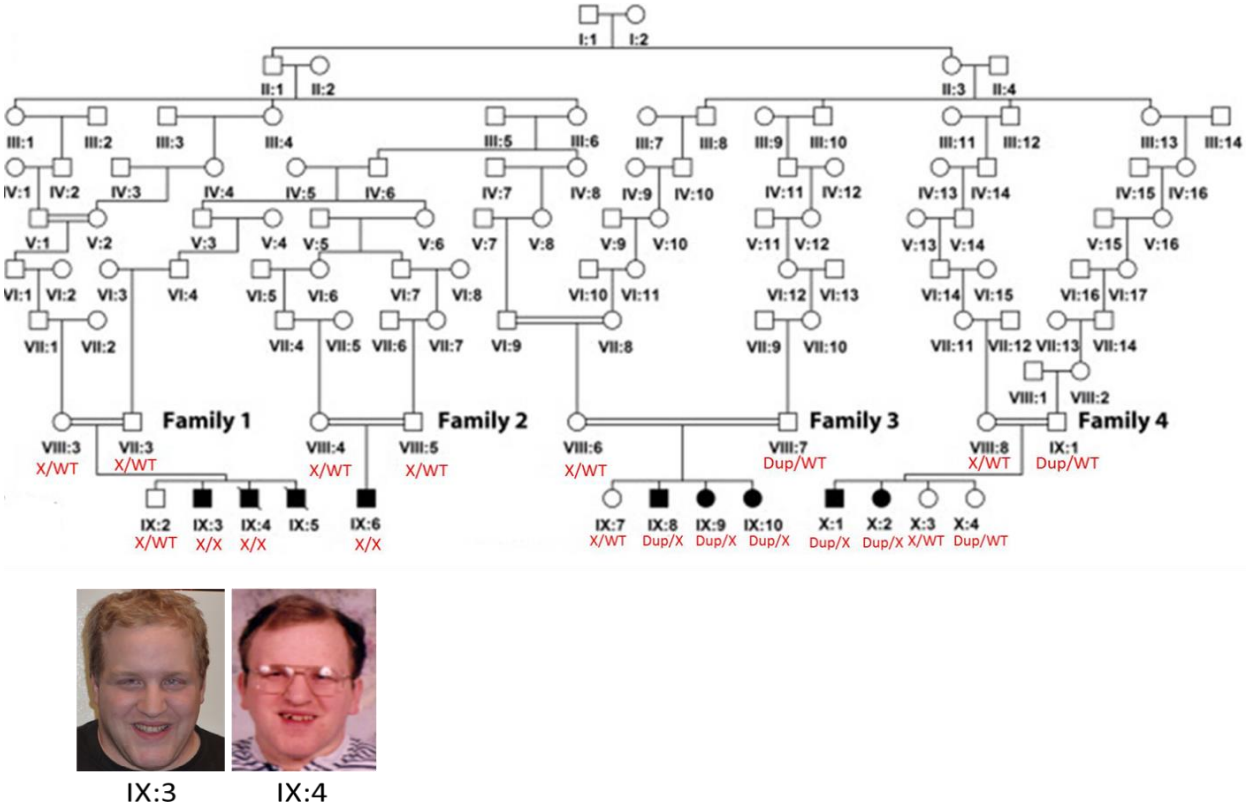


Figure 4.1 Amish family pedigree. Pedigree diagram showing the four nuclear families and the nine affected patients (in black), with circles representing females and squares representing males. The segregation of the two founder mutations is annotated as follows, c.776C>A variant denoted by X and c.714_731dup variant is denoted by Dup. Six out of nine probands are compound heterozygous for both mutations. Bottom left corner – two sibling patients with homozygous X/X mutation. Figure adapted from Baple et al., 2014.

Exome sequencing in the nine probands revealed two founder loss-of-function recessive mutations in exon 8 of *KPTN* gene on chromosome 19 (Fig.4.2). One of the mutations, a nonsense c.776C>A sequence variant (annotated as X, Fig.4.1), causes a premature stop codon. The other mutation is an in-frame 18 bp duplication, c.714-731dup variant (annotated as Dup, Fig.4.1), producing either mislocalised or nonfunctional protein products (Fig.4.2) (Baple *et al.,* 2014). Six out of the nine patients were compound heterozygous for both of the mutations (Fig.4.1). The patients who were homozygous (*X/X*) had a more

severe intellectual impairment and a higher incidence of seizures than the compound heterozygous patients. The authors propose that this may be due to Dup mutation retaining partial *in vivo* functionality (Baple *et al.,* 2014).

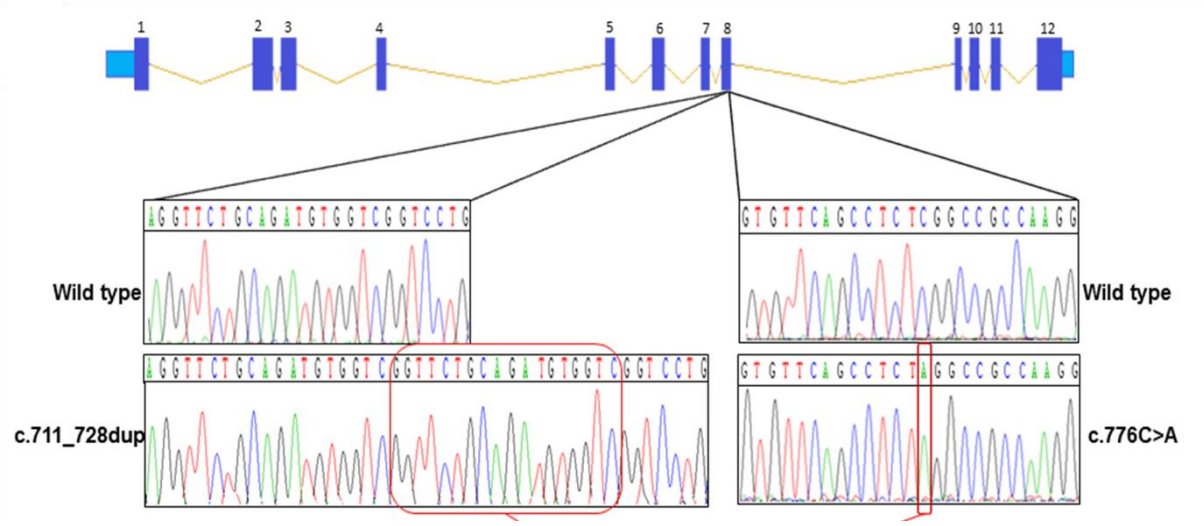


Figure 4.2 *KPTN* gene organisation and position of the two mutations in exon 8 (c.776C>A and c.714-731dup). c.776C>A variant is a nonsense mutation where cytosine is substituted for adenine. c.714-731dup is an in-frame 18 bp duplication mutation. Figure adapted from Baple *et al.,* 2014

Subsequently, two siblings from Estonia were identified with characteristics associated with the *KPTN*-related syndrome and one-nucleotide homozygous loss-of-function duplication in exon 7 of *KPTN* gene (Fig.4.3) (Pajusalu et al., 2015). Both siblings have moderate intellectual disability, progressive macrocephaly (OFC+ 4-5 SD), and anxiety. Out of the two siblings, the brother had more severe behavioural abnormalities and experienced two episodes of generalised seizures, whereas the sister has no remarkable behavioural abnormalities apart from anxiety, and never had seizures (Pajusalu et al., 2015).

Figure 4.3 Appearance of the Estonian siblings with one-nucleotide duplication in exon 7 of *KPTN*. (Pajusalu et al., 2015)

4.1.2 KPTN and actin cytoskeleton

KPTN, also known as 2E4, is an actin-binding protein that is not well characterised. Initially isolated from blood platelets, with a possible role in actin dynamics of platelet activation, it was also found to be localised in sensory epithelium of embryonic inner ear with a suggested role in stereocilia formation (Bearer & Abraham, 1999; Bearer *et al.*, 2000; Hong *et al*., 2004). Moreover, it mapped to DFNA4 nonsyndromic deafness locus and was therefore postulated to be a hearing loss candidate gene (Bearer & Abraham, 1999; Bearer *et al.*, 2000; Chen et al., 1995; Hong *et al*., 2004). However, no mutations were subsequently detected in the coding region of the gene in deaf individuals (Zong *et al.,* 2012). Consistent with this, no hearing deficits were identified in the Amish or the Estonian *KPTN*-syndrome cohorts.

Actin is an abundant highly conserved protein that polymerizes into filaments and is essential for many cellular properties and functions such as cellular motility, the structure and mechanical properties of the cytoplasmic matrix, ion channel activity in the plasma membrane, and for localizing neurotransmitter receptors (Allison et al., 1998; Kneussel & Betz, 2000; Lambrechts et al., 2004; Maximov et al., 1997; Polard and Cooper, 1986; Winder and Ayscough, 2005). The actin cytoskeleton plays an important role in neuron migration and axonal projection during brain development, leading to the formation of highly complex neuronal networks necessary for higher cognitive brain functions (Rivière *et al*, 2012). Cell motility and migration is crucial both during development and throughout the lifetime of the organism and must be tightly controlled (Lambrechts et al., 2004; Kessels et al., 2011). Actin-binding proteins are responsible for regulating the dynamic behaviour of the actin cytoskeleton, and mutations in genes encoding these may, therefore, lead to diseases (Lambrechts et al., 2004). It has been shown that the modulation of the actin system through actin-binding proteins is impaired in the developing brain of individuals with Down's Syndrome, and depolymerization of dynamic actin filaments can affect generation of memory in the hippocampus (Krucker et al., 2000; Weitzdoefer et al., 2002).

Baple *et al*. demonstrated that KPTN is associated with dynamic actin cytoskeletal structures of neuronal cells necessary for dendritic arborization or spine formation (Baple *et al.,* 2014). It was found to be enriched in neuronal growth cones at early developmental stages and at postsynapses of neurons during synaptogenesis, suggesting a role for KPTN in neuromorphogenesis (Baple *et al.,* 2014). Moreover, *KPTN* is expressed throughout the body and seems to be present in all cell types in the brain (Fig.4.4) (Uhlén et al., 2015; Zhang et al., 2014).

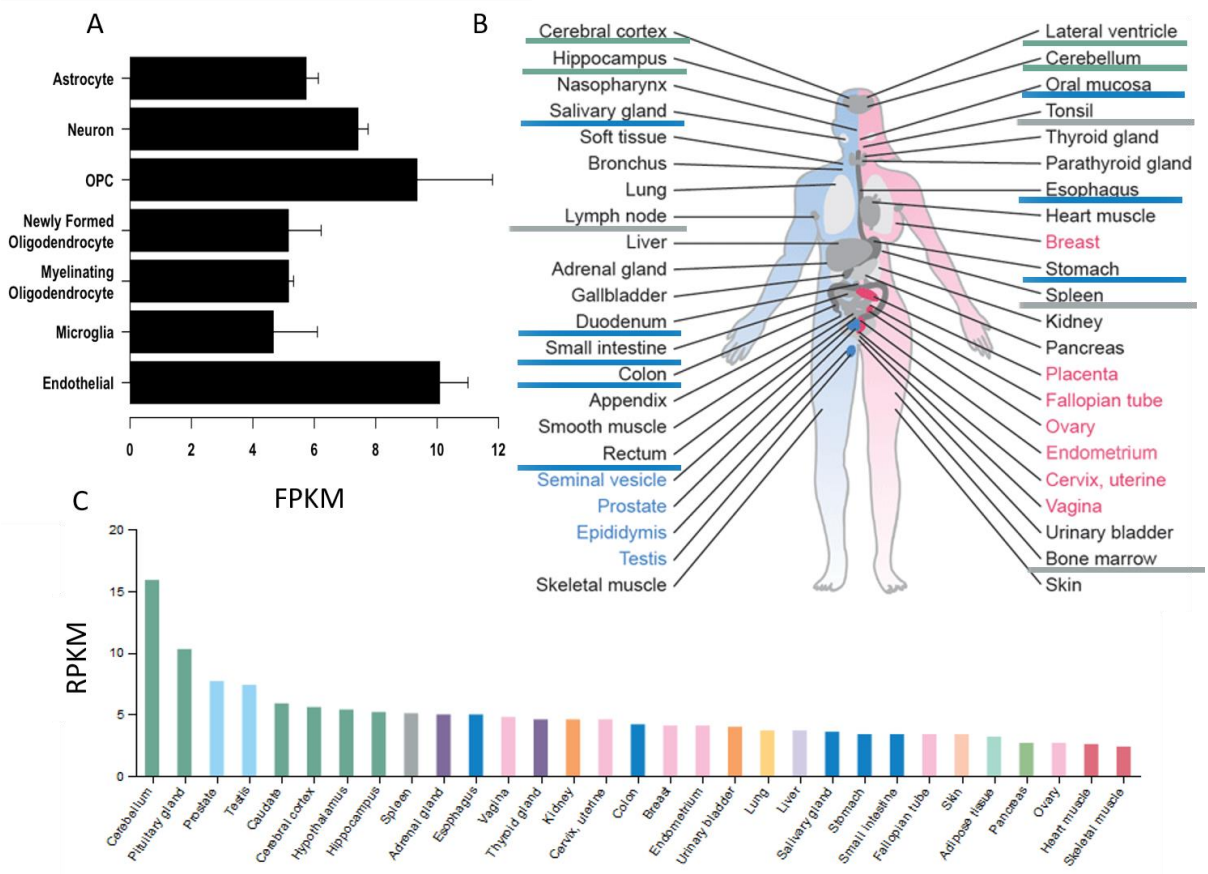


Figure 4.4 *KPTN* RNA-Seq transcriptome. A. Relative *KPTN* expression, as FPKM (fragments per kilobase of transcript per million mapped reads), in glia, neurons, and vascular cells of the cerebral cortex. **B**. Diagram of the human tissue atlas, with some tissue groups, underlined with the same colour scheme as used for relative expression of *KPTN* in these tissues. **C.** The panel of tissues and their relative *KPTN* expression, as RPKM (reads per kilobase of transcript per Million mapped reads). Colour scheme: Brain (green), Bone marrow and immune system (grey), endocrine (purple), muscle tissue (pale red), lung (yellow), liver and gallbladder (lilac), pancreas (light green), gastrointestinal tract (dark blue), kidney and urinary bladder (orange), male tissues (light blue), female tissues (pink), skin (peach), adipose and soft tissue (turquoise). Figure adapted from The Human Gene Database, The Human Protein Atlas, and Zhang et al., 2014.

4.1.3 KPTN and mTOR signalling

More recently KPTN has been implicated as an upstream regulator of the mTORC1 signalling pathway (Wolfson et al., 2017). mTOR, (formerly 'mammalian') mechanistic target of rapamycin, is a serine/threonine protein kinase that forms the catalytic subunit of two distinct protein complexes, known as mTOR Complex 1 (mTORC1) and 2 (mTORC2) (Crino, 2011; Laplante & Sabatini, 2012; Saxton & Sabatini, 2017; Wolfson et al., 2017). mTORC1 comprises three core components: mTOR, Raptor (regulatory protein associated with mTOR), and mLST8 (mammalian lethal with Sec13 protein 8) (Fig.4.5A). Raptor enables the recruitment of substrates to the complex and is necessary for the correct subcellular localization of mTORC1, while mLST8 associates with the catalytic domain and may play a role in stabilizing the kinase activation loop although its importance is still not fully understood (Frey et al., 2014; Hoeffer and Klann, 2010; Saxton & Sabatini, 2017). The complex also contains regulatory subunits PRAS40 (proline-rich Akt substrate of 40 kDa) and DEPTOR (DEP domain containing mTOR interacting protein) that inhibit Raptor (Fig.4.5A) (Peterson et al., 2009; Wang et al., 2007). The mTORC2 contains mTOR, mLST8 and DEPTOR, but instead of Raptor has an unrelated protein Rictor (rapamycin-insensitive companion of mTOR) with an analogous function, and has additional regulatory subunits Protor1/2 (Protein observed with rictor-1 or 2) and mSin1 (mammalian stress-activated protein kinase interacting protein 1) (Fig.4.5B) (Frias et al., 2006; Guertin et al., 2006; Lamming et al., 2012). Rapamycin is a potent antagonist of mTOR, and rapamycin-FKBP12 complexes directly bind and inhibit mTORC1, while mTORC2 is insensitive to its acute exposure (Crino, 2011; Laplante & Sabatini, 2012; Saxton & Sabatini, 2017; Wolfson et al., 2017). However, despite this, prolonged rapamycin treatment does inhibit the mTORC2 signalling, by hindering the incorporation of mTOR into new mTORC2 complexes (Phung et al., 2006; Sarbassov et al., 2006).

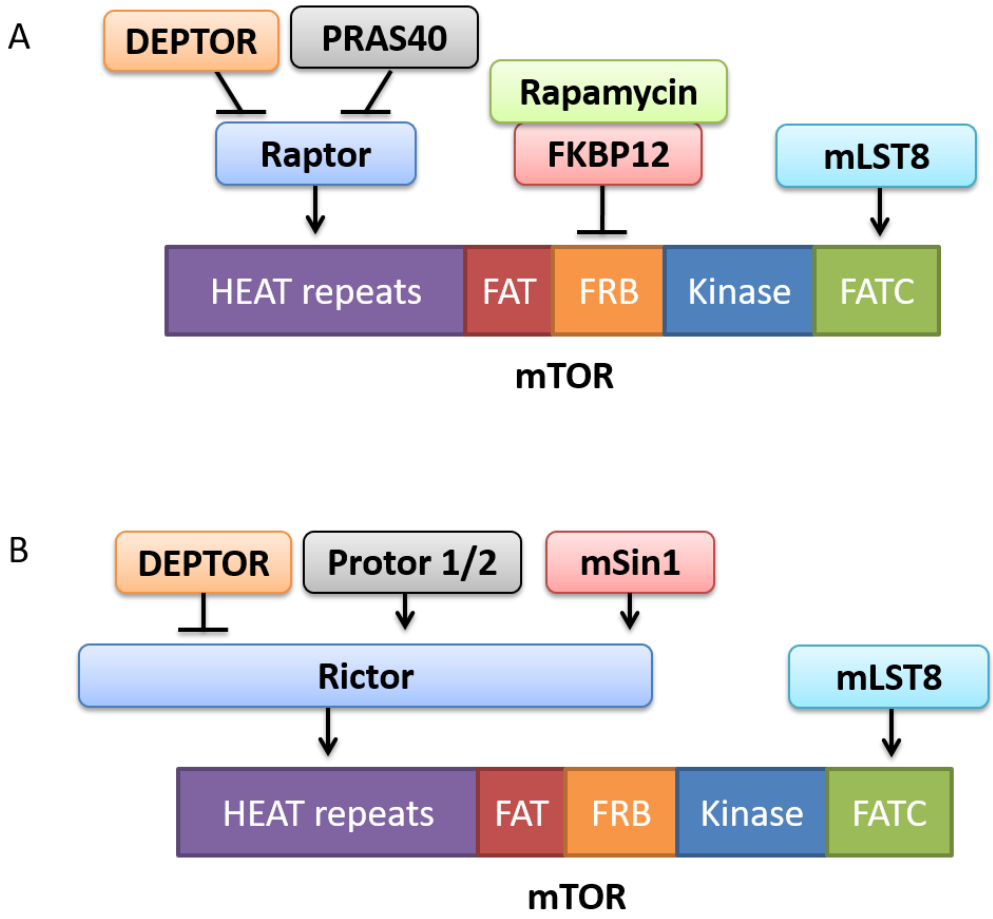


Figure 4.5 mTORC1 and mTORC2 subunits. A. mTORC1, comprising of mTOR, positive regulators Raptor and mLST8, as well as negative regulators DEPTOR, PRAS40, and FKBP12 that forms a complex with rapamycin. **B**. mTORC2, comprising of positive regulators Rictor, mLST8, Protor 1/2, and mSin1, as well as negative regulators DEPTOR. Figure adapted from Saxton & Sabatini, 2017.

mTOR signalling pathways are activated in response to environmental cues such as amino acid, oxygen, stress, and growth factor levels, as well as synaptic activity in the case of neurons (Fig.4.6) (Crino, 2011). Once activated, mTOR signalling regulates many important and distinct cellular processes (Fig.4.6).

mTORC1 signalling plays a central role in nutrient sensing, regulation cell cycle progression (through the regulation of protein, lipid, and nucleotide synthesis, as well as autophagy), and metabolism (by controlling the balance between anabolic and catabolic processes in response to environmental cues) (Fig.4.6) (Kim et al., 2002; Kim et al, 2008; Laplante &

Sabatini, 2012). One of the most characterised functions of mTORC1 is the regulation of translation through multiple processes, including regulation of ribosomal biogenesis, through proteins such as S6 kinase 1 and 2 (S6K1/2) and eIF4E-binding proteins (4E-BPs), and regulation of activity of phosphatases that in turn regulate mTOR substrates, leading to a mTOR-dependent feedback loop (Mayer and Grummt, 2006). Of note, although extensively linked to translational regulation, mTORC1 signalling has also been shown to regulate gene transcription through modulating the activity of transcription regulators, and potentially other processes such as affecting RNA stability and degradation directly or regulating epigenetic mechanisms (Passtoors et al., 2012; Wang et al., 2011; Yokogami et al., 2000).

mTORC2 signalling is important for cell survival and proliferation, ion transport, cell migration, metabolism, insulin/PI3K signalling, and cytoskeleton organisation (Fig.4.6) (Gu et al., 2011; Inoki et al., 2003; Jacinto et al., 2004 Yao et al., 2014). Therefore, not surprisingly, due to the pivotal roles of mTOR in many processes, aberrant mTOR signalling is associated with many diseases such as cancer, diabetes, neurodegenerative disorders, and neurological disorders such as epilepsy and autism (Crino, 2011; Saxton and Sebatini, 2017) (Fig.4.6). Proteins associated with both mTORC1 and mTORC2 pathway regulation, are expressed in progenitor cells of the ventricular zone during early brain development and in early neurons in the nascent cortical plate, and are important for regulating neuronal control of feeding, neural circuit formation, and overall brain development (Choi et al., 2008; Hentges et al., 2001; Kwon et al., 2006; Nie et al., 2010; Tavazoie et al., 2005; Thomanetz et al., 2013). Both mTORC1 and mTORC2 signalling play a role in these neurological processes (Lipton and Sahin, 2014; Tee et al., 2016). Moreover, mTOR signalling is central to the regulation of long-lasting synaptic plasticity, which is reliant on protein synthesis and is critical for the formation and storage of memories (Hoeffer and Klann, 2010).

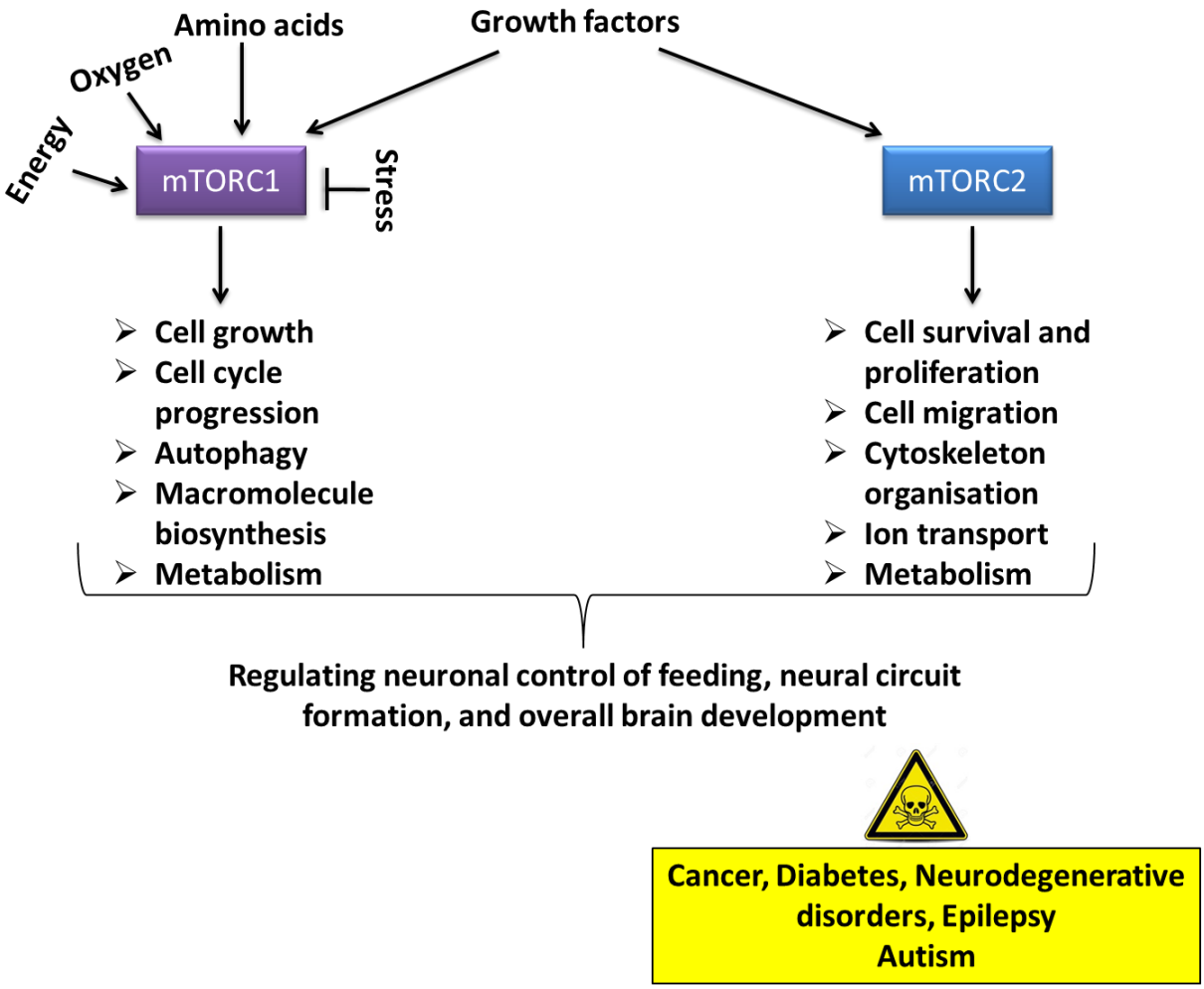


Figure 4.6 mTOR signalling and associated disorders. The upstream regulators of mTORC1 and mTORC2 pathways and the downstream pathways mTOR signalling regulates, as well as the roles mTOR signalling has in brain development. Bottom right (yellow box) – disorders associated with aberrant mTOR signalling.

Interestingly, KPTN has been recently shown to be a part of KICSTOR protein complex, consisting of KPTN, ITFG2, C12orf66 and SZT2 (seizure threshold 2) (Wolfson et al., 2017). KICSTOR is required for inhibition of mTORC1 signalling in response to amino acid or glucose deprivation (Saxton and Sebatini, 2017; Wolfson et al., 2017). The proposed mechanism through which KPTN works involves a heterodimer of KPTN and ITFG2. These then form a complex with the other two proteins (C12orf66 and SZT2) of the KICSTOR complex, with SZT2 being the link between the other three proteins (Fig.4.7). Amino acid levels, detected via intra-lysosomal and cytosolic sensing mechanism, activate Rag GTPases (obligate heterodimers) that are tethered to the lysosome membrane via a pentameric Ragulator complex. Rag GTPases bind Raptor and promote the translocation of mTORC1 to lysosome surface. Once at the surface mTORC1 acts via an 'AND-gate' mechanism - both Rag GTPases and Rheb (Ras-homolog expressed in brain) must be activated - to activate mTORC1 signalling (Fig.4.7). In response to amino acid or glucose deficiency, KICSTOR localizes to lysosome surface and recruits GATOR 1 (GTPase activating protein for RAGA), which inhibits Rag GTPases. Conversely, GATOR 2 inhibits GATOR 1 in response to cytosolic amino acid levels (Fig.4.7). Thus, KICSTOR acts as a negative regulator of the mTORC1 pathway. Complexes, such as TSC1-TSC2 (tuberous sclerosis complex) protein heterodimer, negatively regulate Rheb, integrating signals from cellular energy status and growth factor signalling (Fig.4.7).

Loss-of-function of *TSC1* or *TSC2* causes unregulated mTOR activation, which in turn leads to tuberous sclerosis complex disease, associated with epileptic seizures (Crino, 2011). Importantly, mutations in KICSTOR components have been shown to result in neurological diseases. Homozygous and compound heterozygous mutations in *SZT2* cause a distinct and severe early-onset autosomal-recessive epileptic encephalopathy, characterised by epilepsy, global developmental delay, brain abnormalities such as thick and short corpus callosum, and affected head size (Basel-Vanagaite et al., 2013; Tsuchida et al., 2017). In mice, loss-of-function *Szt2* mutants had epileptogenesis and an increase of mTORC1 signalling in the brain (Peng et al., 2017; Frankel et al., 2009). Mutations in *KPTN*, outlined above, further link the role of mTOR in neurological disorders.

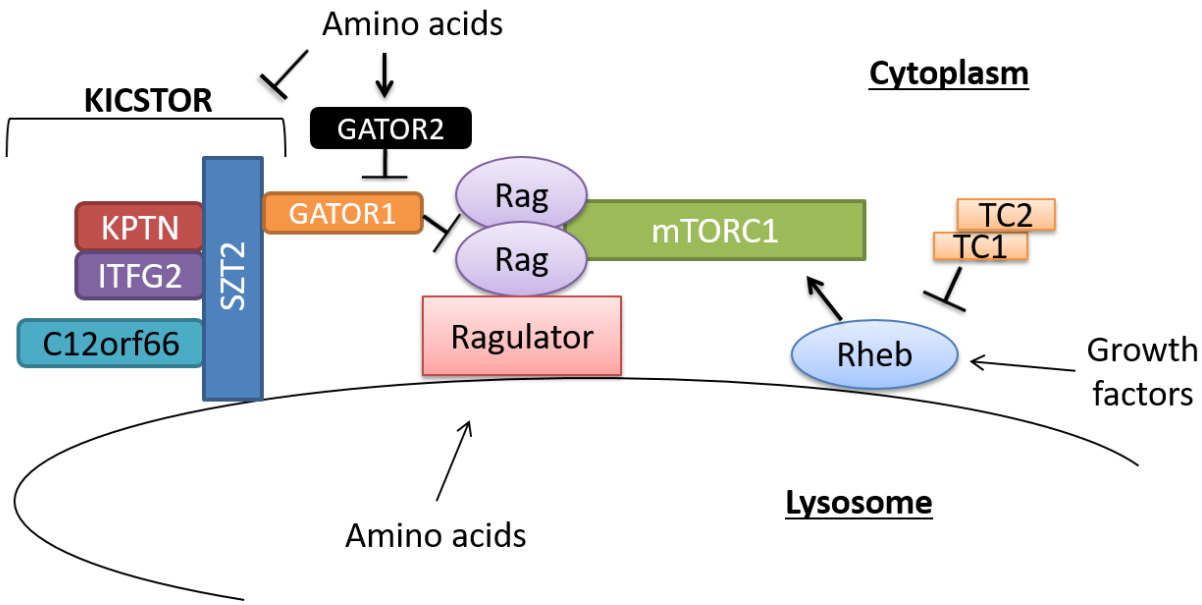


Figure 4.7 KICSTOR complex and mTORC1 signalling. Amino acid levels, detected via intra-lysosomal and cytosolic sensing mechanism, also activate Rag GTPases (obligate heterodimers). Rag GTPases are tethered to the lysosome membrane via a pentameric Ragulator complex. When activated Rag GTPases binds Raptor, which is part of the mTORC1, thus promoting the translocation of mTORC1 to lysosome surface, where it also interacts with Rheb (Ras-homolog expressed in brain). Both Rheb and Rag GTPases must be activated to activate the mTORC1 signalling pathways. KPTN forms a heterodimer with ITFG2, and these then form a complex with the other two proteins (C12orf66 and SZT2) of the KICSTOR complex, with SZT2 being the link between the other three proteins. In response to amino acid or glucose deficiency, KICSTOR localizes to lysosome surface, where it recruits GATOR 1 (GTPase activating protein for RAGA), which inhibits Rag GTPases. While GATOR 2 acts as a negative inhibitor of GATOR 1, in response to cytosolic amino acid levels. Thus KICSTOR acts as a negative regulator of the mTORC1 pathway. TSC1-TSC2 (tuberous sclerosis complex) protein heterodimer, negatively regulate Rheb. Figure adapted from Saxton & Sabatini, 2017.

4.1.3 The approach and aims

My aim was to model the novel *KPTN*-related syndrome, using a *Kptn* mutant mouse line, to demonstrate a causal link between the loss-of-function mutations and phenotype. Once a link was established, I aimed to use the validated animal model to study the mechanisms by which loss of KPTN protein results in the human developmental disorder. No study to date has modelled this syndrome in mice.

My approach was first to test whether I could recapitulate the main patient phenotypes in the homozygous loss-of-function *Kptn* mice, to confirm this is an appropriate model for the disorder. To this end, I employed a set of morphometric, cognitive, behavioural, and molecular tests. Secondly, to identify the potential mechanisms underpinning this neurodevelopmental disorder, I carried out both hypothesis-driven developmental work, as well as genome-wide expression profiling in several brain regions.

Because the most consistent phenotypes found in patients include macrocephaly, global developmental, and anxiety, I focused on testing these features first.

4.2 Results

4.2.1 *Kptn*$^{-/-}$ mouse model

To study the syndrome associated with the loss-of-function mutations in *KPTN* in greater detail, I used an engineered loss-of-function *Kptn*$^{tm1a(EUCOMM)Wtsi}$ "knockout first" model, generated by the Wellcome Trust Sanger Institute's Mouse Genetics Project (MGP) (White *et al.,* 2013). To assess the extent of transcriptional knockdown in the *Kptn* mutant line, the gene expression in three tissues - brown adipose, white adipose, and liver tissues - was assessed in 4 *Kptn*$^{tm1a/tm1a}$ and 4 *Kptn*$^{+/+}$ mice (Fig.4.8). The results indicate the *tm1a* allele has complete loss of expression for *Kptn* and is henceforth referred to as *Kptn*$^{-/-}$.

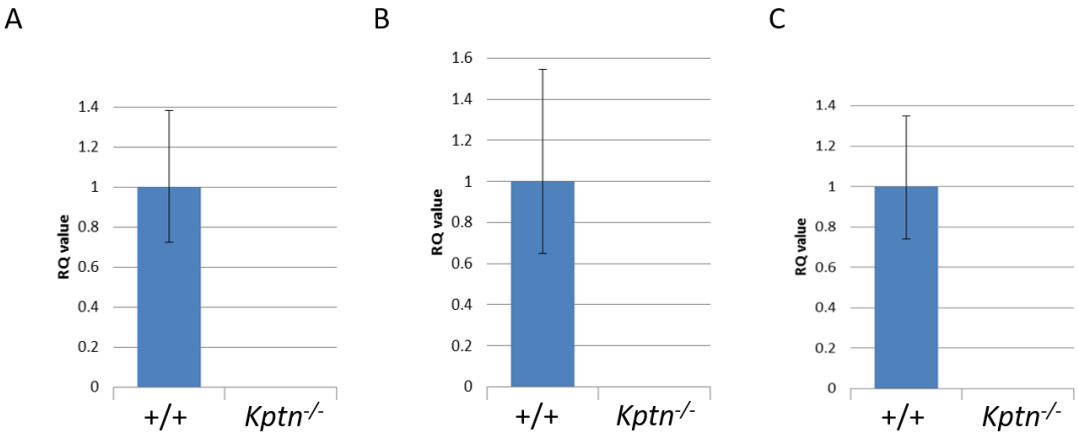


Figure 4.8 *Kptn*$^{tm1a(EUCOMM)Wtsi}$ "knockout first" construct is a complete null. Expression profile of *Kptn*$^{-/-}$ (n=4) and wildtype (+/+; n=4) in **A.** brown adipose, **B.** white adipose, **C.** and liver tissues. The graphs are showing the relative quantity (RQ) of expression in each tissue from quantitative PCR. Values are plotted as mean ± SD.

4.2.2 Morphometric brain analyses

Since macrocephaly was observed in all the patients, in both the Amish and Estonian cohorts, I investigated whether the homozygous mice recapitulate this aspect of the human disorder. To investigate the skull and brain volumes, I employed several approaches, in collaboration with Stephen Sawiak (Magnetic resonance imaging (MRI), Binnaz Yalcin and Perrine Kretz (histological morphometrics), and with support of the Mouse Genetics Project (MGP) team (X-ray cephalometrics).

4.2.2.1 *Kptn* mutants have a severe and global macrocephaly

Using X-ray cephalometric analysis (Chapter 2, section 2.5), we detected an increase in skull perimeter and width and a decrease in skull length in *Kptn*$^{-/-}$ mice (perimeter P=0.0007, skull width P=9.7x10^{-6}, skull length P=0.003, Fig.4.9). The results indicate that the mutant mice

are macrocephalic. The fact that the skull length was decreased, while other parameters were increased, also implies skull shape abnormalities.

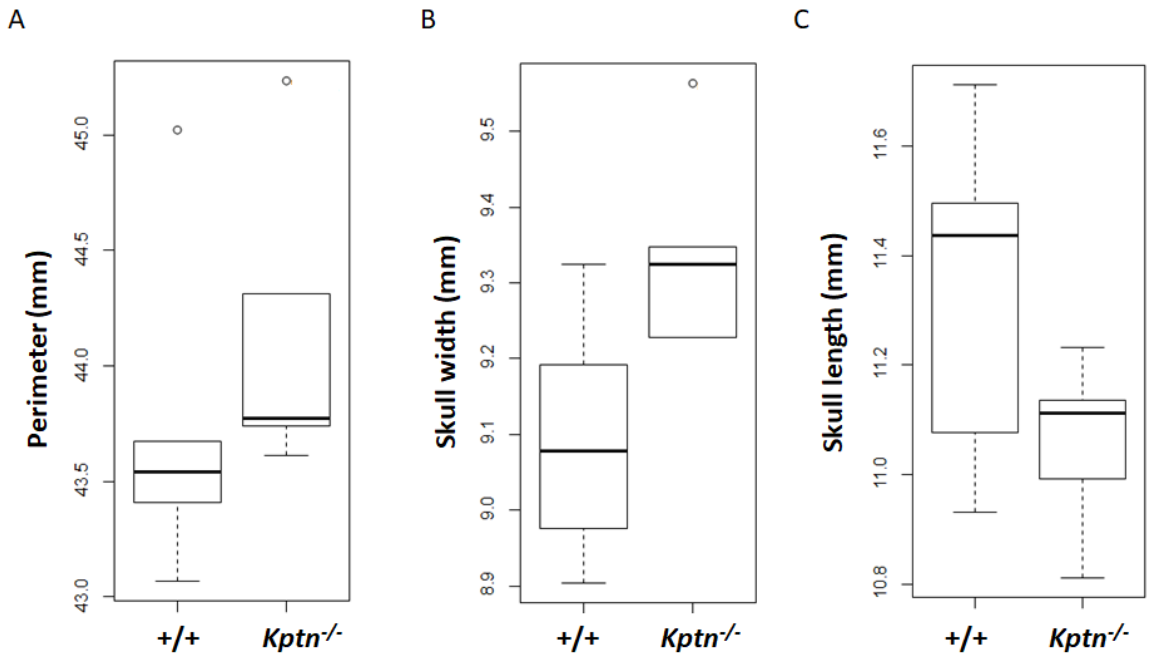


Figure 4.9 *Kptn*$^{-/-}$ mice are macrocephalic. Perimeter, skull width and skull length were measured with X-Ray cephalometrics in mutant (*Kptn*$^{-/-}$) and wildtype (+/+) male mice (n=7 per genotype). There was a significant increase in the perimeter and width of the skull, and a decrease in the length (perimeter length P=0.0007, skull width P=9.7x10-6, skull length P=0.003). Values plotted as a box and whisker plot, with outliers plotted as individual points.

To analyse brain volume directly, we performed Magnetic resonance imaging (MRI) imaging on post-mortem mouse brains (Chapter 2, section 2.4). Using MRI tensor-based morphometry followed by voxel-based quantification of brain volume, we detected intracranial volume increase in the homozygous *Kptn* mice of both genders when compared to controls (Male: P= 0.0009, t=4.526 df=11, two-tailed Student's *t*-test; Female: P= 0.0001, t=5.226 df=14, two-tailed Student's *t*-test). Since *Kptn*$^{-/-}$ mice have larger brains, removing the effect of overall brain volume also removes the *Kptn* effects. I, therefore, did not adjust for overall brain volume. Without regressing out the overall volume, there were detectable

effects at every part of the brain (significant at p<0.05 family-wise error rate corrected). We, therefore, set a more stringent threshold (p < 0.01 family-wise error corrected), so that only the most significant areas are highlighted Fig.4.10.

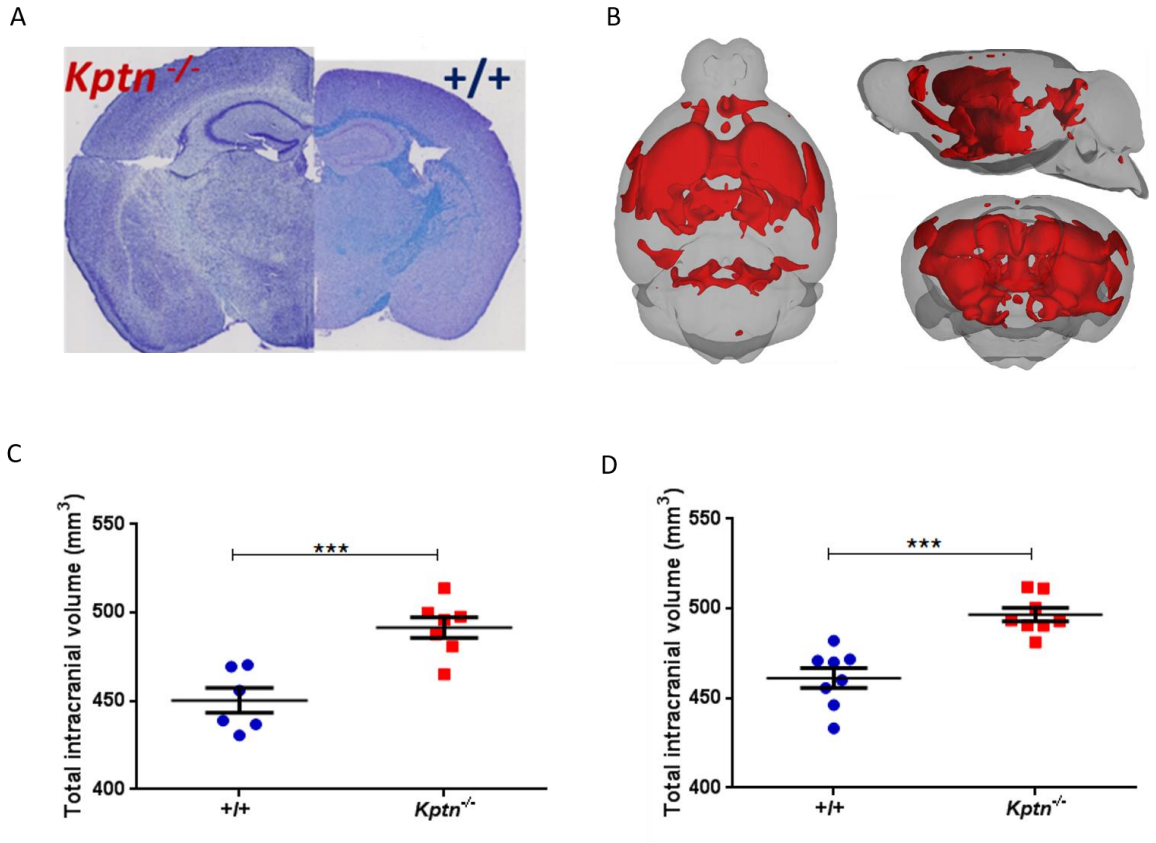


Figure 4.10 *Kptn*$^{-/-}$ mice are macrocephalic. **A.** Representative coronal histology section (Nissl staining) of *Kptn*$^{-/-}$ and control (*+/+*) brain respectively, qualitatively showing the mutant mice have larger brain than controls. **B**. 3D reconstruction MRI image of *Kptn*$^{-/-}$ brain; regions highlighted in red are significantly larger in mutant mice than in controls (with a threshold of significance at p < 0.01 family-wise error corrected). **C.** Total intracranial brain volume, from MRI, of male mutant (*Kptn*$^{-/-}$; red; n=7) and control mice (*+/+*; blue; n=6), showing an enlargement in intracranial volume the *Kptn* mutant mice (P= 0.0009***) **D**. Total intracranial brain volume, from MRI, of female mutant (*Kptn*$^{-/-}$; red) and control mice (*+/+*; blue) (n=8 per genotype), showing an enlargement in intracranial volume the *Kptn* mutant mice (P= 0.0001***). Values are plotted as mean ± SEM.

The global increase in brain volume was confirmed by histological volumetric analysis (Chapter 2, section 2.6). The representative coronal histology section of *Kptn*$^{-/-}$ and control brains visually illustrates the mutant mice have enlarged brains (Fig.4.10A). For a more thorough quantitative approach to analyse the neuroanatomical defects in *Kptn*$^{-/-}$ mice, we used a robust method of assessment of 78 brain parameters across 20 distinct brain regions, developed by my collaborators (Mikhaleva et al., 2016). This consisted of a systematic quantification of the same three coronal brain regions down to cell level resolution and blind to the genotype (Fig.4.11; Chapter 2, section 2.6; Table 3).

In males (n=8 per genotype; Fig.4.11A) the total brain area was significantly increased across the three coronal sections (section 1: +14.4%, P=0.0007; section 2: +9.1%, P=0.00008; section 3: +10.2%, P=0.00002) concomitantly with enlarged cortical regions including the cingulate cortex (+10.8%, P=0.03), motor cortex (section 1: +6.2%, P=0.01; section 2: +10.3%, P=0.0008), somatosensory cortex (section 1: +10.3%, P=0.008; section 2: +6.2%, P=0.004), and the piriform cortex (+19.2%, P=0.004). White matter structures were also affected and included the genu of the corpus callosum (+22.2%, P=0.00006), the soma of the corpus callosum (+16.1%, P=0.0003), and the internal capsule (+12.1%, P=0.0002). The lateral ventricles were the only brain regions exhibiting a decreased size (section 1: -36%, P=0.05; section 2: +54%, P=0.02).

This was consistent in the females (n=8 per genotype; Fig.4.11B), where the total brain area was significantly increased across the three coronal sections (section 1: +13.8%, P=0.00001; section 2: +8.3%, P=0.0002; section 3: +7.9%, P=0.0003). Similar cortical regions and white matter structures were also significantly enlarged in female mice, for example, the somatosensory cortex (+6.9%, P=0.0001) and the soma of the corpus callosum (+27.3%, P=0.000017). The size of the hippocampus is the only brain structure not affected in *Kptn*$^{-/-}$ when compared to controls in both males and females.

Taken together, the results indicate that *Kptn* depletion results in an increase in brain volume, with both white and grey matter affected in both sexes.

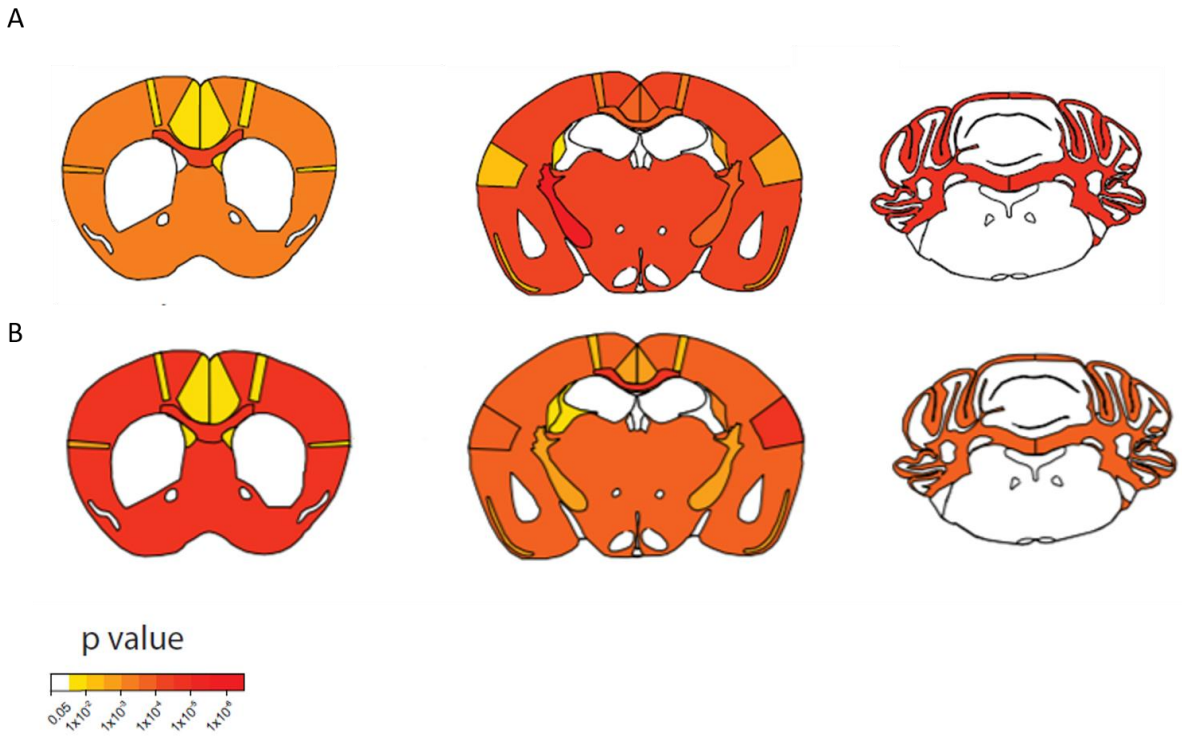


Figure 4.11 There is significant and global increase in size in the *Kptn*$^{-/-}$ brain, compared to controls. Images showing a representative heat map of p values of change in brain size (%) across three coronal sections – section 1 (left), section 2 (middle), and section 3 (right) - for male (**A**) and female (n=8 per genotype) (**B**) mice brains (n=8 per genotype), with red regions denoting areas with highest change in *Kptn* mutants (%) when compared to littermate wildtype controls. Below is the panel of colours and associated p value scores.

4.2.2.2 *Kptn* mutants have a progressive macrocephaly

We next investigated whether the macrocephaly observed in the *Kptn*$^{-/-}$ mice is present from birth (as was observed in the Amish patients) or is progressive (as with the Estonian siblings). To investigate this we performed morphological analysis on brains of postnatal day (P)0 mice (wildtype n=8, *Kptn*$^{-/-}$ n=9) using a quantitative histological approach with 53 parameters of size and surface, similar to the analysis done on adult brains and with most parameters having an equivalent in the adult mice (Chapter 2, section 2.6; Table 4).

We observed no difference in 51 out of 53 parameters, including total brain area (Fig.4.12A). This indicates that macrocephaly phenotype in the *Kptn*$^{-/-}$ mouse model is not present at birth, and is, therefore, a progressive form, likely associated with postnatal development processes. Out of the two parameters that were affected, there was 11% reduction in the area of the hippocampus in the mutants (P=0.04, Fig.4.12B) and a 9% increase in the internal capsule (P=0.03, Fig.4.12C), which suggests that some morphological anomalies originate from prenatal stages, but are restricted to the hippocampus and the internal capsule.

Taken together, these data allowed us to discriminate between different forms of macrocephaly, concluding that *Kptn* deficiency is associated with progressive macrocephaly in the *Kptn*$^{-/-}$ mice.

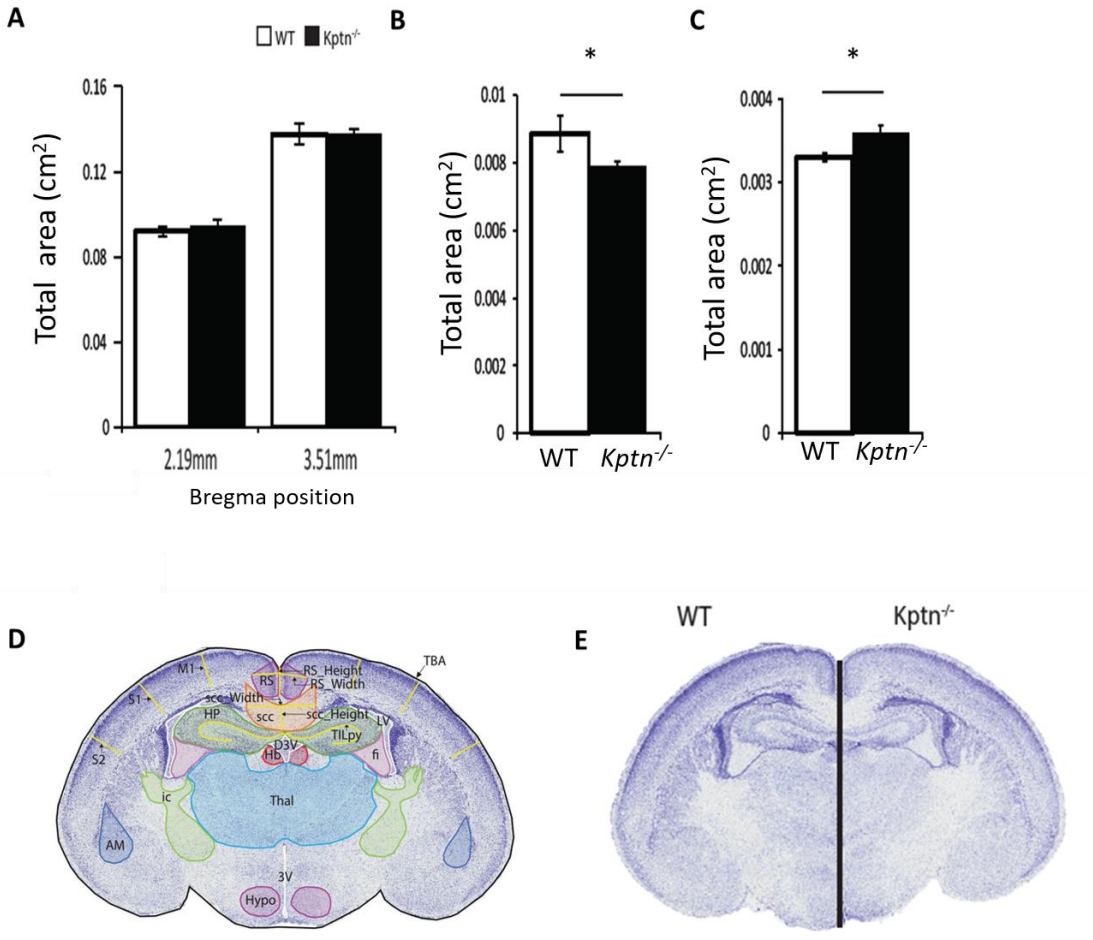


Figure 4.12 P0 mutant mice are not macrocephalic. **A.** Total brain area shown for wildtype (WT) and *Kptn*$^{-/-}$ male P0 mice at Bregma +2.19mm and 3.51mm. In both instances, the total area was not significantly different between genotypes. **B**. The hippocampus total area is reduced by 11% in the *Kptn*$^{-/-}$ mice (P=0.04*). **C.** The internal capsule total area is increased by 9% in the *Kptn*$^{-/-}$ mice when compared to wildtype controls (P=0.03*). **D.** A representative image of the parameters in the P0 brain morphology analysis, including primary somatosensory cortex (S1), secondary somatosensory cortex (S2), motor cortex (M1), retrosplenial granular cortex (RS), splenium of the corpus callosum (scc), habenula (Hb), thalamus (Thal), hippocampus (HP), total internal length of pyramidal layer (TILpy), amygdala (AM), hypothalamus (Hypo), internal capsid (ic), total brain area (TBA), fimbria (fi), third ventricle (3V), dorsal 3rd ventricle (D3V). **E**. Representative image of wildtype (WT) and *Kptn*$^{-/-}$ at Bregma 3.51 mm, with no phenotypic differences observed between genotypes (Nissl staining).

4.2.2.3 Increased cell count and proliferation rate contribute to the adult macrocephaly phenotype in *Kptn* mutants

To elucidate the cellular mechanisms behind the macrocephaly we investigated whether the macrocephaly observed in the adults could be explained by a larger number of cortical cells (assessed in males only, due to lack of sexual dimorphism), rather than cell size, as have been observed in some models of macrocephaly (Kwon et al., 2001). To answer this question we measured the following parameters in male mice (n=8 per genotype; Chapter 2, section 2.6) - cell count, cell area, cell circularity, cell solidity, and layer (region) area in each cortical layer (I - VI) of motor and somatosensory cortical regions (layer area: layer I: +21.1%, P=0.0018; layer II/III: +22.4%, P=0.012; layer IV: +29.9%, P=0.00054; layer V: +17.7%, P=0.013; layer VI: +17.4%, P=0.013) and cell count was increased in layers II to V (layer II/III: +21.7%, P=0.028; layer IV: +20.1%, P=0.008; layer V: +16.3%, P=0.028). There was a 19.4% increase in average cell count across the cortical layers and 21.8% increase in average layer (region) area in the mutants (Fig.4.13A-B). There was no significant increase in cell density or average cell area (a supporting parameter for cell size) (Fig.4.13A-B). There was no significant change in cell circularity and cell solidity, suggesting the rate of cell death is not affected. Taken together these results indicate that the increase in cell number in the brain contributes to the severe macrocephaly phenotype in adult *Kptn*$^{-/-}$ mice.

We next investigated whether there was an increase in the rate of proliferation (Chapter 2, section 2.6). Ki-67 protein is commonly used as a cellular marker for proliferation, and is present in all active phases of the cell cycle, but not in the quiescent (non-dividing cells). We performed a cell count of Ki-67 positive cells, and identified an average 11% increase in Ki-67 positive cells in the dentate gyrus of the hippocampus of male adults (P=0.045), indicating an increase in the rate of proliferation in the hippocampus (Fig.4.13C).

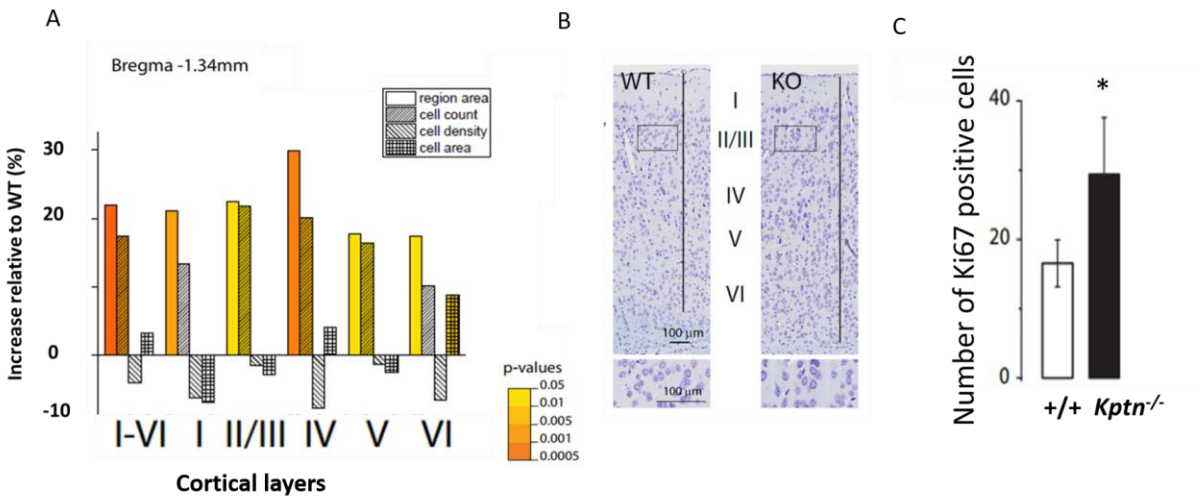


Figure 4.13 *Kptn*$^{-/-}$ mice have an increased cortical cell count and proliferation rate. A. Cell count and total cell area were measured in each cortical layer (I - VI) of motor and somatosensory cortical regions. All six cortical layers were increased in the mutants relative to the WTs (n=8 per genotype) (layer I: +21.1%, P=0.0018; layer II/III: +22.4%, P=0.012; layer IV: +29.9%, P=0.00054; layer V: +17.7%, P=0.013; layer VI: +17.4%, P=0.013) and cell count was increased in layers II to V (layer II/III: +21.7%, P=0.028; layer IV: +20.1%, P=0.008; layer V: +16.3%, P=0.028). The colours of the bars represent the significance level for each measurement (as shown by panel of colours and associated p value scores). Values are plotted as mean ± SEM. **B.** Representative images showing a larger number of cells in the *Kptn*$^{-/-}$ when compared to the control, without a change in density. **C.** There is an increase in Ki67 positive cells in the dentate gyrus of the hippocampus (P=0.045*), indicating an increase in proliferation in the hippocampus. Values are plotted as mean ± SEM.

4.2.3 Behavioural and cognitive consequences of *Kptn* deficiency

Next, I determined whether the cognitive and behavioural phenotypes observed in the patients were phenocopied in this mouse model. Because there is no recorded sexual dimorphism in *KPTN*-syndrome patients, and no identified sexual dimorphism based on brain morphometric analysis, I conducted all subsequent experiments in male mice, unless otherwise stated, in order to reduce the number of animals used.

4.2.3.1 Kptn$^{-/-}$ mice have increased locomotor activity and anxiety-like phenotypes

To characterise the general behaviour of *Kptn*$^{-/-}$ mice, I assessed their overall locomotor capabilities, such as distance covered, time spent moving, and velocity travelled, compared to controls, using the open field test (Chapter 2, section 2.3.1). The *Kptn* mutant mice spent more time moving (mean difference: 82.16 sec, $P \leq 0.05$, t=2.640 df=23, two-tailed Student's t-test; Fig.4.14B) than controls and travelled a greater distance (mean difference: 1320cm, P = 0.0185, t=2.534 df=23, two-tailed Student's t-test; Fig.4.14A). This is consistent with the hyperactivity phenotype observed in a subset of patients with *KPTN*-related syndrome (Baple et al., 2013). All other activity parameters were not significantly different between genotypes.

To test for anxiety-related behaviours, a phenotype shared across all the patients; I used the light/dark box assay (Chapter 2, section 2.3.2). *Kptn*$^{-/-}$ mice spent significantly more time in the dark zone (mean difference: 190.2sec, P<0.0001 t=4.946 df=24, two-tailed Student's t-test; Fig.4.15A) and had a reduced frequency of visits to the light zone (mean difference: -7.09, P= 0.0002, t=4.326 df=25, two-tailed Student's t-test; Fig.4.15B), confirming a strong anxiety-like phenotype, which is concordant with the anxiety observed in the patients. It is important to note that the mutant mice still transitioned into the light zone, thus exploring both areas, but showing a significant increase in preference for the dark zone compared to controls. This excludes the possibility that this increase in preference is due to the lack of exploration.

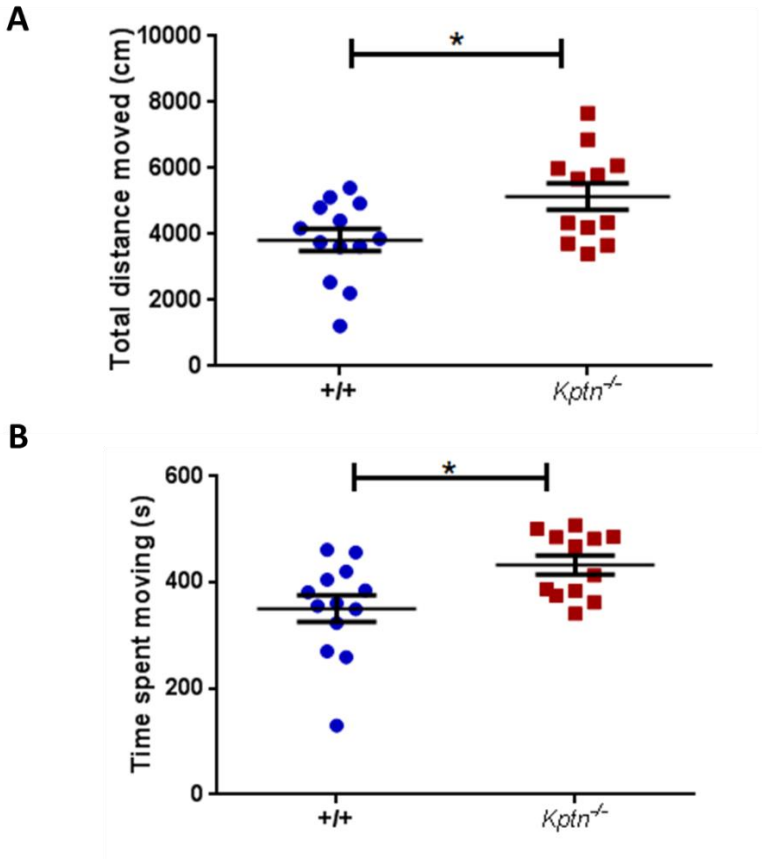


Figure 4.14 *Kptn*$^{-/-}$ mice display increased locomotor activity. The distance covered in an open field and time spent moving were plotted. **A**. *Kptn*$^{-/-}$ mice (red; n=12) covered significantly more distance (P=0.0185*) than wildtype controls (+/+) (blue; n=13) and **B**. spent more time moving (P=0.0146*). Values are plotted as mean ± SEM.

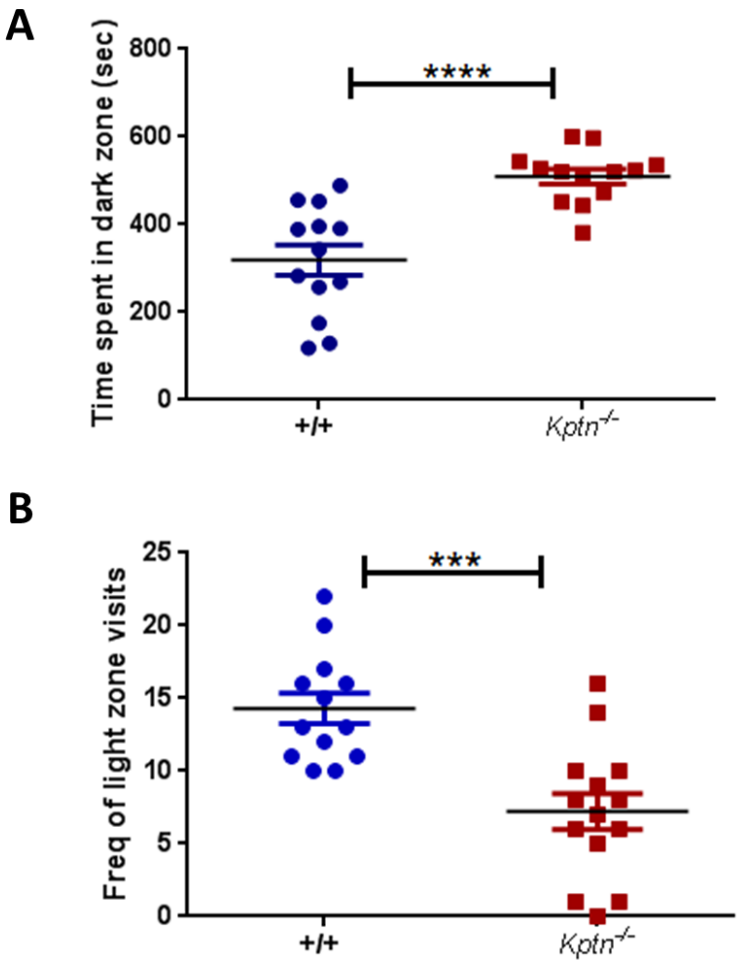


Figure 4.15 *Kptn* homozygous mice have increased anxiety-like behaviour. **A.** *Kptn*$^{-/-}$ mice (red; n=14) spend significantly longer time in the dark zone (P< 0.0001****) of a light/dark box than wildtype (+/+) controls (blue; n=13), and **B.** have reduced frequency of visits to the light zone (P=0.0002***). Values are plotted as mean ± SEM.

4.2.3.2 Kptn$^{-/-}$ have a deficit in hippocampal-dependent memory

All the *KPTN*-related syndrome patients were diagnosed with intellectual disability and global developmental delay, with variable severity. Since learning and memory are integral to cognition, to model the cognitive impairment observed in the patients, I tested the mice using several memory and learning assays (Sweatt, 2004).

4.2.3.2.1 Social recognition

First, I tested the mice using the social recognition assay (Chapter 2, section 2.3.4), which exploits in the innate preference of mice for investigating novel over familiar conspecifics and assesses olfactory-mediated hippocampus-dependent memory (Dias et al., 2016; Ferguson et al., 2001; Guan et al., 1993; Kogan et al., 2000; Sanchez-Andrade et al., 2005; Winslow and Insel, 2004). The *Kptn*$^{-/-}$ mutant mice did not differ from the wildtype controls in their levels of social approach behaviour, measured as time spent investigating the novel conspecific (stimulus A) on trial 1, implying the mice do not have a deficit in social interaction. Both wildtypes and *Kptn*$^{-/-}$ mice habituated to the stimulus over four trials, and significantly increased their investigation time on trial 5 (no significant interaction, genotype difference $F(1, 17) = 1.004$, $P=0.3304$; trial 4 vs trial 5 wildtype $P<0.0001$, mutants $P=0.0023$, *post-hoc* analysis after two-way ANOVA), indicating they were able to differentiate between the stimuli in trial 5 (stimulus B) and trial 4 (stimulus A), which is suggestive of an acquired memory for stimulus A (Fig.4.16A) and indicative of a functional olfactory system.

During 24h memory test, the wildtype controls spent longer (investigating the unfamiliar stimuli on Day 2, whereas the *Kptn* mutants did not show a preference for the unfamiliar mouse when compared to controls, indicating a 24h memory impairment (wildtype: $P=0.0002$, $t= 4.79$ $df=2$; Mutant: $P=0.985$, $t=0.0185$ $df=14$, two-tailed multiple *t*-test with multiple comparison corrections; Fig.4.16B).

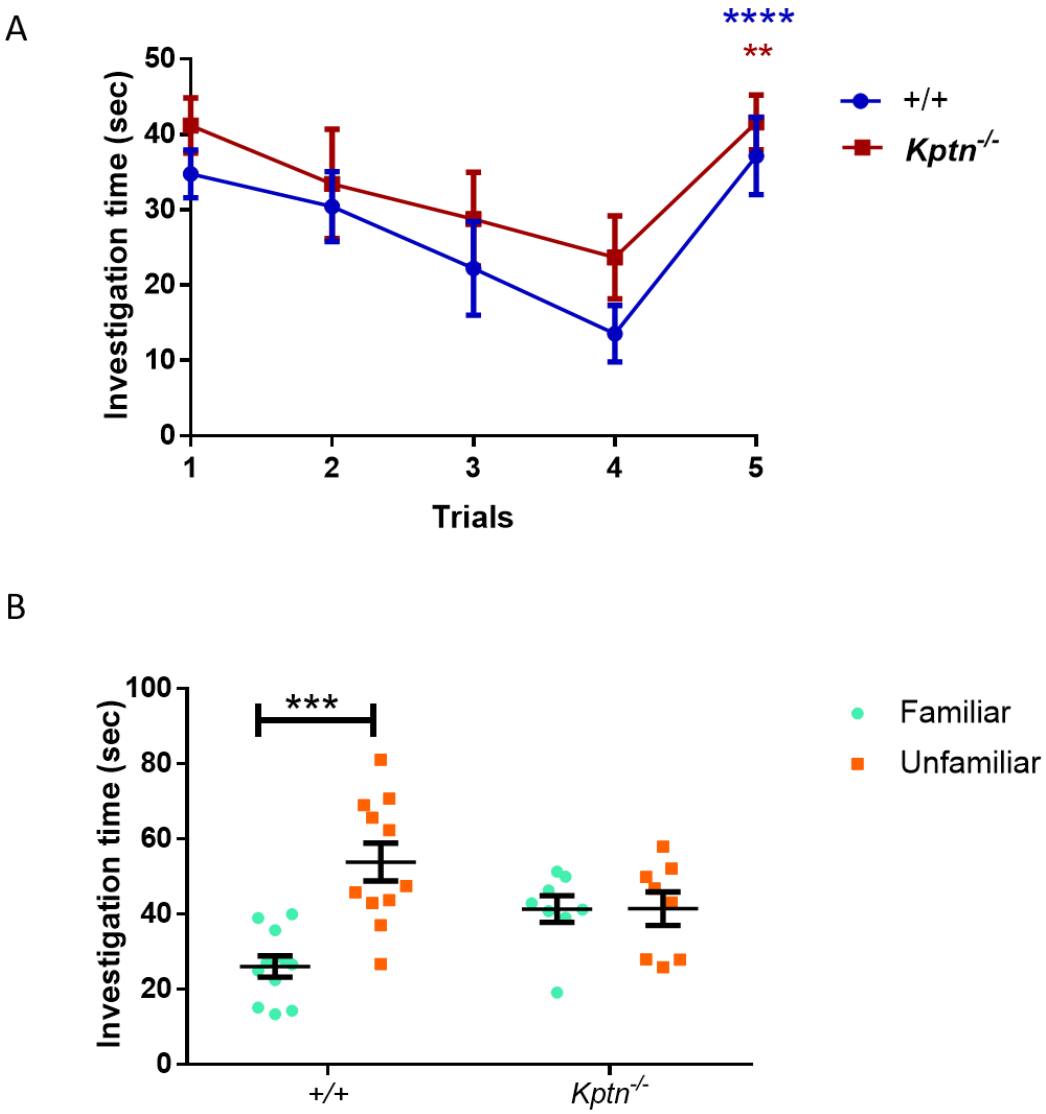


Figure 4.16 24h memory impairment in *Kptn*$^{-/-}$ mice. A. Memory acquisition (Day 1). Both controls (n=11, +/+) and *Kptn*$^{-/-}$ mutant (n=8, *Kptn*$^{-/-}$) (only females tested) recognize stimulus animal repeatedly presented to them over the course of four trials, as shown by a decline in the investigation time over trials 1-4. Both mutant and wildtype mice display an increase in the investigation time on trial 5 when presented with a novel stimulus animal (wildtype P <0.0001****, mutants P=0.0023**). **B.** 24h memory test (memory retention test). The controls, but not the mutant mice show a significant increase in time spent investigating the unfamiliar stimulus vs the familiar from Day 1), suggesting that *Kptn*$^{-/-}$ mice do not retain the memory of a familiar animal over 24h (wildtype: P=0.0002***, Mutant: P= 0.985).

Taken together, *Kptn*$^{-/-}$ mice indicate likely deficits in hippocampus-dependent memory. To confirm this phenotype, I next tested the mice in a hippocampus-dependent spatial memory assay, the Barnes maze.

4.2.3.2.2 Barnes maze

Barnes maze is a dry alternative of the Morris water maze and relies on the inherent tendency of mice to want to escape an aversive environment (Chapter 2, section 2.3.7). Mice placed on a brightly lit open table surface with 20 holes around the periphery must use spatial cues to repeatedly locate an escape box beneath one of the holes, the *target* (Harrison et al., 2006; Harrison et al., 2009; Koopmans et al., 2003).

4.2.3.2.2.1. Acquisition results

During training, there was no difference between *Kptn*$^{-/-}$ and wildtype in the time taken to approach the escape box (measured by primary latency to approach) during both training 1 (T1.1-T1.2) (no significant interaction, genotype difference: F (1, 207) = 0.1512, P=0.6978, two-way ANOVA) and training N2 (T2.1-2.3) (no significant interaction, genotype difference: F (1, 262) = 0.4383, P=0.5085, two-way ANOVA), indicating that both *Kptn* mutant and wildtype mice were able to locate the target zone with equivalent speed (Fig.4.17A). There was a significant overall reduction in the time taken to reach the escape box for both genotypes during training (F (4, 469) = 11.33, P<0.0001, two-way ANOVA) (Fig.4.17A).

There was no detected difference in the total errors in wildtypes and mutants during Training 1 (T1.1-1.2; Chapter 2, section 2.3.7) (Fig.17B). However, during Training 2 (T2.1-T2.3; Chapter 2, section 2.3.7) the *Kptn* mutant mice made overall significantly more errors (variable interaction P= 0.0157, genotype difference: F (1, 471) = 8.76, P = 0.0032; T2.2: mutant made on average 12.61 errors more than wildtypes (P<0.001), T2.3: *Kptn* mutant made on average 9.2 more errors than wildtypes (P<0.05); *posthoc* analysis after two-way ANOVA) (Fig.4.17B).

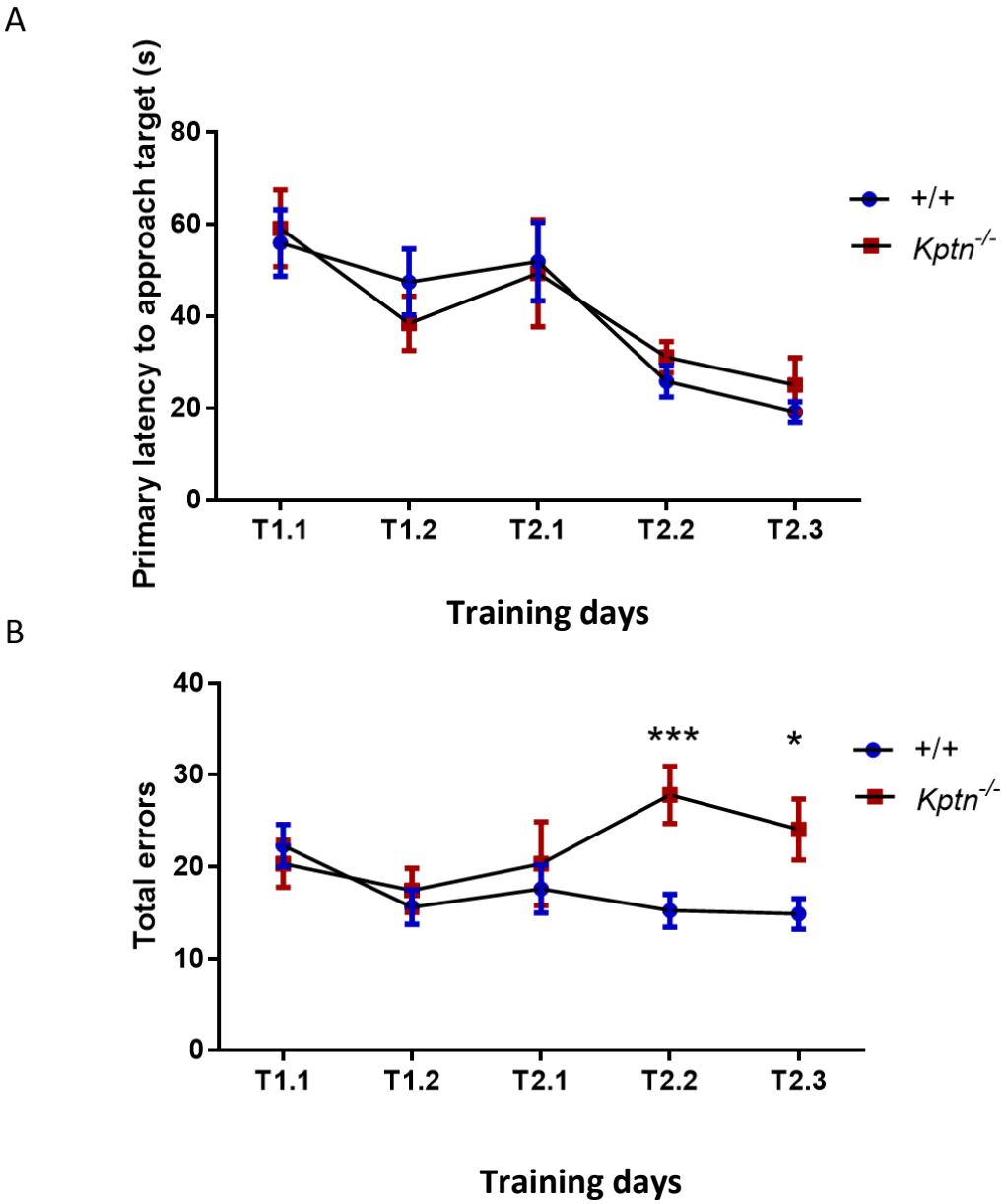


Figure 4.17 Barnes maze memory acquisition results. A. Primary latency across all days of training (Training 1 (T1.1-1.2) and Training 2 (T2.1-2.3)) was measured for wildtypes (+/+) (n=16) and *Kptn*$^{-/-}$ (n=11). There was no genotype difference across all training days and a reduction in time taken to get to the escape box from T1.1 to T2.3. **B**. The total number of errors recorded before the mice went into the escape box, overtraining days (Training 1 (T1.1-1.2) and Training 2 (T2.1-2.3)) was plotted. There was no significant genotype difference in total errors made on T1.1-1.2, but there was a genotype difference at T2.1 (P<0.001***) and T2.3 (P<0.05*). Values are plotted as mean ± SEM.

4.3.3.2.2.2 Memory retention results

During 24h probe, the percentage of time spent around the target hole relative to other holes was calculated (Chapter 2, section 2.3.7). Both genotypes spent significantly longer around the target hole when compared to other holes (no significant interaction, genotype difference: $F(1, 25) = 2.663$, $P=0.1153$; genotype difference at target: $P>0.05$, *post-hoc* analysis after two-way ANOV; Fig.4.18A, Fig.4.19). However at 72h probe, the mutant mice spent 5% less time at the target zone than wildtype controls (significant interaction $P=0.0326$, genotype difference: $F(1, 500) = 0.8423$, $P=0.3592$; duration at hole $P < 0.0001$, *post-hoc* analysis after a two-way ANOVA; Fig.4.18C, Fig.4.19). The percentage difference in time spent at each hole between mutant and controls is also depicted spatially, showing the location of the target zones during 24h and 72h probe trials, and the difference between genotypes of the proportion of time spent at the target during both probe trials (Fig.4.19).

There was a significant difference between the mean time spent in all holes vs target hole on 24h probe for both genotypes ($P \leq 0.001$ for both genotypes, *post-hoc* analysis after two-way ANOVA, Fig.4.18B), but the mutants did not have a significant difference between time spent at target vs non-target on 72h probe because of the overall reduction in time spent at target (controls: $P \leq 0.0001$, mutants: $P>0.05$, *post-hoc* analysis after two-way ANOVA, Fig.4.18D). These results, therefore, indicate that *Kptn* mutants cannot retain spatial memory for 72h in this task.

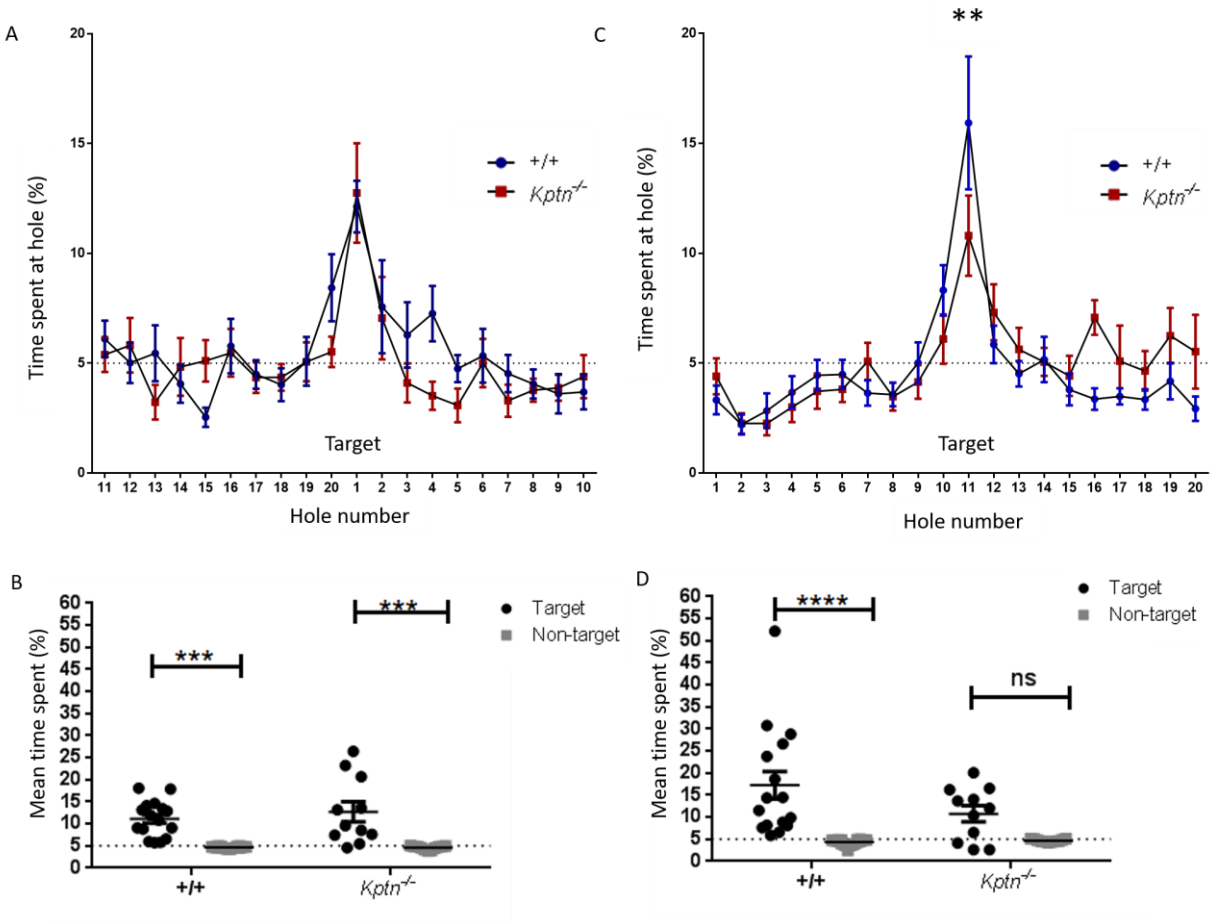


Figure 4.18 *Kptn* have a 72h memory impairment. A. The percentage of time male wildtype (+/+) (n=16) and *Kptn*$^{-/-}$ (n=11) mice spent each hole (1-20; with hole N1 being the target) of the Barnes maze during 24h probe trial was plotted. There was no difference in time spent near target hole (where during training there was an escape box). **B.** Mean percentage of time spent at target hole relative to all the other holes during 24h probe trial. Both genotypes spent significantly more time at target vs all other holes (P ≤ 0.0001***). **C** The percentage of time male wildtype (+/+) (n=16) and *Kptn*$^{-/-}$ (n=11) mice spent each hole (1-20; with hole N11 being the target) of the Barnes maze during 72h probe trial was plotted. The controls spent a significantly more time (5.13% more) at the target (hole N=11) than the *Kptn* mutants (P≤ 0.001**). **D.** Mean percentage of time spent at target hole relative to all the other holes during 72h probe trial. The mutants did not have a significant difference between target vs non-target due to their reduced time at the target, while the wildtypes spent significantly more time at target vs all other holes (P ≤ 0.0001***). Values are plotted as mean ± SEM.

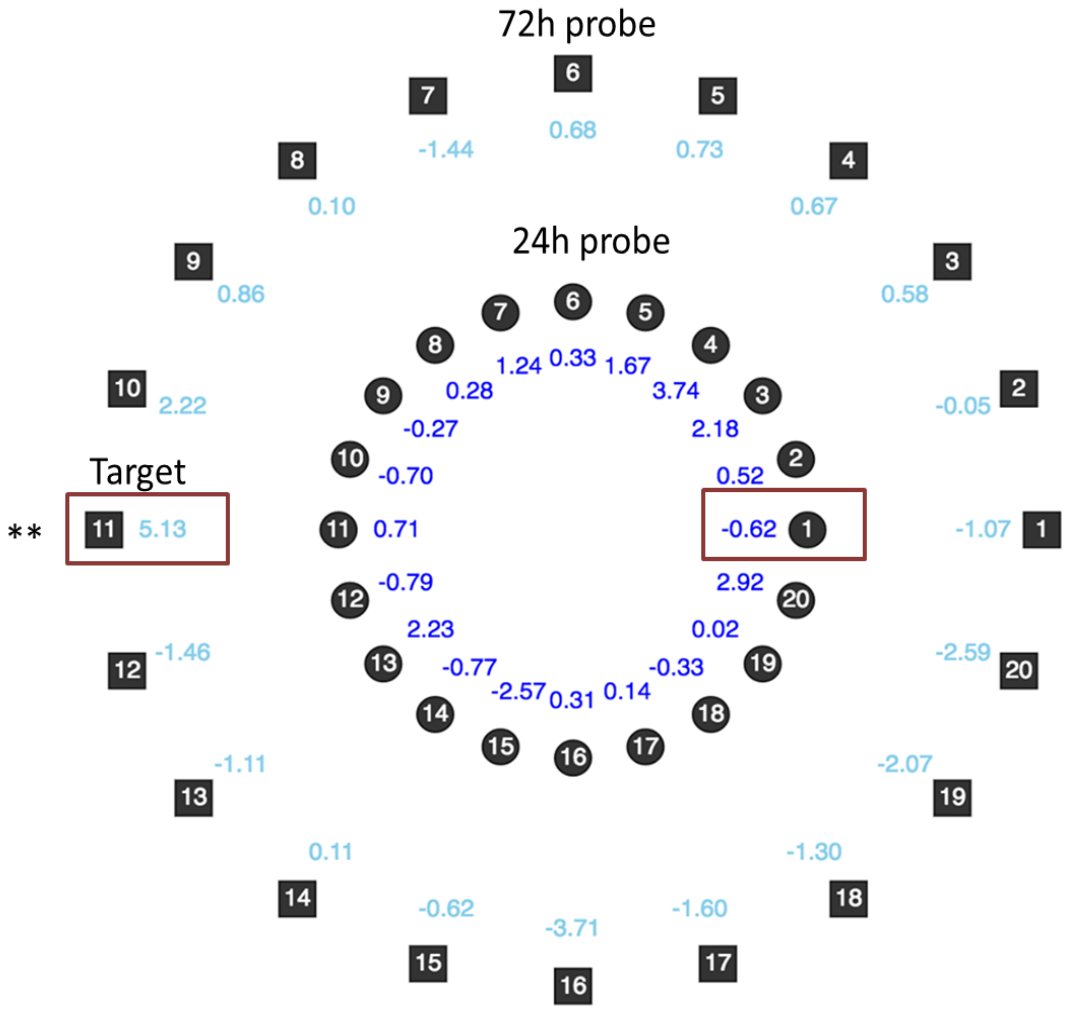


Figure 4.19 A spatial representation of the time *Kptn*$^{-/-}$ mice spent at each hole during 24h and 72h probes relative to controls. The two circles represent the spatial distribution of holes (labelled 1-20 in black) in Barnes maze during 24h and 72h probes. The target hole is annotated by a red box in both instances (24h probe: hole N1; 72h probes: hole N11). The numbers in blue represent percentage of time difference between genotypes at each hole, with 0 being no difference and positive values meaning controls spend longer at the hole. There is no significant difference at target hole between genotypes on 24h probe and therefore, but the controls spend 5.13 % longer at target hole in 72h probe ($P \leq 0.01$**).

The results from this assay testing hippocampus-dependent memory strongly support the hypothesis that the *Kptn* mutant mice have a deficit in hippocampus-dependent memory retention.

4.2.3.3 *Kptn*$^{-/-}$ mutants do not have impairment in spatial memory of up to 2h

I next tested the mice (n=12 for both genotypes) using the object discrimination paradigm, which incorporates object displacement (OD) and novel object recognition (NOR) testing (Chapter 2, section 2.3.3; pilots described in Chapter 3).

There was no significant genotype difference observed in the preference of mice (n=7 wildtype, n=9 *Kptn*$^{-/-}$) for the moved object (P=0.9498, t=0.06405 df=14, two-tailed Student's *t*-test) (Fig.4.20A). Both genotypes spent longer investigating the moved object in OD part of the paradigm than reference value mean of 50% chance (wildtype: mean: 74.67, P= 0.0242, t=2.995 df=6; *Kptn* mutants: mean: 75.42, P=0.0141, t=3.124 df=8; one-tailed *t*-test). Therefore despite lower sample size relative to the estimated cohort size of n=12 (power analysis, Chapter 2, section 2.3.3.3), both genotypes displayed sufficiently high preference score (%) to be able to detect a difference from the reference value of 50% ('no preference').

Those mice that were not excluded from the OD phase, were then tested 24h later in NOR part of the paradigm. There was no significant genotype difference in their preference for the novel object (P=0.4875, t=0.7146 df=13, two-tailed Student's *t*-test). Both genotypes spent significantly more with the novel object than the theoretical mean of 50% chance (wildtype: mean: 67.58, t=2.653 df=5; *Kptn*$^{-/-}$: mean: 61.41, t=2.075 df=8; one-tailed *t* test) (Fig.4.20B).

Overall, *Kptn*$^{-/-}$ did not exhibit 1-2h spatial memory impairment, nor an impairment in the NOR performance.

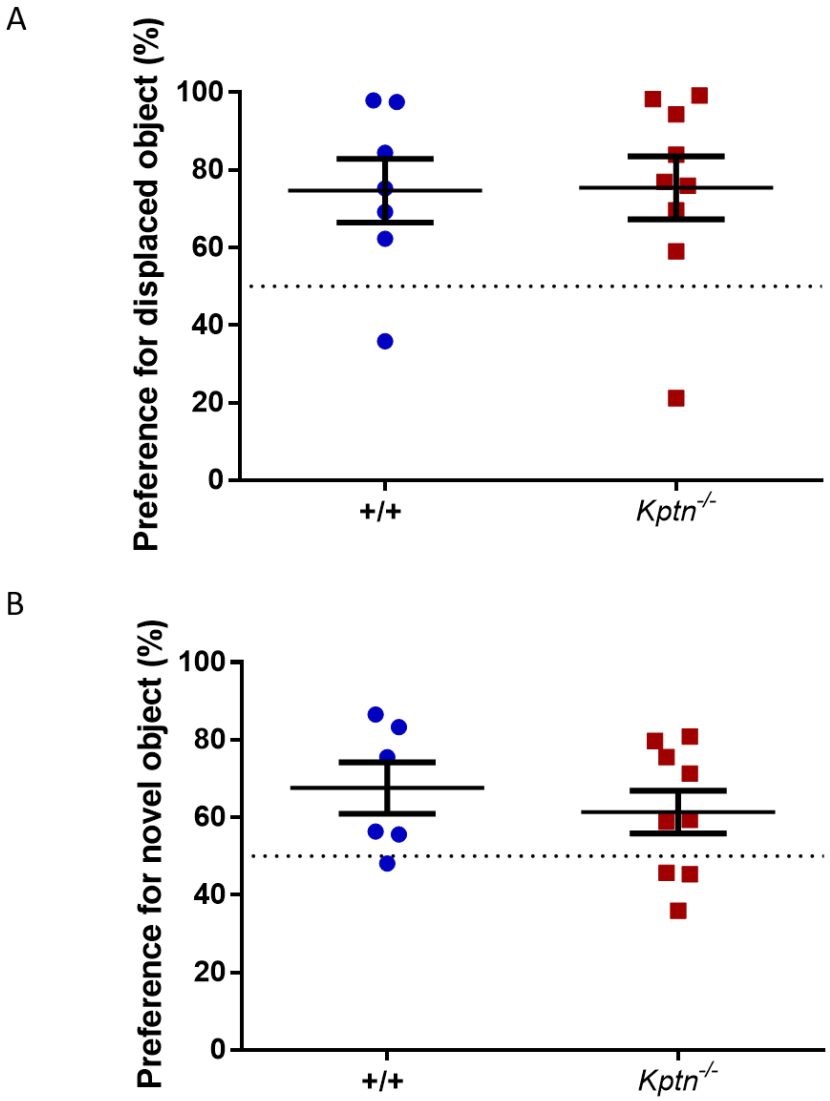


Figure 4.20 Object discrimination paradigm, displaying no impairment in 1-2h spatial memory and NOR performance in the mutants. **A.** Object displacement testing. The percentage of time spent (out of total investigation time of sample and displaced objects) with displaced object was plotted. Both genotypes (*Kptn*$^{-/-}$ n=9 and controls (+/+) n=7) had a significant preference for the displaced object, above 50% chance (dotted line). **B**. NOR testing phase. The percentage of time spent (out of total investigation of sample and novel objects) with novel objects was plotted. Both mutants (n=9) and wildtype (+/+) (n=5) had a significant preference for the novel object, above 50% chance (dotted line). Values are plotted as mean ± SEM.

4.2.3.4 *Kptn*$^{-/-}$ do not have an impaired perceptual non-hippocampal memory and learning

To test whether the mutants have non-hippocampal dependent cognitive impairments, I used the touchscreen pairwise discrimination (PD) task, an operant conditioning task that is run using a touchscreen platform (Chapter 2, section 2.3.6; Chapter 3).

N=12 per genotype were put through training, however only n=10 wildtype and n=8 *Kptn*$^{-/-}$ were able to learn the initial set of rules associated with the training phase. Two wildtypes and four mutants were excluded because they did not reach criteria within 25 day cut-off (Chapter 2, section 2.3.6). During the PD task, both genotypes started off at around 50% of correct trials in Session 1, which implies the mice had no bias for either of the images at the start of PD task. Both genotypes were able to reach the criteria 80% of correct trials for two consecutive days ($F(1, 16) = 1.144$, $P=0.3007$, two-way ANOVA; Fig.4.21), after a similar number of sessions (WT: 14.6 days, *Kptn*$^{-/-}$: 11.25 days, $P=0.2093$, $t=1.162$ df=16, two-tailed Student's *t*-test; Fig.4.22).

To ensure that the mice did indeed learn, the images were swapped (CS- image became the CS+ and vice versa) after the last PD task session and the percentage of correct trials was recorded for a subset of the mice (wildtype n=9, *Kptn*$^{-/-}$ n=4). As expected, the percentage of correct trials dropped significantly below 50% for both genotypes (genotype difference: ($F(1, 11) = 0.02269$, $P=0.8830$, two-way ANOVA)) between the last day of pairwise discrimination (session 25) and the first day of images being swapped from CS+ to CS- (session 1, with an arrow) ($F(6, 66) = 68.19$, $P<0.0001$, two-way ANOVA), (Fig.4.21). The percentage of correct trials increased with subsequent sessions, reaching the 50% correct at fifth session (Fig.4.21).

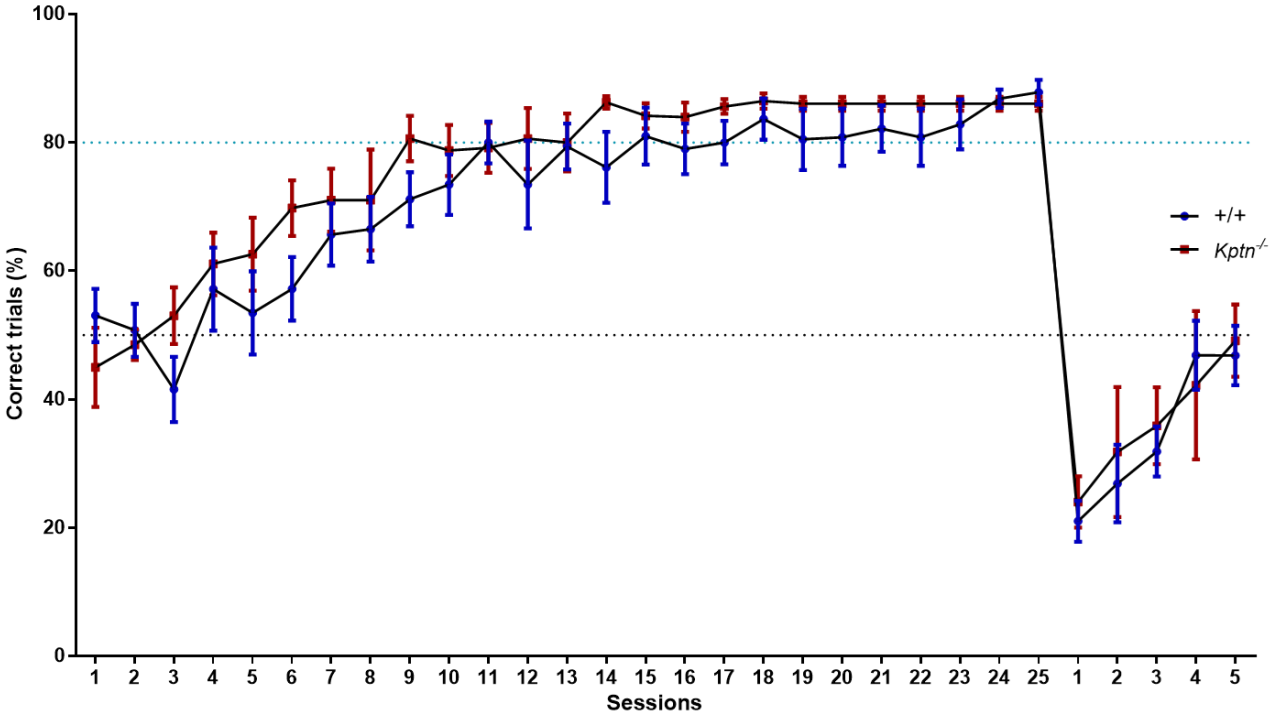


Figure 4.21 *Kptn* mutants do not display a difference in performance in the hippocampus-independent pairwise discrimination task. Percentage of correct trials (when CS+ image was nose-poked) out of the total trials done per session were plotted for *Kptn*$^{-/-}$ (n=8) and wildtype (n=10) mice. There is no significant difference in genotype between learning in the pairwise discrimination (P=0.3007). Once the mice reached criteria (80% correct trials; blue dotted line), the mean of percentage of correct trials over their last two days was taken and used in the overall mean for the subsequent sessions. Therefore, each session has an equal number of mice, even though some mice finished the assay faster than other. Mice (wildtype n=9, *Kptn*$^{-/-}$ n=4) showed significant decrease in the percentage of correct trials completed, below 50% chance (black dotted line), between the last day of pairwise discrimination (session 25) and the first day of images being swapped from CS+ to CS- (session 1, with an arrow) (P<0.0001), but there was no genotype difference in percentage of correct trials between new session 1 and 5 (P=0.8830, two-way ANOVA). Values are plotted as mean ± SEM.

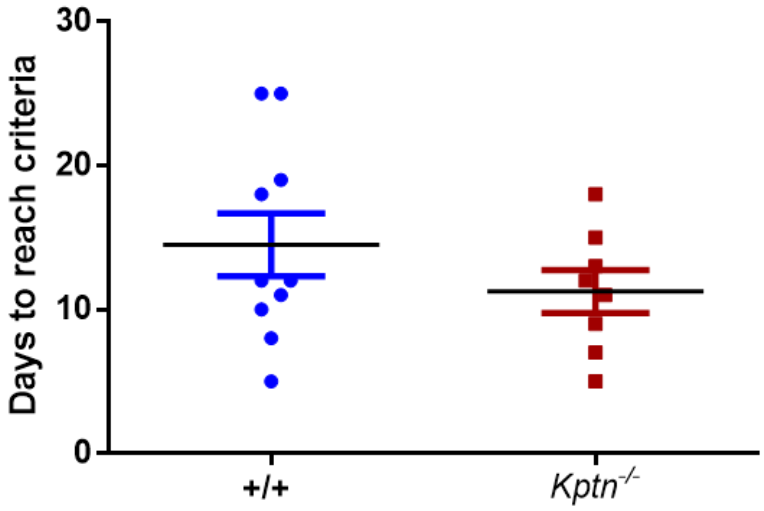


Figure 4.22 There was no detected genotype difference of days needed to reach criteria in pairwise discrimination task. The number of sessions (each session is one day) it took for *Kptn*$^{-/-}$ (n=8) and wildtypes (+/+) (n=10) to get 80% of trials correct on two consecutive days was plotted. No genotype difference was observed (P=0.2093). Values are plotted as mean ± SEM.

4.2.4 *Kptn* deficient transcriptome

In order to assess postnatal consequences of *Kptn* deficiency on a molecular level and explore affected mechanisms in different brain regions, I carried out gene expression analysis on four separate brain regions from male *Kptn* mutant and wildtype mice (n=6 per genotype) (Chapter 2, section 2.7). The regions analysed, in collaboration with Fernando Riveros-McKay at the Sanger Institute, were hippocampus, striatum, prefrontal cortex, and cerebellum. These structures are of particular interest due to the cognitive testing results and morphometric analysis I performed. Although no behavioural or cognitive assays specifically tested cerebellar function in *Kptn*$^{-/-}$ mice, this region was collected to provide additional information on the global consequences *Kptn*$^{-/-}$ mouse brain.

Genotypes were confirmed for each tissue by examining the counts of reads for *Kptn* across *Kptn*$^{-/-}$ and wildtype controls (Fig.4.23). In all of the four tissues, normalised read counts for *Kptn* clustered by genotype as expected (Fig.4.23).

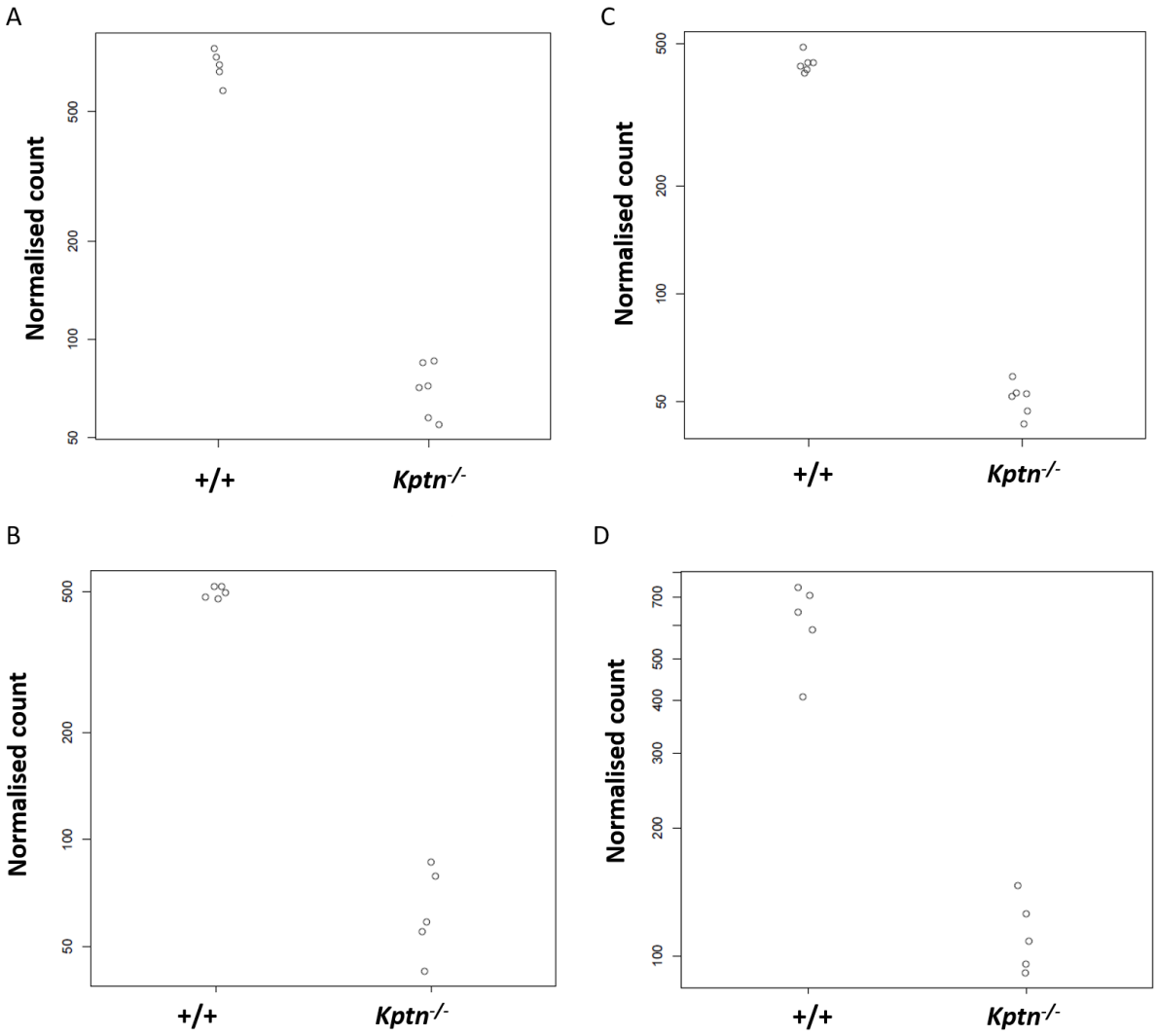


Figure 4.23 Normalised read counts for *Kptn* in wildtype (+/+) and *Kptn*$^{-/-}$ in cerebellum, prefrontal cortex, hippocampus, striatum. The genotype of the samples was confirmed by observing clustering of normalised *Kptn* counts by genotype in each of the four tissues collected: **A.** cerebellum, **B**. prefrontal cortex, **C**. hippocampus, and **D**. striatum.

A comparison between mutant and wildtype samples revealed that within each tissue, there are numerous differently expressed (DE) genes. Using a threshold of 5% FDR, we identified 578 DE genes in the cerebellum (out of which 326 were unique to the tissue), 776 DE genes in prefrontal cortex (PFC) (out of which 396 were unique to the tissue), 1315 DE genes in hippocampus (out of which 841 were unique to the tissue), and 3896 DE genes in the striatum (out of which 3232 were unique to the tissue) (Fig.4.24A). Only nine of these DE genes overlapped between all the tissues, with *Kptn* being one of them (Fig.4.24A). These nine common DE genes are enriched in seven KEGG pathways including antigen processing and presentation, allograft rejection, type I diabetes mellitus, autoimmune thyroid disease, salivary secretion, PPAR signalling pathway, viral myocarditis, and graft-versus-host disease.

Among the upregulated DE genes, of which there are 262 in the cerebellum, 606 in the hippocampus, 1776 in the striatum, and 439 in PFC (Fig.4.24B, D-G), there are only 3 that overlap all tissue (Fig.4.24B). Regarding the downregulated DE genes, there are 199 in the cerebellum, 709 in the hippocampus, 2120 in the striatum, 337 in the PFC (Fig.4.24C, D-G), and only four overlaps between all tissues (Fig.4.24C). Interestingly, when analysing the subset of DE genes in the hippocampus that are upregulated, negative regulation of mTOR signalling was the only significantly associated biological pathway (Fig.4.25D). It was also one out of 284 biological pathway GO terms enriched in PFC (Fig.4.25B).

In order to identify the biological and functional enrichment within the DE genes, I performed a pathway and GO enrichment analysis of DE genes in each tissue, using as background all genes expressed in the particular tissue being analysed (Chapter 2, section 2.7.3). There was an over-representation of only 2 KEGG pathways – one in the hippocampus (‘ribosome’) and one in the cerebellum (‘graft-versus-host disease’) (Table 6). However, there was enrichment for many GO terms in each tissue (Fig.4.25A-D).

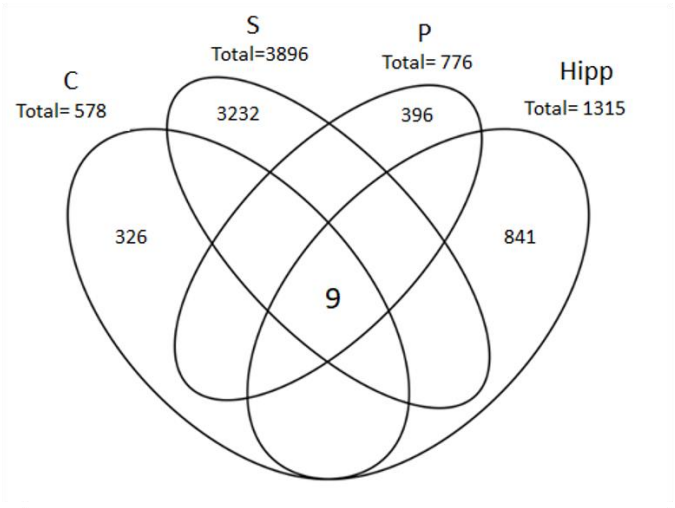
Table 6 KEGG pathway results of DE genes for different brain tissues in *Kptn*$^{-/-}$.

Tissues	KEGG term	Number of genes	Total N associated with pathway	Proportion of recall
Cerebellum	Graft-versus-host disease	8	18	0.444
Prefrontal Cortex	-	-	-	-
Hippocampus	Ribosome	84	125	0.672
Striatum	-	-	-	-

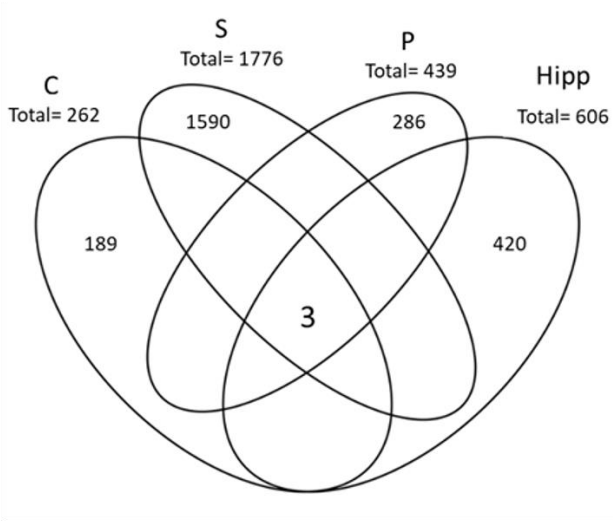
Enriched gene ontology categories included terms related to neuronal and central nervous system development, negative regulation of mTOR signalling, regulation of neurogenesis, proliferation, immune system, cell-cell signalling, adhesion, negative regulation of cell death, behaviour, actin cytoskeleton, synapse transmission, ion transport, translation and ribosome. Due to the recently described role of *KPTN* as a negative regulator of mTOR pathway, I tested whether the DE genes were enriched for mTOR signalling in the four tissues. Only PFC and hippocampus had an enrichment of mTOR regulation as one of the GO terms (Fig.4.25). No other tissues had mTOR signalling enrichment, suggestive that aberrant mTOR signalling is associated with these tissues.

Through this preliminary analysis of the RNAseq results from mutant *Kptn* mouse brains, there is a significant signal of aberrant mTOR signalling in specific tissues. These data are in line with the recently proposed role of *Kptn* in mTOR, and will drive future work on this model.

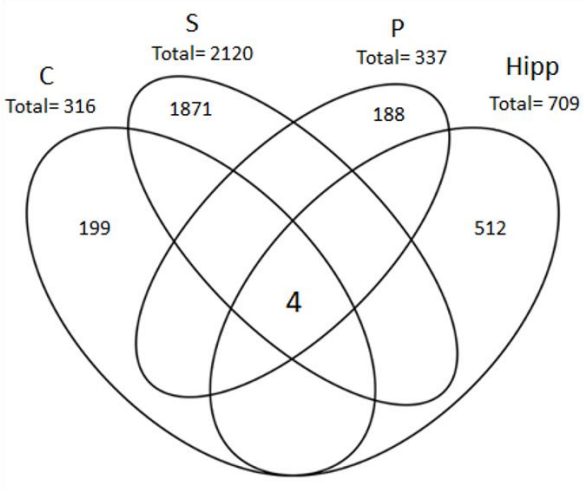
A


S
Total=3896
P
Total= 776
C
Total= 578
Hipp
Total= 1315
3232
396
326
841
9

B


S
Total= 1776
P
Total= 439
C
Total= 262
Hipp
Total= 606
1590
286
189
420
3

C


S
Total= 2120
P
Total= 337
C
Total= 316
Hipp
Total= 709
1871
188
199
512
4

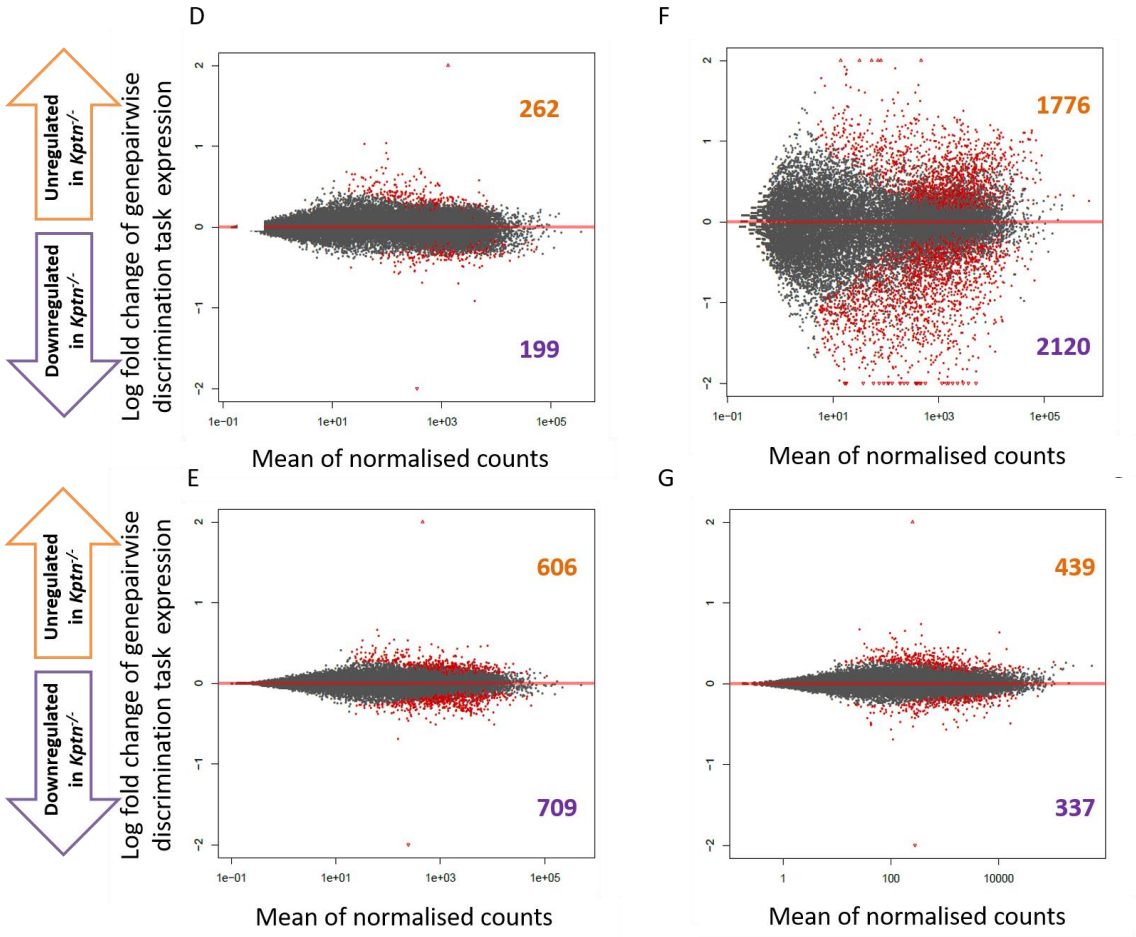
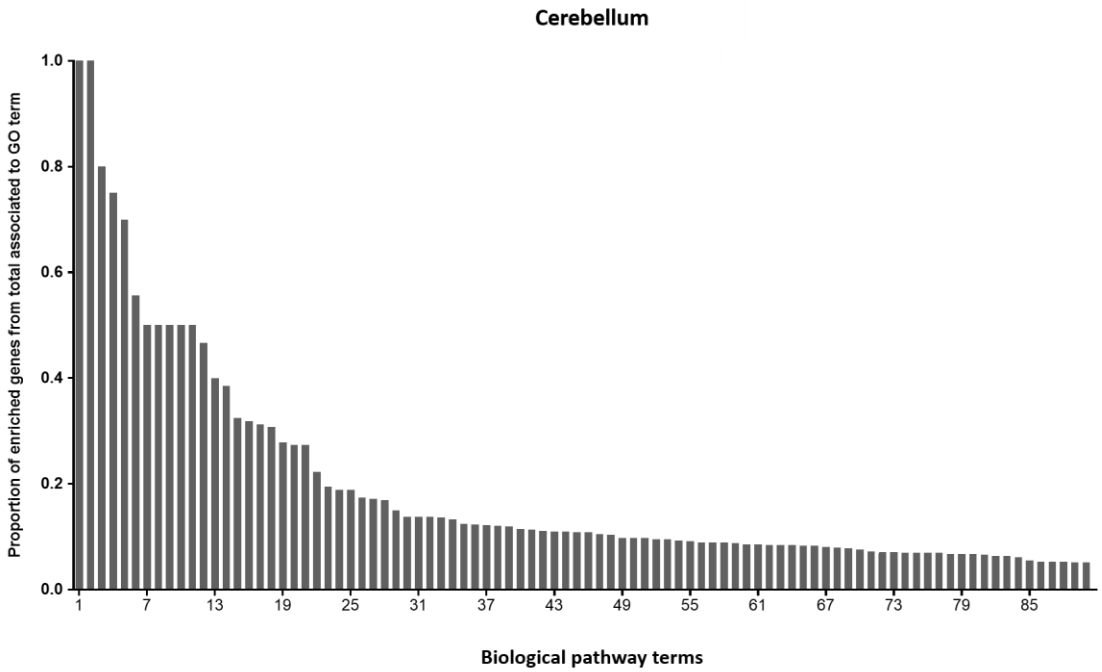
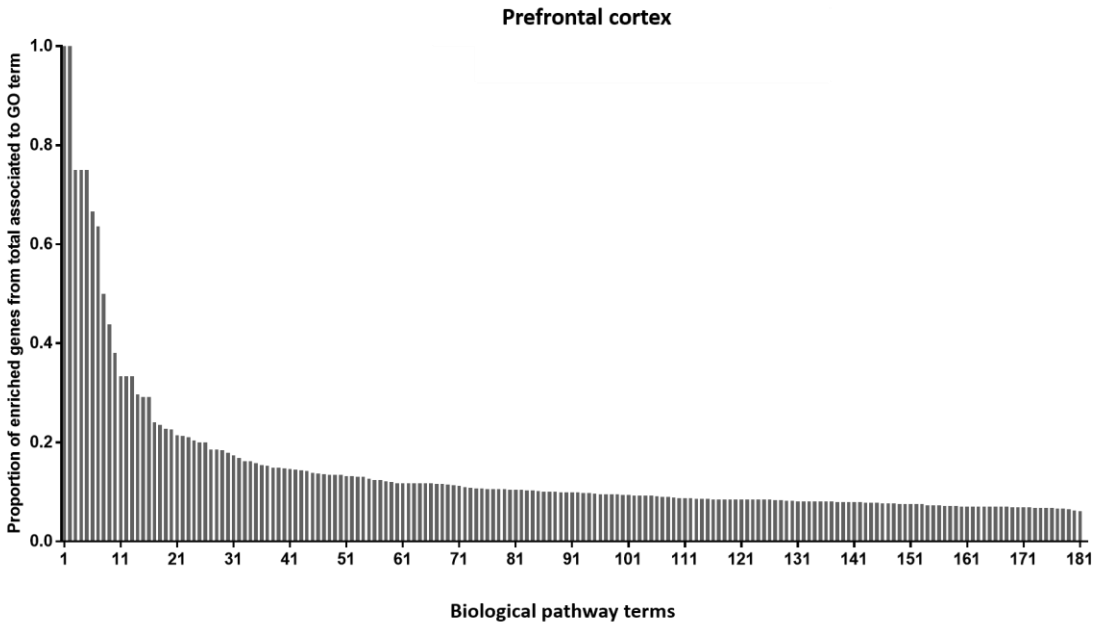


Figure 4.24 *Kptn*$^{-/-}$ mice have altered gene expression profiles in all brain tissues examined. A. Differentially expressed genes in all the tissues: 578 DE genes in the cerebellum (C) (326 unique to the tissue), 776 DE genes in prefrontal cortex (P) (396 unique to the tissue), 1315 DE genes in hippocampus (Hipp) (841 unique to the tissue), and 3896 DE genes in the striatum (S) (3232 unique to the tissue). **B**. Differentially expressed genes upregulated in *Kptn*$^{-/-}$ brain tissues**.** Among the total upregulated DE genes: 262 are in the cerebellum (C), 606 in the hippocampus (Hipp), 1776 in the striatum (S), and 439 in prefrontal cortex (P). **C**. Differentially expressed genes downregulated in *Kptn*$^{-/-}$ brain tissues. Among the total downregulated DE genes: there are 199 in the cerebellum (C), 709 in the hippocampus (Hipp), 2120 in the striatum (S), 337 in the prefrontal cortex (P). **D-G**. Plots of differential gene expression between *Kptn*$^{-/-}$ and wildtype in **D.** cerebellum, **E.** hippocampus, **F.** striatum, **G.** prefrontal cortex. The x-axis is the log10 average expression of all samples (normalised counts); the y-axis is DESeq2's shrinkage estimation of log2 fold changes between genotypes. Each gene is represented as a dot; significantly differentially expressed genes (Benjamini-Hochberg adjusted p-value< 0.05) are highlighted in red. Orange numbers indicate total upregulated genes per tissue, while purple numbers indicate total downregulated numbers per tissue.

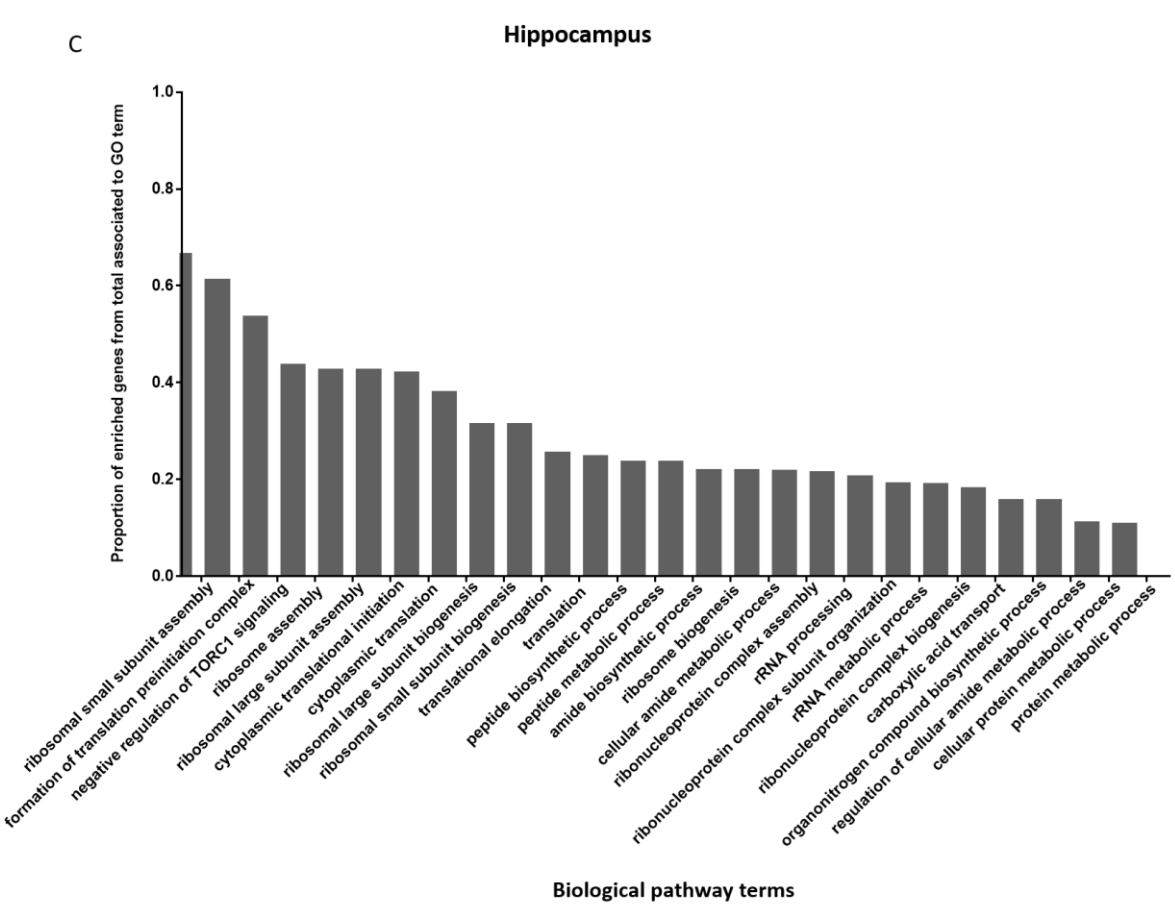
A


Cerebellum
Proportion of enriched genes from total associated to GO term
1.0
0.8
0.6
0.4
0.2
0.0
1
7
13
19
25
31
37
43
49
55
61
67
73
79
85
Biological pathway terms

B


Prefrontal cortex
Proportion of enriched genes from total associated to GO term
1.0
0.8
0.6
0.4
0.2
0.0
1
11
21
31
41
51
61
71
81
91
101
111
121
131
141
151
161
171
181
Biological pathway terms

C


Hippocampus
Proportion of enriched genes from total associated to GO term
1.0
0.8
0.6
0.4
0.2
0.0
ribosomal small subunit assembly
formation of translation preinitiation complex
negative regulation of TORC1 signaling
ribosome assembly
ribosomal large subunit assembly
cytoplasmic translational initiation
cytoplasmic translation
ribosomal large subunit biogenesis
ribosomal small subunit biogenesis
translational elongation
translation
peptide biosynthetic process
peptide metabolic process
amide biosynthetic process
ribosome biogenesis
cellular amide metabolic process
ribonucleoprotein complex assembly
rRNA processing
ribonucleoprotein complex subunit organization
rRNA metabolic process
ribonucleoprotein complex biogenesis
carboxylic acid transport
organonitrogen compound biosynthetic process
regulation of cellular amide metabolic process
cellular protein metabolic process
protein metabolic process
Biological pathway terms

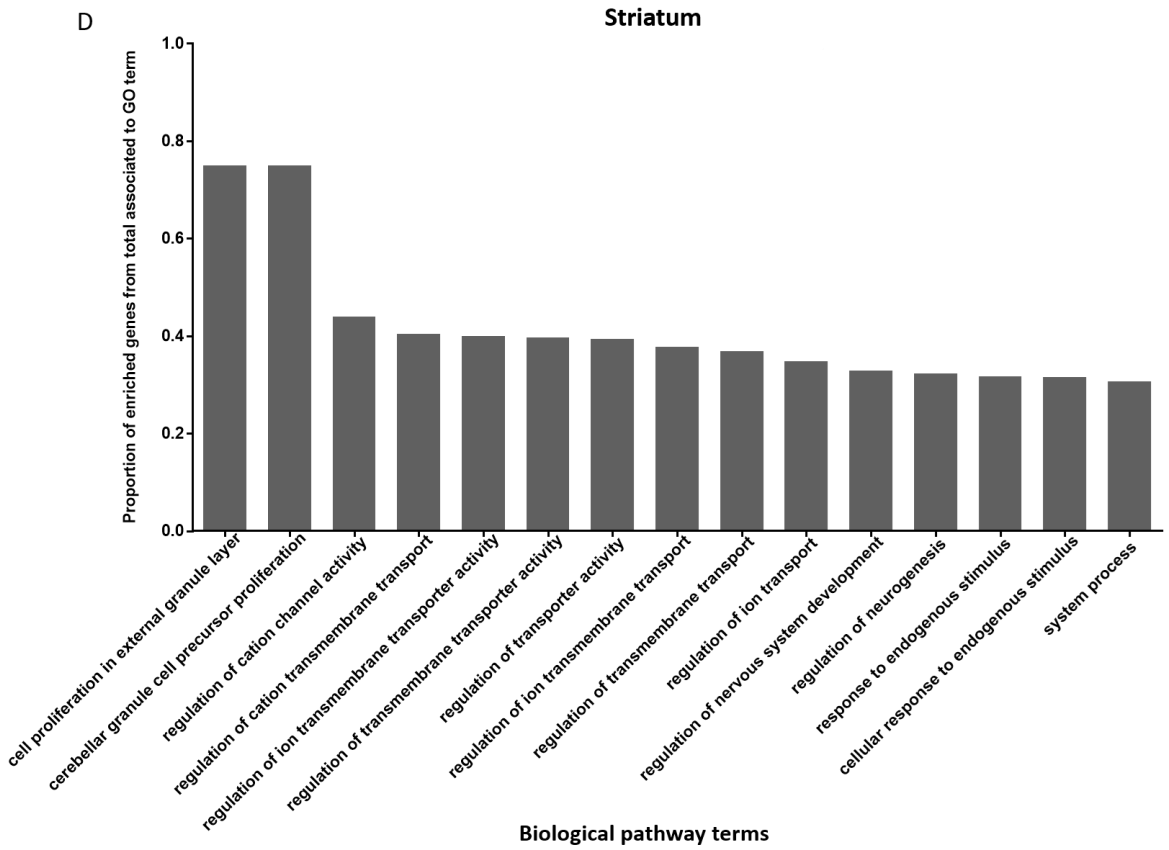


Figure 4.25 Biological pathway GO terms significantly enriched in the differentially expressed genes in *Kptn*$^{-/-}$ compared to reference set of total genes expressed in the tissue. All the significantly enriched biological pathways associated with the differentially expressed genes from the *Kptn*$^{-/-}$ brain tissues plotted as a proportion of the total number of genes associated with that pathway. Due to a high number of biological pathways associated with cerebellum (**A**) and prefrontal cortex (**B**), the terms are displayed as numbers (with pathways listed in Table 7), while the biological pathways in the hippocampus (**C**) and striatum (**D**) are displayed fully.

Table 7. Biological pathways enriched in *Kptn*$^{-/-}$ cerebellum (cereb.) and prefrontal cortex (PFC), with the number labels (N) used in Fig.4.25.

Tissue	N	Biological pathway
Cereb.	1	antigen processing and presentation of exogenous peptide antigen via MHC class I
	2	clustering of voltage-gated potassium channels
	3	antigen processing and presentation of exogenous peptide antigen via MHC class II
	4	regulation of systemic arterial blood pressure by baroreceptor feedback
	5	antigen processing and presentation of exogenous peptide antigen
	6	establishment or maintenance of transmembrane electrochemical gradient
	7	sodium ion export
	8	sodium ion export from cell
	9	neuronal ion channel clustering
	10	cellular potassium ion homeostasis
	11	cellular sodium ion homeostasis
	12	antigen processing and presentation of exogenous antigen
	13	antigen processing and presentation of peptide antigen via MHC class II
	14	antigen processing and presentation of peptide or polysaccharide antigen via MHC class II
	15	antigen processing and presentation of peptide antigen
	16	regulation of the force of heart contraction
	17	regulation of calcineurin-NFAT signaling cascade

	18	antigen processing and presentation of peptide antigen via MHC class I
	19	humoral immune response mediated by circulating immunoglobulin
	20	calcineurin-mediated signaling
	21	calcineurin-NFAT signaling cascade
	22	sodium ion homeostasis
	23	antigen processing and presentation
	24	regulation of sodium ion transport
	25	regulation of sodium ion transmembrane transport
	26	import into cell
	27	immunoglobulin mediated immune response
	28	B cell mediated immunity
	29	positive regulation of adaptive immune response
	30	sodium ion transmembrane transport
	31	signal release from synapse
	32	neurotransmitter secretion
	33	neurotransmitter transport
	34	presynaptic process involved in chemical synaptic transmission
	35	regulation of metal ion transport
	36	regulation of neurotransmitter levels
	37	sodium ion transport
	38	homophilic cell adhesion via plasma membrane adhesion molecules
	39	regulation of calcium ion transport

	40	synapse assembly
	41	calcium ion transmembrane transport
	42	regulation of cation transmembrane transport
	43	adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains
	44	adult behavior
	45	lymphocyte mediated immunity
	46	calcium ion transport
	47	cell-cell adhesion via plasma-membrane adhesion molecules
	48	adaptive immune response
	49	positive regulation of immune response
	50	divalent inorganic cation transport
	51	divalent metal ion transport
	52	activation of immune response
	53	regulation of ion transmembrane transport
	54	synapse organization
	55	regulation of transmembrane transport
	56	innate immune response
	57	regulation of membrane potential
	58	cellular metal ion homeostasis
	59	regulation of ion transport
	60	regulation of immune response

	61	metal ion homeostasis
	62	metal ion transport
	63	synaptic signaling
	64	trans-synaptic signaling
	65	anterograde trans-synaptic signaling
	66	chemical synaptic transmission
	67	cell-cell adhesion
	68	cellular cation homeostasis
	69	cellular ion homeostasis
	70	cation homeostasis
	71	cell adhesion
	72	immune response
	73	biological adhesion
	74	defense response
	75	cation transport
	76	ion transmembrane transport
	77	regulation of immune system process
	78	neuron development
	79	neuron projection development
	80	ion transport
	81	response to external stimulus
	82	immune system process

	83	neuron differentiation
	84	regulation of transport
	85	regulation of response to stimulus
	86	cell communication
	87	signaling
	88	single organism signaling
	89	multicellular organismal process
	90	single-multicellular organism process
PFC	1	antigen processing and presentation of endogenous peptide antigen via MHC class I via ER pathway
	2	antigen processing and presentation of endogenous peptide antigen via MHC class I via ER pathway, TAP-dependent
	3	D-aspartate transport
	4	D-aspartate import
	5	canonical Wnt signaling pathway involved in osteoblast differentiation
	6	antigen processing and presentation of exogenous peptide antigen via MHC class II
	7	antigen processing and presentation of exogenous peptide antigen
	8	antigen processing and presentation of endogenous peptide antigen via MHC class I
	9	antigen processing and presentation of exogenous antigen
	10	T cell mediated cytotoxicity
	11	antigen processing and presentation of peptide or polysaccharide antigen via MHC class II
	12	negative regulation of potassium ion transport

	13	basement membrane organization
	14	antigen processing and presentation of peptide antigen
	15	antigen processing and presentation of peptide antigen via MHC class I
	16	positive regulation of adenylate cyclase activity
	17	epithelial cell morphogenesis
	18	membrane repolarization
	19	positive regulation of cAMP biosynthetic process
	20	positive regulation of lyase activity
	21	negative regulation of TOR signaling
	22	positive regulation of cAMP metabolic process
	23	regulation of adenylate cyclase activity
	24	leukocyte mediated cytotoxicity
	25	embryo implantation
	26	cell killing
	27	positive regulation of cyclic nucleotide biosynthetic process
	28	actin-mediated cell contraction
	29	regulation of phospholipid metabolic process
	30	membrane depolarization
	31	antigen processing and presentation
	32	regulation of potassium ion transport
	33	cAMP biosynthetic process
	34	regulation of cAMP biosynthetic process

	35	response to cAMP
	36	regulation of cAMP metabolic process
	37	cAMP metabolic process
	38	regulation of nucleotide biosynthetic process
	39	regulation of purine nucleotide biosynthetic process
	40	cyclic purine nucleotide metabolic process
	41	cyclic nucleotide biosynthetic process
	42	response to corticosteroid
	43	response to purine-containing compound
	44	lymphocyte mediated immunity
	45	response to organophosphorus
	46	circadian rhythm
	47	transforming growth factor beta receptor signaling pathway
	48	response to glucocorticoid
	49	extracellular structure organization
	50	extracellular matrix organization
	51	adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains
	52	cyclic nucleotide metabolic process
	53	response to transforming growth factor beta
	54	cellular response to transforming growth factor beta stimulus
	55	regulation of purine nucleotide metabolic process

	56	single-organism behavior
	57	regulation of blood circulation
	58	regulation of metal ion transport
	59	angiogenesis
	60	muscle system process
	61	response to decreased oxygen levels
	62	regulation of angiogenesis
	63	muscle contraction
	64	response to hypoxia
	65	adaptive immune response
	66	leukocyte mediated immunity
	67	response to oxygen levels
	68	regulation of transporter activity
	69	leukocyte migration
	70	cellular carbohydrate metabolic process
	71	rhythmic process
	72	positive regulation of cell migration
	73	positive regulation of cell motility
	74	response to steroid hormone
	75	transmembrane receptor protein serine/threonine kinase signaling pathway
	76	response to drug
	77	response to extracellular stimulus

	78	positive regulation of locomotion
	79	positive regulation of cellular component movement
	80	behavior
	81	response to hormone
	82	signal release
	83	response to peptide hormone
	84	blood vessel morphogenesis
	85	response to inorganic substance
	86	response to nutrient levels
	87	circulatory system process
	88	response to peptide
	89	divalent inorganic cation transport
	90	regulation of cell motility
	91	regulation of cell migration
	92	divalent metal ion transport
	93	regulation of cellular component movement
	94	positive regulation of immune response
	95	blood vessel development
	96	response to organic cyclic compound
	97	anterograde trans-synaptic signaling
	98	chemical synaptic transmission
	99	regulation of locomotion

	100	synaptic signaling
	101	trans-synaptic signaling
	102	vasculature development
	103	regulation of ion transport
	104	G-protein coupled receptor signaling pathway
	105	cardiovascular system development
	106	positive regulation of immune system process
	107	response to organonitrogen compound
	108	metal ion transport
	109	response to nitrogen compound
	110	response to endogenous stimulus
	111	anatomical structure formation involved in morphogenesis
	112	positive regulation of cell differentiation
	113	immune response
	114	positive regulation of phosphorus metabolic process
	115	positive regulation of phosphate metabolic process
	116	response to lipid
	117	cellular response to endogenous stimulus
	118	defense response
	119	brain development
	120	negative regulation of cell death
	121	biological adhesion

	122	system process
	123	regulation of immune system process
	124	circulatory system development
	125	secretion
	126	cell adhesion
	127	response to abiotic stimulus
	128	positive regulation of multicellular organismal process
	129	response to external stimulus
	130	regulation of programmed cell death
	131	localization of cell
	132	regulation of multicellular organismal process
	133	positive regulation of developmental process
	134	regulation of localization
	135	cell motility
	136	cell migration
	137	regulation of cell differentiation
	138	response to oxygen-containing compound
	139	response to organic substance
	140	positive regulation of response to stimulus
	141	regulation of apoptotic process
	142	regulation of cell death
	143	response to chemical

	144	cellular response to chemical stimulus
	145	cell-cell signaling
	146	cellular response to organic substance
	147	regulation of transport
	148	movement of cell or subcellular component
	149	regulation of phosphorus metabolic process
	150	regulation of phosphate metabolic process
	151	immune system process
	152	regulation of developmental process
	153	regulation of multicellular organismal development
	154	signaling
	155	single organism signaling
	156	nervous system development
	157	cell communication
	158	regulation of response to stimulus
	159	cell surface receptor signaling pathway
	160	regulation of molecular function
	161	regulation of cell communication
	162	regulation of signaling
	163	cell differentiation
	164	regulation of biological quality
	165	single-multicellular organism process

	166	system development
	167	animal organ development
	168	cellular developmental process
	169	multicellular organismal process
	170	response to stress
	171	positive regulation of biological process
	172	signal transduction
	173	response to stimulus
	174	developmental process
	175	anatomical structure development
	176	single-organism developmental process
	177	positive regulation of cellular process
	178	multicellular organism development
	179	cellular response to stimulus
	180	biological regulation
	181	regulation of biological process

4.3 Discussion and future directions

By using a series of behavioural and cognitive tests, testing overlapping behavioural and cognitive domains, combined with morphometric brain analyses, I successfully recapitulated in the *Kptn*$^{-/-}$ mice the main phenotypes observed in the patients with *KPTN*-related disorder, namely macrocephaly, anxiety, hyperactivity, and cognitive impairment. Moreover, I gained further insight into the underlying processes affected by *Kptn* deficiency.

4.3.1 Kptn$^{-/-}$ mice exhibit behavioural abnormalities and hippocampus-dependent memory impairment

To test for behavioural abnormalities associated with the disorder, I tested the mice in the open field under dim rather than bright light (stressor), in order to capture the general activity and locomotion of mice in a less stressful environment (dim light). Under dim light, the *Kptn* mutant mice spent more time moving and covered a larger overall distance than the controls, suggestive of hyperactivity-like phenotype, which parallels the hyperactivity phenotype observed in a subset of patients with *KPTN*-related disorder. When tested in the light/dark assay, the mice spent longer in the dark zone than *Ktpn*$^{-/-}$ mutants. Moreover, taking into account that *Kptn*$^{-/-}$ have increased activity, the fact that they spent significantly longer in the dark zone, which is a smaller area than the light zone, implies that the performance of the mutants in this test is due to increased anxiety and not due to other confounding effects such as increased overall activity.

The results of several cognitive assays strongly point to hippocampal function being affected in *Kptn*$^{-/-}$ adult mice. *Kptn*$^{-/-}$ mice displayed impaired performance in cognitive tasks that rely on hippocampal function, such as the Barnes maze and social recognition assay, while there was no difference in performance between genotypes in pairwise discrimination task, which is hippocampus-independent task. Because the male wildtype controls did not habituate correctly to the initial stimuli in acquisition phase, I was not able to analyse males in the SOR assay. However, no sexual dimorphism has been reported in the patients, nor was there sexual dimorphism in our brain morphology data. For these reasons, I chose to test the females instead, and detected 24h memory impairment in *Kptn*$^{-/-}$ mice (Day 2), as they

showed no significant bias for either familiar or unfamiliar stimuli. In subsequent experiments, I used only males in order to reduce number of animals used.

In the Barnes maze, testing spatial memory, both genotypes successfully completed initial training sessions (T1-2) taking the same time to approach the escape hole, without a difference in total errors made. On 24h probe, *Kptn*$^{-/-}$ mice spent an equivalent time near the target zone as the wildtypes, indicating no deficit in memory of the location of the target. However, *Kptn*$^{-/-}$ showed impaired memory retention when tested after 72h (during 72h probe session), indicating that spatial, hippocampus-dependent memory is affected in the mutant *Kptn* mice.

When the escape box was moved in second part of training (training 2), *Kptn*$^{-/-}$ identified the escape box with an equivalent speed (primary latency) to the wildtypes, but made significantly more errors, measured as frequency of visits to other holes, before going into the escape box (which terminates the session). The fact that the *Kptn*$^{-/-}$ mice displayed a primary latency reduction, equivalent to wildtypes, over this training period, as well as a significant reduction from the first day of initial training to last day of training N2, implies *Kptn*$^{-/-}$ mice learned where the escape box is. The difference in total errors could be indicative of the *Kptn*$^{-/-}$ mice locating the target, but not going into the escape box straight away. Therefore, as a result, despite having learned the association between spatial cues and the escape location and identifying it, the total errors may be reflective on an increased exploratory behaviour in the mice during training N2 or reduced motivation to go into the escape box as soon as it is found. The performance in training N2 requires an inhibition of previously learned information and the ability to learn the new location. Performance in this learning task is associated with cognitive flexibility and is dependent on orbitofrontal region of prefrontal cortex (OFC) (Brigman et al. 2013; Chudasama and Robbins, 2003). It, therefore, could be possible that *Kptn* deficiency affects cognitive flexibility. Unfortunately, reversal was not done after pairwise discrimination task to be able to confirm this.

When interpreting the Barnes maze results, it is important to be mindful of potential non-cognitive factors that may affect the results of this assay, such as anxiety and activity levels (Wolfer et al., 1998). It is formally possible that results observed are not solely due to

memory impairment but are affected by the mutant's greater responsiveness to stress due to their increased anxiety and overall locomotion. Studies have shown that stress and anxiety have varying effects on hippocampus function depending on the task under consideration, as well as the duration of stress (Hölscher, 1999; Luine et al., 1996; Miyakawa et al., 2001). Stress has been shown to negatively affect hippocampus-dependent performance in the water maze (Hölscher, 1999; Harrison et al., 2009). Interestingly, however, Harrison et al. (2009) have shown that spatial memory is inversely correlated with corticosterone levels in the water maze, but not Barnes maze, suggesting the Morris water maze performance is more affected by stress than the Barnes maze. Moreover, the Barnes maze has been shown to be less anxiogenic than the water maze (Harrison et al., 2009). It is therefore unlikely that my results are purely non-cognitive, especially since the results are consistent across two assays (Barnes maze and social recognition).

After identifying *Kptn*$^{-/-}$ hippocampus-dependent memory impairment phenotypes with social recognition and Barnes maze assays, I tested the animals with the object discrimination paradigm that combines object displacement (OD) and novel object recognition (NOR) assays, which I piloted in Chapter 3. Both genotypes performed well in the paradigm, and I was unable to detect impairment in the mutants in both the OD part of the assay nor the NOR part. There are several likely reasons for this: Firstly, the lack of phenotype in the OD part of the paradigm, which assesses spatial hippocampus-dependent memory, indicates that the mutant mice do not have 1h spatial memory impairment. This suggests that the memory defect seen in *Kptn* mice is sensitive to the interval time between acquisition and training, in agreement with the results from SOR and the Barnes maze. During piloting (Chapter 3), in order to optimise the paradigm to be suited for a more high-throughput approach, I shortened the time between acquisition training and testing from 24h to 1-2h. Testing the *Kptn*$^{-/-}$ mice with a longer time interval should confirm whether this assumption. Secondly, although the critical roles of the PFC and parahippocampal regions in the NOR task have been well established, there is controversy, due to the great variety in testing methodology between studies, regarding the association of the hippocampus in the NOR task (Bussey et al., 1999; Brown and Aggleton 2001; Banks et al. 2012; Hammond et al., 2004). The lack of phenotype in the *Kptn* mutants in the NOR assay is in line with studies

that have shown that hippocampus involvement in this assay is reliant on complex environmental conditions with spatial and contextual cues, which was not the case in my paradigm (Winters et al. 2004; Forwood et al. 2005). Lastly, in this OD/NOR paradigm I excluded a number of animals from the analysis because of their suboptimal overall investigation of the objects (Chapter 2, section 3.2.3) Because of the high dropout rate of the mice, I decided not to incorporate it as part of the cognitive screen for characterizing multiple novel DD mouse lines (described in Chapter 5).

The pairwise discrimination (PD) task (Chapter 2, section 2.3.6; Chapter 3) assesses reward-based associative learning, associated with the dorsal striatum (Brigman et al. 2013; Hamilton and Brigman 2015). Both *Kptn* mutant and wildtypes completed a series of procedural training stages, confirming there were no visual, motor, or motivational issues in these mice tested in the platform. When tested in the PD task, both genotypes were able to complete the PD criteria of 80% correct nose-pokes of the CS+ stimuli, in 60 minutes, over two consecutive days, suggestive of no memory and learning impairment in both genotypes. Learning was confirmed by swapping CS+ and CS- images (new CS+ image was previously not associated with the reward). Both genotypes significantly dropped in their performance in this case, indicating a strong association to the original CS+. The results from PD are suggestive of intact striatum and hippocampal-independent learning in the mutants, however care must be taken when interpreting the results as it is possible that subtle differences were missed due to lower number of animals tested in the PD task than the estimated sample size (Chapter 2, section 2.3.6; Chapter 3, section 3.3.2).

4.3.2 Kptn deficiency contributes to a progressive global macrocephaly phenotype

A consistent phenotype across all patients was macrocephaly. Interestingly, the age of onset of this phenotype varied between the Amish (9 probands, all macrocephalic at birth) and the Estonian cohorts (2 siblings, born with normal head parameters and display macrocephaly in adulthood). Our analysis revealed that *Kptn*$^{-/-}$ mice display severe macrocephaly in adulthood (measured in here by X-ray cephalometric analysis, MRI, and

histological morphometrics), but are not macrocephalic at birth (P0). This indicates that the deficit in *Kptn* causes progressive macrocephaly in our mouse model. Importantly, P0 stage in the mice is equivalent of that of gestation week 23–32 in the humans (pre-birth), as discussed in Chapter 1. This may imply that the progressive form of macrocephaly in the mouse has comparable relative timing to the Amish patients, who are born macrocephalic. In order to elucidate this further, several postnatal mouse time points should be collected and the brain morphology analysed in a similar manner as was done with P0 brains. If the mice display macrocephaly at P7-10, which is equivalent to infant stage in humans, then this finding would be in accordance with the Estonian siblings who have progressive macrocephaly.

Observing phenotypic manifestations of neurodevelopmental disorders postnatally is not uncommon. For example, both autism and Rett syndrome are associated with normal initial development and an onset of pathology occurring at later postnatal stages, associated with aberrant cell growth, synaptic maturation, connectivity, and stabilization (Zoghbi, 2003; Rubenstein, 2010). Aberrant increase in brain volume is most commonly associated with autism spectrum disorder (ASD), where the brain volume and head circumference of patients' increases during early postnatal stages and subsequently decelerates at later stages (Piven et al., 1995; Carper et al., 2006; Amaral et al., 2008). Since *KPTN* is involved with actin cytoskeleton, it would be important to look at whether axonal projections and overall connectivity is affected in the mutants during development. This could be done by performing neural outgrowth assays at different embryonic ages to assess KPTN involvement in neuromorphogenesis (Radio & Mund, 2008). In adult mice, our results indicate that the mutants have global macrocephaly, with most brain regions significantly enlarged, and both grey and white matters affected. Interestingly, only two structures were not enlarged in the mutant adult brain – the hippocampus and the ventricular zone. The hippocampus was the same size as the wildtypes, and the ventricular zones were reduced. In the P0 brains on the contrary, none of the parameters were affected in the mutants, with the exception of an observed reduction in the hippocampus and an increase in the internal capsule areas.

The reduction of the ventricular size in the adult mutants suggests a potential compensation mechanism of the cavities to accommodate the enlarged size of the overall brain. The finding of a normal sized hippocampus in the adult *Kptn*$^{-/-}$ mice is somewhat unexpected, given my results pointing to cognitive deficits involving hippocampus-dependent processes. The hippocampus plays a role in the multi-modal sensory integration system in the CNS, by being both upstream and downstream of cortical areas of the brain (Sweatt, 2004). Since the hippocampus sends projections to the cortex and receives information from the various cortical areas, having the cortex enlarged while the hippocampus remains unchanged in size may lead to aberrant cognitive pathways. It has been shown that hippocampus is significantly reduced in people with mild cognitive impairment (MCI) when compared to normal cognition controls (Schuff et al., 2001). Therefore, even though the hippocampus is not reduced in size in adult *Kptn*$^{-/-}$ mice relative to the controls, the reduced size ratio of the hippocampus to cortex may affect the cognitive processes, similarly that a reduction of hippocampal size in a normal sized brain would have. Moreover, *Kptn* mutants display a reduction of hippocampus size when compared to controls at birth (P0), while other brain structures apart from internal capsule remain within normal size range. This suggests that the abnormalities in the hippocampus originate prenatally, potentially causing a delay in its development, which in turn is associated with aberrant function in adulthood despite the normal size of the hippocampus in the adult brain of *Kptn* mutant mice.

Our data suggest that the increase in internal capsule starts prenatally. Internal capsule is a white matter structure linking cerebral cortex with other parts of the brain, with a role in influencing higher cognitive function (Rousset et al., 2006). The abnormal signal intensity in the internal capsule from MRI studies in infants has been shown to be a good predictor of aberrant neurodevelopmental outcomes (Rutherford et al., 1998). Moreover, damage in white matter structures, such as internal capsule, have been associated with cognitive impairment in Alzheimers Disease (Duan et al., 2006).

4.3.3 Kptn deficient transcriptome

Combining the results from morphometric brain analyses and cognitive testing, an interesting pattern emerges, associated with the hippocampus. Firstly, several independent cognitive tests identified hippocampus-dependent memory impairment in the *Kptn*$^{-/-}$, but no memory impairment was identified when *Kptn*$^{-/-}$ were tested in hippocampus-independent cognitive tasks. Secondly, based on morphometric analyses hippocampus has a significantly reduced size in *Kptn*$^{-/-}$ at P0 relative to the wildtypes, implying a pre-natal developmental defect. However, in adult *Kptn*$^{-/-}$ mice, hippocampus and ventricle zones are the only brain structures not enlarged. Therefore, to shed more light on the effect of *Kptn* deficiency on the postnatally and detect potential mechanisms that may be contributing to the adult phenotypes described, I analysed the transcriptome of hippocampus, striatum, prefrontal cortex and cerebellum in *Kptn*$^{-/-}$ adult mice. The four brain structures were chosen for transcriptomic analysis due to the morphometric and cognitive phenotyping data, with the first three neuroanatomical structures associated with the cognitive testing described earlier in this chapter and all the structures apart from the hippocampus are enlarged in the *Kptn*$^{-/-}$ adults based on our morphometric data.

The transcriptome of all four neuroanatomical structures analysed were affected by *Kptn* deficiency (compared to the expression levels of wildtype controls), with many differently expressed (DE) genes identified in each tissue. The largest number of significantly differentially expressed genes was in the striatum, where 3896 genes were differentially expressed, followed by hippocampus with an affected expression in 1315 genes. Based on the gene set enrichment analysis, the DE genes identified are enriched in many pathways, including those associated with neuronal development, brain function, and behaviour. Interestingly, mTORC1 signalling was enriched in the DE genes only in the hippocampus and prefrontal cortex. Moreover, 'ribosome' was the only significantly enriched KEGG pathway associated with the DE genes in the hippocampus, with 85 DE genes associated with ribosome function (68% of the total proportion of functional genes associated with this KEGG pathway). As outlined here, mTOR signalling is an important regulator of ribosomal biogenesis and has a role in gene transcription regulation in part through the activation of multiple transcription factors (Mayer and Grummt, 2006; Passtoors et al., 2012; Wang et al.,

2011; Yokogami et al., 2000). Therefore, the enrichment of DE genes associated with ribosome function may be an indication of an increased activity of the mTORC1 in the hippocampus, which could be one of the contributing factors to the functional deficit detected in the hippocampus. This is in accordance with recent finding by Wolfson et al. (2017) of the role of *KPTN* as a negative regulator of mTORC1 signalling, as part of a KICSTOR complex, which suggests *Kptn* deficiency should lead to mTORC1 signalling hyperactivity in the brain. This was previously confirmed in another protein of the KICSTOR complex, *Szt2,* where deficiency in *Szt2* was associated with an increase in mTORC1 signalling in the brain (Peng et al., 2017; Frankel et al., 2009). To my knowledge, no study to date has tested mTOR activity levels in *Kptn*$^{-/-}$ mice.

Interestingly, when gene enrichment was analysed for upregulated DE genes, there was enrichment for negative regulation of mTORC1 in the hippocampus and prefrontal cortex. In the hippocampus, this was the only pathway enriched in the upregulated DE genes. This could suggest that the loss of negative regulation of mTORC1 in these structures, due to deficiency in *Kptn,* could potentially lead to upregulation of other mTOR negative regulators in these brain regions, as part of the feedback loop mechanism. However, despite this potential compensatory mechanism, the full function of hippocampus was not rescued, as demonstrated by the cognitive results in the *Kptn*$^{-/-}$. Increase in mTOR signalling due to loss-of-function of negative regulators, such as *Pten* and *Tc1*, have been previously associated with increased neuron soma size and macrocephaly (Ehninger et al., 2008; Kwon et al., 2003). Therefore, it is possible that an increase of activity of mTORC1 signalling in the *Kptn*$^{-/-}$ brain could be contributing to the macrocephaly phenotype in this model as well. One hypothesis could be that the compensatory mechanism of increasing expression of other negative mTOR regulators could be one of the processes that are responsible for the lack of volume change in the adult hippocampus relative to the other brain structures. Distinguishing this hypothesis from one in which cell death counteracts the overgrowth phenotypes will be the subject of further study in the lab.

In order to confirm the involvement of mTOR signaling in *Kptn* deficient brain and to assess the distribution of mTOR activity between different brain regions, the activity of mTOR signalling in *Kptn*$^{-/-}$ mice should be assessed by investigating the difference in

phosphorylation levels of the mTOR downstream targets, such as S6K and 4E-BP, between controls and mutants, by western blotting (Sharma et al., 2010; Way et al. 2012). If this confirms the dysregulation of the mTORC1 pathway, it opens an exciting opportunity for therapeutic intervention, by targeting the hyperactive pathway with rapamycin, a clinically approved drug, to rescue some of the phenotypes associated with the disease.

4.3.4 Future directions

4.3.4.1 Phenotypic rescue with rapamycin treatment

Rapamycin has been extensively and successfully used to target many of the mTORC1 components, leading to increased lifespan and rescuing of cognitive deficits in neurodevelopmental and neurodegenerative mouse models (Ehninger et al., 2008; Harrison et al., 2009; Johnson et al., 2013; Spilman et al., 2010). *Tsc2*$^{+/-}$ mice showed learning deficits in three hippocampus-dependent tasks (Morris water maze, eight-arm radial maze, context discrimination) and had hyperactive mTOR signalling in the hippocampus, suggesting that enhanced mTOR signalling leads to these deficits. Suppression of the aberrant signalling with rapamycin in these mice, rescued the cognitive deficits associated with this mouse model (Ehninger et al., 2008). Moreover, rapamycin treatment reduced brain weight (*Tsc1*) and reversed neuronal soma enlargement (*Pten*) in the conditional mouse models, with minimal effects on normal brain growth and function in wildtype mice (Ehninger et al., 2008; Kwon et al., 2003).

Therefore, once hyperactivity of the mTOR pathway in *Kptn*$^{-/-}$ confirms my hypothesis that treating *Kptn*$^{-/-}$ mutant mice with rapamycin from birth could rescue either the macrocephaly, some aspects of behavioural and cognitive impairments identified in the adults, or possibly both. Cognitive rescue could be tested in the adults using the light/dark and social recognition assays, supported by morphometric and transcriptomic brain analysis. Furthermore, by identifying the postnatal stage at which the progressive macrocephaly is first manifested in the *Kptn*$^{-/-}$ and performing rapamycin treatment after this time point, could allow to dissociate the macrocephaly phenotype from the cognitive

defects, and assess if these are independent. If these treatments show promise, it will be crucial to consider how this might be moved into the clinic to potentially treat patients with *KPTN*-related syndrome.

4.3.4.2 Macrocephaly mechanism

An important question to answer is at what postnatal stage the macrocephaly phenotype first appears. Our morphometric data indicate that the *Kptn*$^{-/-}$ mice are born with normal growth parameters and are severely macrocephalic in adulthood. These data, in the context of the recently identified role of *KPTN* as a negative regulator mTORC1 pathway, which is involved in cell growth and cell cycle progression, support the hypothesis that *Kptn* may be involved in the regulation of brain growth during a critical growth window (Szulc et al., 2015; Zhang et al., 2005). As outlined in Chapter 1, the peak brain growth spurt in mice is at P7-P10, with brain reaching 95% of its volume by P20 (Bockhorst et al., 2008; Keshavan et al., 2002). Measuring the brain volume before this growth window (P7-P20) and after should elucidate when the mice become significantly macrocephalic. One can then analyse proliferation and apoptosis rates at this stage of initial macrocephaly manifestations to elucidate the mechanisms responsible for the increased growth. One possibility is that the deficiency in Kptn is causing a reduction in apoptosis at the peak of brain growth, leading to an abnormal increase in brain size. Alternatively, *Kptn* mutants may experience a longer period of brain growth beyond the normal plateau around P20 in wildtype mice.

In the adult *Kptn* mutant mice, we were able to detect increased number of proliferating cells in the hippocampus, increased cell counts in the cortex, and no increase in apoptosis in the cortex. This suggests that increased proliferation rates could be contributing to the macrocephaly phenotype. However, because these assays were performed only in adults, likely after the overgrowth has occurred, the observations on cell cycle and apoptosis may not be relevant to the underlying mechanism acting during the acquisition of the phenotype. It will, therefore, be critically important to analyse changes during the brain growth period identified before drawing any firm conclusions on the mechanism involved.

4.3.4.3 Elucidating the roles of neurogenesis in the adult

The increased proliferation rate identified in the hippocampus of *Kptn*$^{-/-}$ mice could be suggestive of an increased rate of adult neurogenesis in the hippocampus. Adult neurogenesis is a process through which new neurons are generated from a pool of progenitor cells, and is believed to be restricted to the hippocampus (Hill et al., 2015; Sahay et al., 2011). However, some controversial evidence points to the presence of neurogenesis in other areas of the brain such as the neocortex and striatum (Gould et al., 2007; Lledo et al., 2006; Bernier et al., 2002). Adult neurogenesis persists throughout life and the rate of proliferation, maturation, and survival of the progenitors is affected by environmental cues such as exercise, antidepressants, stress, and age (Hill et al., 2015; Ishikawa et al., 2016; Kitamura and Inokuchi, 2014; Sahay et al., 2011).

Co-staining the adult *Kptn*$^{-/-}$ mouse brain sections with cell-type specific markers and markers of proliferation will identify the cell types that are undergoing cell division. If adult *Kptn*$^{-/-}$ mice have increased neurogenesis in the hippocampus in adults, it may further explain why mutant mice have specific impairment in the hippocampus-dependent and not hippocampus-independent tasks. The role of neurogenesis in cognition is complex. On the one hand, studies outline the importance of neurogenesis in neural circuit plasticity and report positive effects of neurogenesis on learning and cognitive plasticity (Sahay et al., 2011). Neurogenesis has been implicated in enhanced pattern separation, and reduced rates have been associated with age-dependent decline and neurodegenerative disorders (Aimone et al., 2014; Bishop et al., 2010; Chen et al., 2008; Zhao et al., 2008). On the other hand, increased neurogenesis in hippocampus has also been linked to loss of hippocampus-dependent recent memories, but not remote (long-term) and hippocampus-independent memories (Epp et al, 2016; Ishikawa et al., 2016; Kitamura et al., 2009).

Further cognitive tests associated with cognitive flexibility, such as the five-choice serial-reaction test and operant reversal learning, could be used to elucidate this further (Bussey et al. 2012; Semenova, 2012). The five-choice serial-reaction test is similar to the human continuous performance tasks and tests attention, motor impulsivity, decision-making abilities, and cognitive flexibility in the rodents (Mar et al., 2013; Semenova, 2012).

4.3.5 Concluding remarks

KPTN-syndrome, a novel developmental disorder, was identified in homozygous and compound heterozygous patients with loss-of-function mutations in the *KPTN* gene. Here I report that *Kptn*$^{-/-}$ mice phenocopy the main behavioural and cognitive features of the human condition namely anxiety, cognitive impairment, and macrocephaly, thus successfully characterising a novel developmental disorder associated with intellectual disability. Through the developmental and molecular experiments done in this thesis work, I gained preliminary insights into the underlying mechanisms associated with the disorder, but further work needs to be done to fully elucidate the roles of KPTN in the brain. The work outlined here, in the light of recent insight into *KPTN* role in the mTOR pathway, offers promising leads into potential treatments for this disorder.

Chapter 5. Behavioural and Cognitive Neurodevelopmental Disorder screen

5.1 Introduction

5.1.1 Decipher Developmental Disorders (DDD) study

As discussed in Chapter 1, many neurodevelopmental disorders have a genetic cause, yet few affected children receive a genetic diagnosis (Fitzgerald et al., 2015; Wright et al., 2015). Successful diagnosis is further challenged by disorders that are not well characterised, have highly variable manifestations, and are hard to distinguish from other disorders phenotypically. Due to the clinical and genetic heterogeneity of neurodevelopmental disorders, there is a growing appreciation that using unbiased genome-wide approaches, such as genomic microarrays and next-generation sequencing (whole genome and exome), can dramatically improve diagnosis of neurodevelopmental disorders (Fitzgerald et al., 2015; Firth et al., 2011; Vries et al., 2005; Wright et al., 2015). It has been estimated that the per-generation mutation rate in humans is between 7.6×10^{-9} and 2.2×10^{-8}, such that, on average, a newborn child has 50 to 100 *de novo* mutations in their genome (Lynch et al., 2010; Roach et al., 2010; Vissers et al., 2010). Mutations occurring spontaneously in the germline persist in the population despite serious phenotypic consequences, such as intellectual disability (ID) and reduced fecundity, (Uher, 2009). Because ID negatively affects reproductive fitness, sporadic forms are more severe than the familial the more severe forms of are mainly sporadic, while the familial forms are milder (Vissers et al., 2016).

Deciphering Developmental Disorders (DDD) is a large-scale study that uses parent/child trio exome sequencing to identify causative mutations in the genomes of children with undiagnosed developmental disorders (Wright et al., 2015; Fitzgerald et al., 2015; McRae et al., 2017). The study uses a genotype-driven genome-wide diagnostic approach for identification of groups of patients with similar pathogenic variants from a large number of patients with diverse DDs (Wright et al., 2015; Fitzgerald et al., 2015; McRae et al., 2017). Moreover, this approach has the potential of expanding the phenotypic range of known DD conditions, by genetically characterizing patients lacking the clinical features used for phenotype-driven diagnosis (Wright et al., 2015; Fitzgerald et al., 2015; McRae et al., 2017).

The DDD study began in 2011 and had recruited 13,600 patients from all over UK and Ireland, through UK National Health Service (NHS) and Republic of Ireland genetics services by 2017. Children with severe undiagnosed DD and their parents were recruited, in order to maximise the chances of capturing highly penetrant monogenic conditions (Wright et al., 2015; Fitzgerald et al., 2015; McRae et al., 2017). Growth measurements, relevant family history, developmental milestones were recorded, pertinent pregnancy and neonatal parameters, and the detailed clinical phenotypes were noted using Human Phenotype Ontology terms. The median age of children assessed in the study was around 5.5 years old (Fitzgerald et al., 2015). The most prevalent phenotypes (of the first 1,133 children analysed) were intellectual disability or developmental delay (87%), cranial abnormalities (30%), seizures (24%), and congenital heart defects (11%) (Fitzgerald et al., 2015). The genetic ancestry of the recruited patients is predominantly (90%) of Northwest European ancestry (Fitzgerald et al., 2015). Overall the disorders in the affected patients were genetically and phenotypically heterogeneous.

De novo mutations (DNM) have been reported as a major cause of neurodevelopmental disorders by multiple studies, although the reported diagnostic yield varies between studies due to experimental design, while recessive inherited mutations are enriched in populations with frequent parental consanguinity (de Ligt et al., 2012; Firth et al., 2011; Hamdan et al., 2014; Musante and Ropers, 2014; Najmabadi et al., 2011; Rauch et al., 2012; Vissers et al., 2010; Wright et al., 2015). In the DDD cohort, d*e novo* variants have the highest diagnostic yield, and it was estimated that 42% of individuals in the cohort carry pathogenic DNMs in the coding sequences, with relatively equal representation of loss-of-function mutations and those resulting in altered protein function (Wright *et al.,* 2015; McRae et al., 2017).

5.1.2 Systemic behavioural testing paradigms

Although next-generation sequencing approaches play a critical role in identifying disease-associated mutations, animal and cellular modelling are essential to verify the candidate mutation, validate novel candidate genes, and elucidate pathophysiological mechanisms. Mouse models are a critical resource for modelling human disease and dissecting gene function. To fully understand the neurophysiological and neuroanatomical implications of a disorder, it is critical to apply a broad set of behavioural and cognitive tests to assess overlapping behavioural domains when modelling a new mouse model (Rogers et al., 1999; Tarantino et al., 2000). As outlined in previous chapters, rodent behaviour is susceptible to environmental influence, background strain, sex, and the overall experimental design. These present a particularly difficult challenge, not only to achieve sufficient sensitivity and consistency but also to maintain inter-operator and inter-lab reproducibility (McGoniglea and Rugger, 2014). Studies have tried to dissect these problems by assessing the contribution of different strains on behavioural outputs and/or test mice in different paradigms to explore the effect of differences in environment and experimental history on experimental outcomes (Contet et al., 2001; McIlwain et al., 2001; Rogers et al., 1997; Rogers et al., 1999; Tarantino et al., 2000; Võikar et al., 2001; Võikar et al., 2004). In line with this, there is a delicate balance to be reached between on the one hand, allowing for flexibility of experimental design, for example to tailor a study to fit a particular model in question, and on the other hand striving for a level of standardisation of methodology, to allow for comparisons across laboratories and studies, and to maximise reproducibility. Accurate and reproducible phenotype assessment is thus "*the jewel in the crown of genetic manipulation*" (Rogers et al., 1997).

In the late 1990s, the SHIRPA testing protocol was established by a multicentre consortium to screen mouse models in a systematic way using a wide range of behavioural and functional tests (Rogers et al., 1997; Rogers et al., 1999). More recently the Mouse Genetics Project (MGP) and the European Mouse Disease Clinic (EUMODIC) were established (late 2000s). These two large-scale phenotyping projects pioneered high-throughput production

and analysis of mouse knockout lines, provided valuable information on effective design of phenotyping platforms and overall optimal operation of such big initiatives, and provided resources for over 900 lines (Ayadi et al., 2012; White et al., 2013). Since then, new initiatives have been set up by the International Mouse Phenotyping Consortium (IMPC), including a new adult pipeline, aiming to characterize new knockout mouse strains and creating a genome-wide gene function catalogue (Meehan et al., 2017). These approaches aim to avoid inter-study discrepancies by testing multiple mouse lines without the inherent variation that would exist if these models were studied separately (Bucan and Abel, 2002; Fuchs et al., 2012; Laughlin et al., 2012; Marston et al., 2001; Wakana et al., 2009). These paradigms are powerful yet logistically challenging, as they are reliant on the ability to scale up phenotypic testing and ensure the tests are both rapid and sensitive to detect different behavioural and cognitive repertoires in multiple mouse lines.

5.1.3 The approach and aims

Although there are several efforts, outlined above, to characterize multiple mouse knockout lines, there is a relative lack of specialised high-throughput behavioural and cognitive approaches focusing on intellectual disability and other neurodevelopmental disorders. In this Chapter, I will describe a mouse phenotyping paradigm, designed to support a large, specialised functional screen that I applied to systematically characterise an initial four novel neurodevelopmental disorders associated with intellectual disability, identified by the DDD project. The four loss-of-function mouse lines under investigation, *Arid1b Setd5, Setd1a, Zmynd11*, have mutations in genes encoding chromatin remodelling factors (CRFs) (Fitzgerald et al., 2015). Due to the druggable nature of CRFs and their involvement in many types of cancer, there is a high level of investment for developing drugs targeting of making them an attractive set of proteins to work on (Filippakopoulos et al., 2010; Helin and Dhanak, 2013; Huether et al., 2014).

Since I was able to successfully recapitulate the main patient phenotypes in the homozygous loss-of-function *Kptn* mouse model (Chapter 4), my next aim was to assess whether the strategy used for modelling *Kptn* mice could be scaled up to characterize multiple novel

neurodevelopmental disorders associated with intellectual disability. The phenotyping was not intended to be an exhaustive characterisation of each mutant line. The strategy had three critical criteria: (1) It was required to operate in a standardised and relatively high-throughput manner, ensuring the testing paradigm doesn't last longer than 4-6 weeks to enable multiple lines to be tested a year, with (2) sufficient breadth to detect a variety of cognitive and behavioural defects, and (3) have the sensitivity to capture phenotypic differences between disorders under investigation. For this I employed the most robust techniques from the *Kptn* study, as well as an additional assay testing for features commonly associated with autism spectrum disorders (ASD). This paradigm will henceforth be referred to as BCND (Behavioural and Cognitive Neurodevelopmental Disorder) screen.

This pilot project lay the groundwork for a larger five-year project at the Sanger Institute, aimed at characterising a range of novel neurodevelopmental disorders arising from mutations identified in the DDD and Prenatal Assessment of Genomes and Exomes (PAGE) projects. The plans for the project include generation and phenotypic characterization of approximately 40 mouse lines, prioritising CRFs. Gene expression profiling of several brain regions from each line is also being carried out, as well as the generation and characterization of human induced pluripotent stem cell (hiPSC) lines harbouring the same mutations as the mouse models. This will aid biological validation of the causality of the mutations under investigation, provide insight into the underlying mechanisms for each disorder, as well as provide a large-scale platform for comparison between multiple neurodevelopmental disorders. When a set of robust pathological consequences is identified in a mutant mouse line, the possibility of reversing these phenotypes will be assessed, through inducible reversion of mutant alleles to wild-type. This proof of reversibility is a critical step towards devising therapeutic strategies.

The strategy for BCND was to contribute to the refinement and piloting of the experimental design and high-throughput nature of this larger project.

5.1.4 Candidate genes

5.1.4.1 Epigenetic modifications – histone lysine methylation and acetylation

As described in Chapter 1, more than 700 genes have been associated with ID and related cognitive disorders (Vissers et al., 2016). Despite the genetic heterogeneity, it is becoming apparent that ID can be dissociated into distinct modules of genes, functioning as part of a common pathway or complex and associated with interrelated phenotypes (van Bokhoven, 2011). Transcriptional regulators and chromatin remodelling factors (CRFs) have been shown to be enriched in neurodevelopmental disorders (Hamdan et al., 2014; Kleefstra et al., 2012; Kleefstra et al., 2014; McCarthy et al., 2014; Ronan et al., 2013; van Bokhoven, 2011; Vogel-Ciernia and Wood, 2014).

Epigenetic modifications are responsible for creating differences in tissue and cell type-specific gene expression (van Bokhoven, 2011). Chromatin modifiers have been shown to play crucial roles in brain development, as well as being involved in numerous types of cancer, suggesting a potential overlap between cognitive impairment and carcinogenesis (Vissers et al., 2016). The timing of the *de novo* mutations in CRFs is critical in directing whether the mutations will cause intellectual disability if they occur early in development, or cancer, if they occur later in life (Vissers et al., 2016).

The DNA and histone modifications, such as methylation and acetylation, regulate patterns of gene expression. DNA methylation leads to more stable long-term modifications of DNA accessibility, whereas histone modifications are more flexible and shorter term, causing changes in chromatin structure (Handy et al., 2012). There are dedicated proteins which function as writers, erasers, and readers of epigenetic tags, as well as proteins that act as chromatin remodelers (Fahrner and Bjornsson, 2014). Writers are responsible for placing the markers on particular regions of the genome and histone tails, which are recognised and interpreted by readers and counterbalanced by erasers which favour the opposite chromatin state to the writers. Chromatin remodelers usually act as part of larger protein complexes and are responsible for modifying chromatin architecture and thus transcriptional activity.

Histone acetylation at lysine residues in H3 and H4 tails promotes open chromatin structure and therefore most often associated with transcriptional activation, even though the open chromatin state may also give access to transcriptional repressors (Shogren-Knaak et al., 2006;). The reverse process of histone deacetylation is associated with inactive chromatin state (Nan et al., 1998). Histone deacetylase enzymes are themselves subject to regulation by acetylation, phosphorylation, and sumoylation, which in turn can affect their function, distribution, and protein-protein associations (Mellert and McMahon, 2009).

Histone lysine methylation is a more complex than acetylation because methylation sites can be associated both with transcriptionally permissive (euchromatin) and repressive chromatin (heterochromatin) (Martin et al., 2005). Moreover, lysine residues can be mono, di, or tri-methylated. Most histone lysine methyltransferases have Su(var)3-9, Enhancer of Zeste, Trithorax (SET) homology domain. SET domain containing proteins are broadly subdivided into seven subfamilies based on their structure - SUV39, SET1, SET2, EZ, RIZ, SMYD, and SUV4-20 families (as well as a few orphan members such as SET7/9 and SET8) (Dillon et al., 2005). SET1 methyltransferase members promote active chromatin state by methylating H3K4, whereas other methyltransferases can methylate several histone targets. Some of the methyl transferases also contain additional domains that allow them to bind to methylated DNA and other proteins.

Initially it was proposed that the only mechanism for removal of histone methylation is histone turnover, however, subsequently multiple histone demethylases have been identified that demethylate histone lysine methyl groups (Shi et al., 2004). Histone methylation is nevertheless less dynamic than acetylation and has been implicated in cellular memory of transcriptional states (Völkel and Angrand, 2007).

5.1.4.2 ARID1B (BAF250B and ELD/OSA1)

AT-rich interaction domain-containing protein 1B (ARID1B) (also known as *BAF250B* and *ELD/OSA1*), encodes a protein with four isoforms and is part of the SWI/SNF chromatin remodelling complex, where it functions in an opposing, mutually exclusive way to ARID1A

(Flores-Alcantar et al., 2011). ARID1B directly binds double-stranded DNA, and has been implicated in cell-cycle activation and progression (Flores-Alcantar et al., 2011; Yan et al.,2008). Knockdown and haploinsufficiency of ARID1B have been shown to delay cell cycle re-entry (Nagl et al., 2005; Sim et al., 2014).

ARID1B is ubiquitously expressed in most tissues. In the adult brain, it is predominantly expressed in neuronal cell bodies in the cortex, with moderate expression in neurons and glial cells in the hippocampus, and is most highly expressed in the Purkinje cells in the cerebellum (Uhlén et al., 2015). Although *ARID1B* is predominantly expressed in differentiated cells, it has also been implicated to have a role in the developing brain (Flores-Alcantar et al., 2011).

Deficits in *ARID1B* have been identified as a frequent cause of ID (Fitzgerald et al., 2015 Hoyer et al., 2012). Moreover, haploinsufficiency in *ARID1B* have been associated with abnormalities in corpus callosum, developmental delay and speech impairment, as well as implicated in autism spectrum disorder and as a major cause of Coffin-Siris syndrome (Backx et al., 2011, 167; Halgren et al., 2011; O'Roak et al., 2012; Santen et al., 2012; Tsurusaki et al., 2012; Vals et al., 2014). Two studies, published during the writing of this thesis, reported that *Arid1b*$^{+/-}$ mice display anxiety-like behaviour, social deficits, and growth impairment (Celen et al., 2017; Shibutani et al., 2017). Celen et al. also demonstrated that growth hormone treatment rescued the growth impairment and motor issues (Celen et al., 2017).

5.1.4.3 SETD5

SET Domain-Containing Protein 5 (SETD5) is a ubiquitously expressed gene, with high levels of expression in the brain (Grozeva et al., 2014). It is conserved among mammalian species and is involved in histone modification and transcriptional regulation. Recent studies have identified a number of loss-of-function mutations, *de novo* as well as some familial cases, in *SETD5* in patients with ID*,* implicating *SETD5* variants as a relatively frequent (0.67-0.7%) cause of ID (Grozeva et al., 2014; Kuechler et al., 2015; Fitzgerald et al., 2015; Szczałuba et

al., 2016; Stur et al., 2017). Moreover, *SETD5* is in the critical genomic region for *3p25* microdeletion syndrome, a rare spectrum disorder caused by deletions within the short arm of chromosome 3 (Grozeva et al., 2014; Peltekova et al., 2012). The patients with *de novo SETD5* mutations show phenotypic similarity to those with deletions in the *3p25* region, suggesting that *SETD5* haploinsufficiency may be at the core of the phenotype in the microdeletion syndrome (Grozeva et al., 2014; Kuechler et al., 2015).

SETD5 was proposed to a 'writer' histone lysine methyltransferase due to the presence of SET domain and putative methyl lysine-recognising plant homeodomain (Grozeva et al., 2014; Kuechler et al., 2015). However functional work in mice did not confirm this, and instead, *Setd5* was suggested to lack the PHD domain and thus methyltransferase activity. The *Setd5* mutant mice were shown to be homozygous lethal and no viable null embryos were observed after E10.5 (Osipovic et al., 2016). Of note, there was no difference observed between heterozygous and wildtype embryos. *Setd5* thus has been reported to be critical for embryo development, cell cycle progression, and chromatin modification (Osipovic et al., 2016). Osipovic et al. proposed that Setd5 regulates gene expression through the co-transcriptional regulation of histone acetylation, by interacting with PAF1 and NCoR complexes. Moreover, the study has identified ROSA26 as a bidirectional transcript pair of Setd5. The authors propose that unaltered transcription in ROSA26 direction may aid in maintaining a consistent level of Setd5 expression.

5.1.4.4 SETD1A (KMT2F)

SET Domain Containing 1A (SETD1A), also known as *KMT2F,* encodes a catalytic subunit of the histone lysine methyltransferase protein complex, Set/COMPASS (complex protein associated with Set1). The complex mediates mono-, di-, and trimethylation methylation of histone H3 at Lysine 4 (H3K4) and thus associated with active chromatin structure that is permissive to transcription (Lee and Skalnik, 2008; Schneider et al., 2003).

There is increasing evidence for a role of chromatin modifiers in psychiatric and neurodevelopmental disorders (Dudley et al., 2011; Ronan et al., 2013). Recently two

studies have implicated *SETD1A* in schizophrenia (Singh et al., 2016; Takata et al., 2014). Interestingly, many of the individuals with schizophrenia identified in these studies, heterozygous for *SETD1A* loss-of-function variants, have learning difficulties or developmental delay. It is yet to be elucidated whether *SETD1A* is specifically responsible for the cognitive phenotype in the disorder. The individuals in the DDD cohort were all below the typical age of onset for schizophrenia, so it remains possible that some could develop psychiatric co-morbidities.

4.1.4.5 ZMYND11 (BS69)

Zinc finger MYND domain-containing protein 11 (ZMYND11, also known as BS69), is localised to the nucleus, ubiquitously expressed, and has an inhibitory role in muscle and neuron differentiation (Velasco et al., 2006; Yu et al., 2009). It was first identified as a transcriptional suppressor that interacts with the transactivation domain (conserved region 3) of the 289R adenovirus type 5 E1A protein, involved in cell cycle (Hateboer et al., 1995). More recently, it was demonstrated that ZMYND11 is a chromatin 'reader', specific to the H3.3 variant of histone 3 and that it recognises the tri-methylated histone at Lys-36 (H3K36me3) but does not bind other H3 subtypes (Wen et al., 2014; Wen et al., 2014). H3.3 is enriched in transcribed regions and H3K36me3 is associated with an increase in mRNA expression levels. Therefore ZMYND11 activity co-localizes with highly expressed genes. It has been proposed that ZMYND11 acts as a transcriptional co-repressor with NCoR complex through its MYND domain (Masselink and Bernards, 2000). It is thought to be recruited when sufficient levels of H3.3K36me3 have accumulated, after several initial rounds of transcription, and is responsible for the fine-tuning of gene expression (Wen et al., 2014). Expression of E1A inhibits repression mediated by ZMYND11, therefore modulating its function (Masselink and Bernards, 2000).

Several cases of a 10p15.3 microdeletion, with *ZMYND11* mapped within the region, have been reported with patients with the following clinical features: cognitive, behavioural, and developmental difficulties, speech and motor delay, dysmorphism and brain anomalies, hypotonia, and seizures (DeScipio et al., 2012). Cases with truncating *ZMYND11* mutations

show phenotypes similar to the microdeletion cases suggesting that the haploinsufficiency in *ZMYND11* contributes to these clinical features and that *ZMYND11* mutations are causative of intellectual disability (Coe et al., 2014; Cobben et al., 2014; DeScipio et al., 2012; McRae et al., 2016). *ZMYND11* truncating mutations have also been associated with autistic traits, aggression, and complex neuropsychiatric features (Coe et al., 2014; Vissers et al., 2016). A recent genome-wide association study (GWAS) of behavioural, physiological and gene expression phenotypes in a commercially available Carworth Farms White outbred mouse population identified an association between anxiety-like behaviour and *Zmynd11,* implicating *Zmynd11* in anxiety-like behaviour (Parker et al., 2016).

5.2 Results

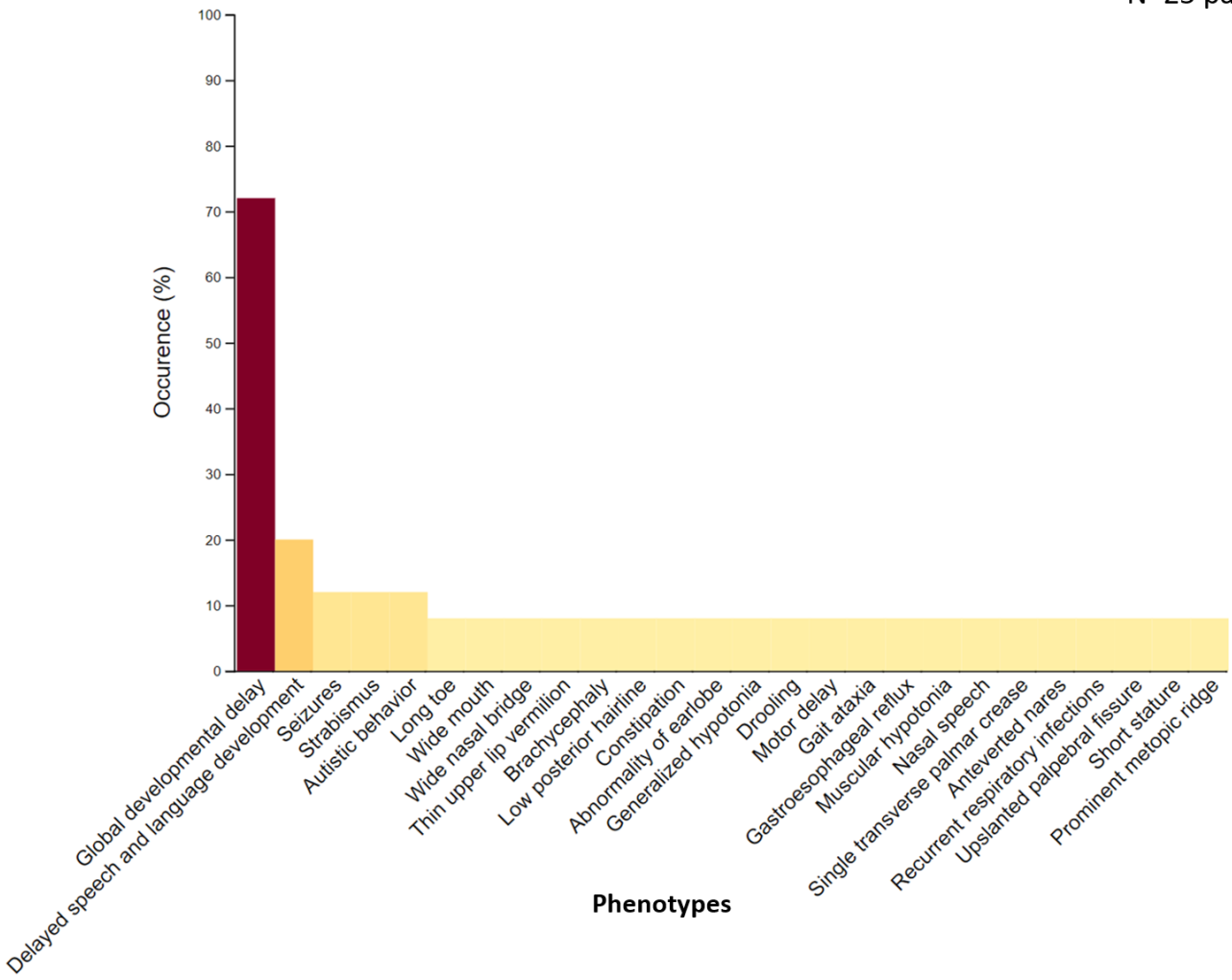
5.2.1 There is high variability in phenotypes within each subset of DDD patients.

From the recent 10,000 patient DDD dataset, 25 probands have mutations in *SETD5* gene, 60 in *ARID1B,* 11 in *SETD1A*, and 6 in *ZMYND11* (DDD study, unpublished work) (Fig.1). *ARID1B* is the most significantly mutated gene in the DDD dataset, with 11 independent loss-of-function mutations identified (Fitzgerald et al., 2015). Either due to the mutational spectrum, the influence of environmental and genetic heterogeneity of the affected individuals, or both, there is a large variability of clinical phenotypes within each set of affected DDD patients, despite all carrying mutations in the same gene. Because of the large number of phenotypes associated with each disorder (254 in the *ARID1B* cohort, 116 phenotypes in *SETD5* cohort, 73 in the *SETD1A* cohort, 40 in *ZMYND11* cohort), as well as many phenotypes described once per cohort, I filtered for phenotypes with occurrence of more than one proband per disorder for the *SETD5, SETD1A,* and *ZMYND11* cohorts (Fig.5.1). For display purposes, filtering of more than two probands was applied for *ARID1B* cohort due to large number of phenotypes (Fig.5.1C). All the affected individuals were haploinsufficient for putative loss-of-function mutations in genes encoding chromatin modifiers. There was only one case of reported consanguinity (in the *ARID1B* cohort). Most parents in the trios were unaffected except one parent in the *ARID1B* cohort and both parents in one of the trios in the *SETD5* cohort.

The most prevalent phenotypes across all the four disorders can be broken down into several broad groups (Fig.5.1). The majority of patients in the four disorders have global developmental delay, a term that encompasses both intellectual disability (ID) (ascribed to patients over the age of 18 who can be diagnosed using an IQ test) and cognitive impairment in children (diagnosed as a delay in achieving developmental milestones). The second most frequent set of phenotypes involves delay in speech and language. Craniofacial deformities, including defects of eyes/ears/face/head, are also common among the patients although these are not always consistent within each disorder. For example, in the *SETD1A* cohort, one patient was macrocephalic, while another was microcephalic. Motor issues, such as motor development delay and ataxia, and/or autistic features have been identified in several patients in each cohort. Aggression was the most common behavioural abnormality described, but often was not found in more than two patients.

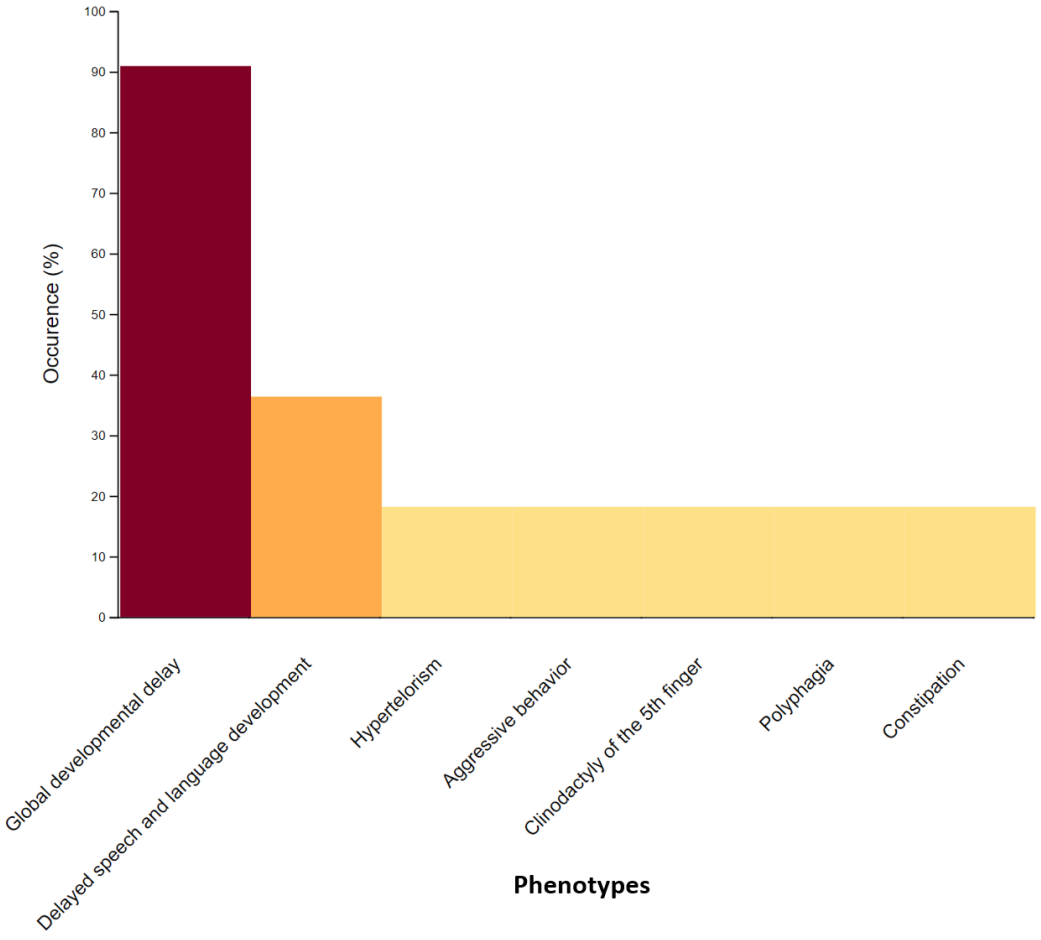
A

SETD5
N=25 patients



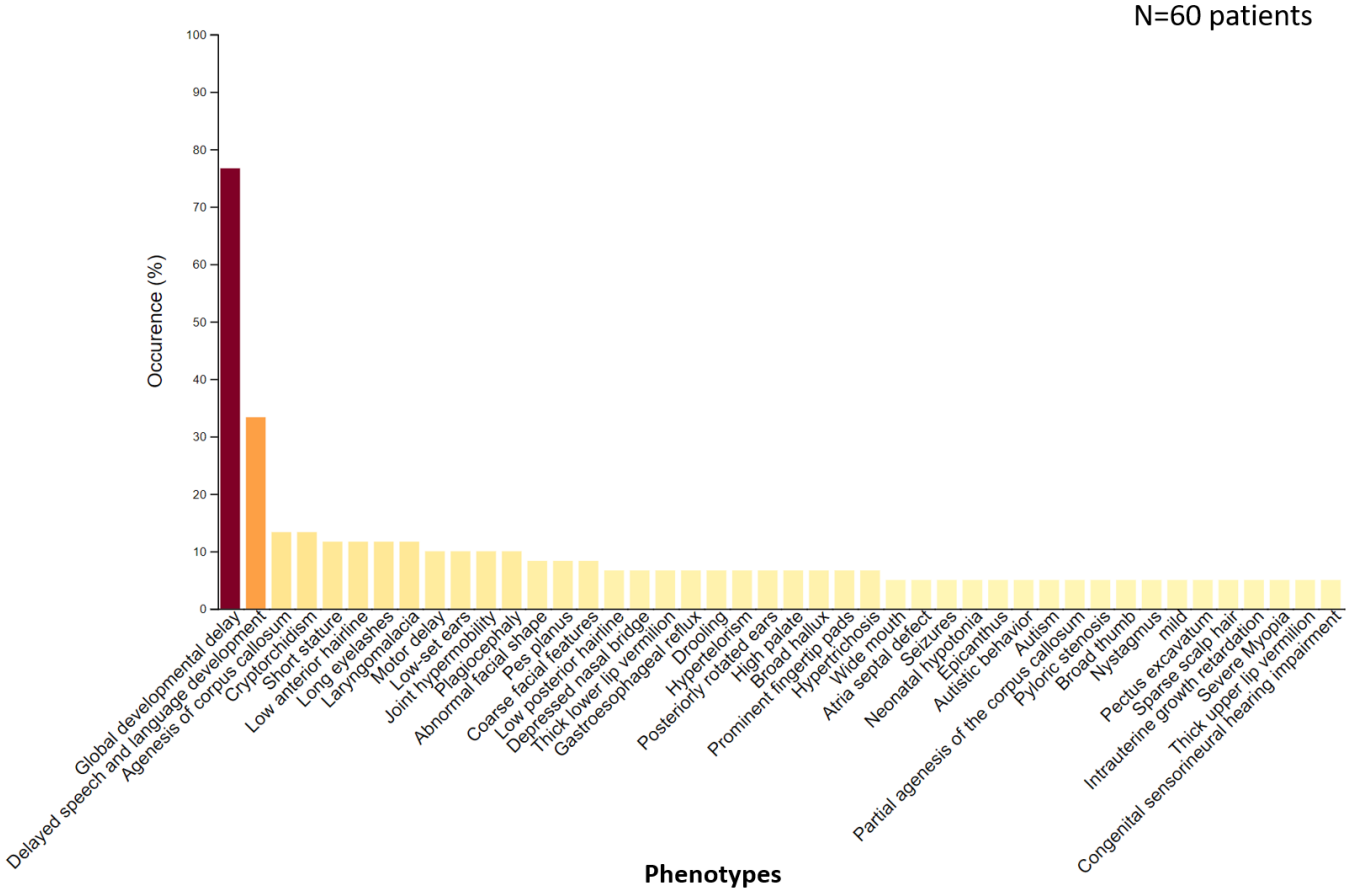
B

SETD1A
N=11 patients



C

ARID1B
N=60 patients



D

ZMYND11
N=6 patients

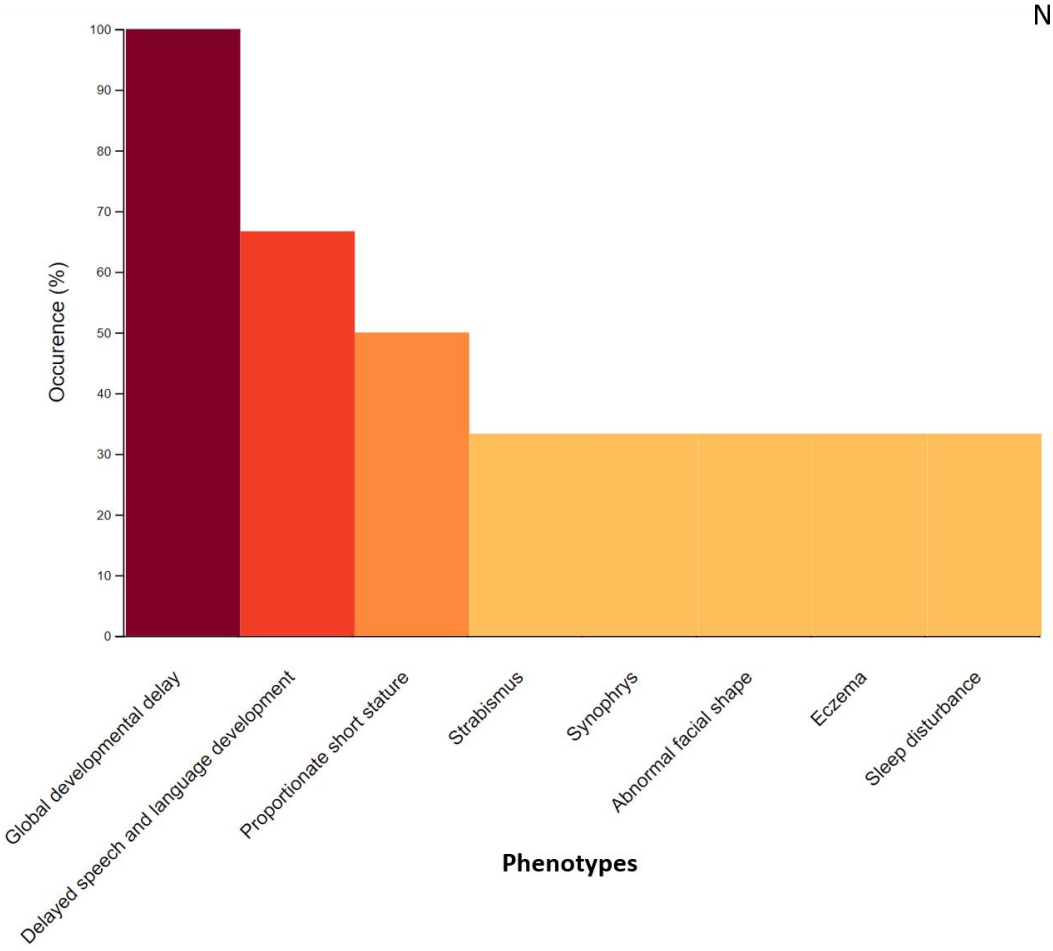


Figure 5.1 Distribution of phenotypes and their occurrences in four novel monogenic disorders identified in the DDD study, associated with mutations in (A) *SETD5* (N=25 probands), (B) *SETD1A* (N=11 probands), (C) *ARID1B* (N=60 probands), *and* (D) *ZMYND11* (N=6 probands). Phenotypes with occurrence (%) in more than one proband, out of the total (displayed in the upper right corner) per disorder have been plotted for all cohorts, except *ARID1B* where filtering of more than two probands was applied due to the large number of phenotypes associated with this disorder. Those phenotypes with the highest occurrence (%) per disorder are displayed first in darker colours, with the colour scale going from darker red orange to lighter orange.

5.2.2 BCND screen design

A series of tests were used, covering multiple behavioural domains. Less invasive assays, such as that testing locomotion and activity were run first to decrease the chance of behavioural outputs being affected by prior test history (McIlwain et al., 2001). This was followed by cognitive tests, which require more handling procedures prior to testing (Võikar et al., 2004). The assays used in BCND, listed in Table 8, were designed to test the most common overlapping phenotypes observed in patients under investigation, with a particular focus on cognitive impairment because the main consistent phenotype shared by patients in all the four disorders is global developmental delay (Fig.5.1).

The most robust assays from *Kptn* work were used in BCND. The efficiency of an assay was judged based on three factors:

1. The length of the protocol. Those that lasted up to a week were used, in order for the overall testing paradigm per mouse line to last 4 weeks.
2. The robustness of the wildtype data, without which conclusive comparisons with mutant mice cannot be made.
3. The ability of the assay to detect subtle behavioural and cognitive deficits.

Because all the affected individuals were heterozygous for mutations in the genes under investigation, I tested heterozygous mice and littermate controls (n=10-15 per line per genotype). Testing began at 10-12 weeks of age, lasted 3-4 weeks, and culminated in the collection of brain samples (hippocampus, prefrontal cortex, striatum, and cerebellum) from n=6 of each genotype, as well as skulls from n=6 per genotype (Fig.5.2). The brain and skull samples were collected for RNA sequencing and micro CT imaging, respectively, for future work as part of the wider study

Table 8 Assays used for testing *Kptn* mice (Chapter 4) and in BCND screen.

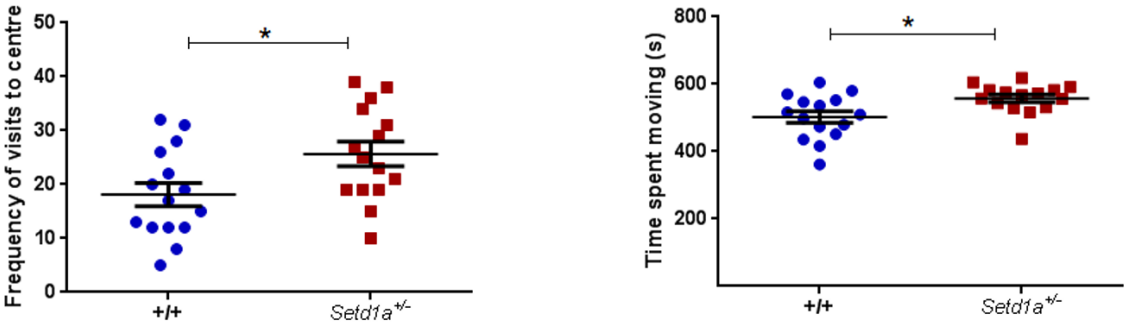
Assays	*Kptn*	BCND	Reasons for inclusion/exclusion
Open Field	Y	Y	
Light/dark box	Y	Y	
Social recognition (SOR)	Y	Y	
Barnes maze	Y	Y	
Sociability	N	Y	Added due to prevalence of autistic features in DDD patient cohort
Object recognition paradigm	Y	N	Too many mice were excluded in the *Kptn* analysis to be viable for a screen (Chapter 4)
Pairwise Discrimination	Y	N	The assay duration is too long (>1 months, including training; Chapter 3)

Figure 5.2 BCND testing schedule flowchart. Sequence of tests in BCND screen are outlined and the duration of each assay, culminating in tissue collection (skulls for micro-CT imaging and brain samples for RNA sequencing).

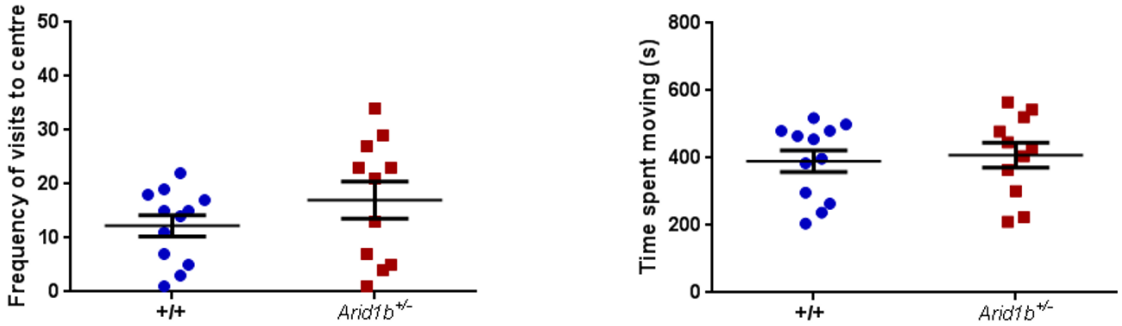
5.2.3 Behavioural abnormalities

The mice were first tested using the open field (OF) assay under bright light (Chapter 2, section 2.3.1). The time spent moving, time in the centre, frequency of visits to the centre, velocity, and total distance covered was measured (Fig.5.3; Fig.5.4). Out of the four lines, only *Setd1a* mutants displayed a behavioural phenotype in this assay (Fig.5.3A). The *Setd1a*$^{+/-}$ mice spent longer moving than their littermate controls (mean difference: 55.41sec, P=0.0121, t=2.684 df=28, two-tailed Student's t-test), suggestive of an increased exploratory behaviour (Fig.5.3A). There was no significant difference between *Setd1a* mutants and controls in the overall distance covered, velocity of movement (Fig.5.4A). *Setd1a*$^{+/-}$ mice did not spend longer in the centre compared to littermate controls, an index of anxiety, but visited the centre more frequently than the controls (mean difference: 7.53, P=0.0221, t=2.422 df=28, two-tailed Student's t-test) (Fig.5.3, Fig.5.4). None of the mutants from the other lines had significant differences relative to their littermate controls in any of the OF parameters (Fig.5.3, Fig.5.4). The data for percentage of time spent in the centre failed the D'Agostino & Pearson normality test for all the lines (*Setd1a*$^{+/-}$: P=0.0437; *Arid1b*$^{+/-}$: P=0.0003; *Setd5*$^{+/+}$: P=0.0021; *Zmynd11*$^{+/+}$: P=0.0094), therefore the two-tailed Mann Whitney test was used to assess genotype differences.

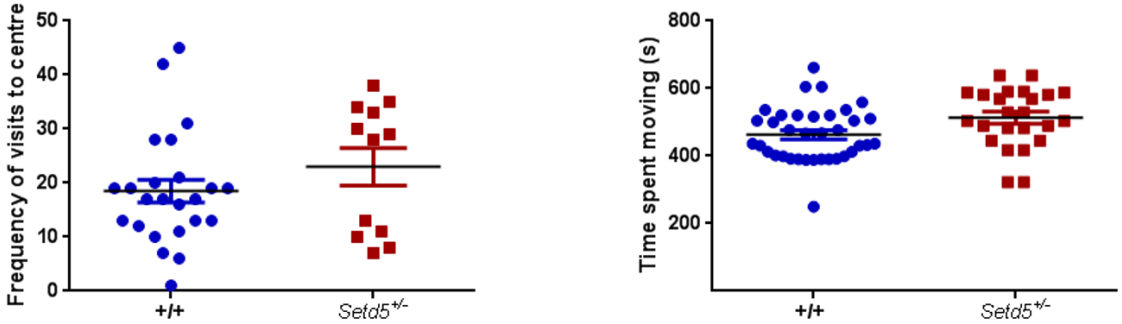
A


Frequency of visits to centre
50
40
30
20
10
0
*
+/+
Setd1a+/-
Time spent moving (s)
800
600
400
200
*
+/+
Setd1a+/-

B


Frequency of visits to centre
50
40
30
20
10
0
+/+
Arid1b+/-
Time spent moving (s)
800
600
400
200
0
+/+
Arid1b+/-

C


Frequency of visits to centre
50
40
30
20
10
0
+/+
Setd5+/-
Time spent moving (s)
800
600
400
200
+/+
Setd5+/-

D

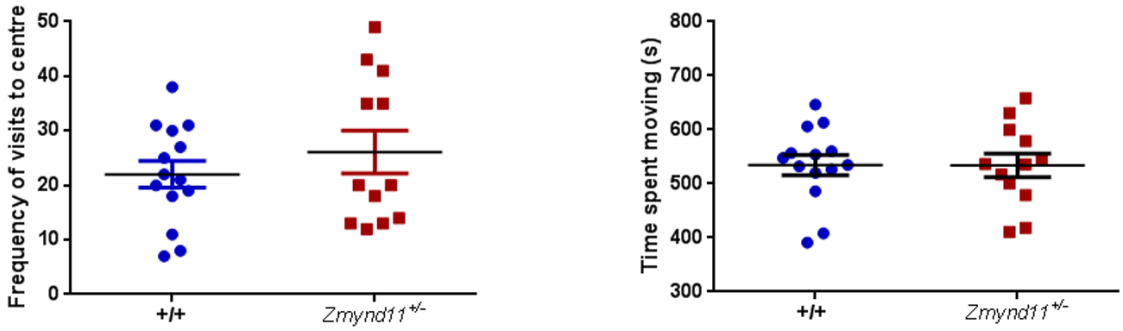
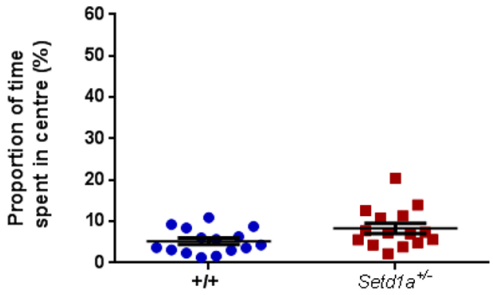
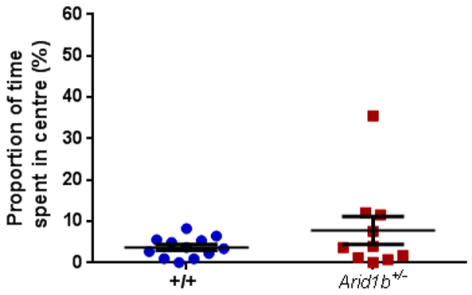

Frequency of visits to centre
50
40
30
20
10
0
+/+
Zmynd11+/-
Time spent moving (s)
800
700
600
500
400
300
+/+
Zmynd11+/-

Figure 5.3 BCND open field results, part 1. The frequency of visits to the centre and time spent moving for four mouse lines is displayed. **A.** *Setd1a* mutants and wildtypes (+/+) littermate controls (n=15 per genotype). *Setd1a* mutant mice made more visits to the centre (P=0.0221*) and spent more time moving overall (P=0.0121*). **B.** *Arid1b* mutants (n=11) and wildtypes (+/+) littermate controls (n=12) had no significant difference in the frequency of visits (P= 0.2344) and time spent moving (P=0.7191). **C**. *Setd5* mutant (n=12) and wildtypes (+/+) littermate controls (n=24) had no significant difference in the frequency of visits (P= 0.2525) and time spent moving (P= 0.1055). **D**. *Zmynd11* mutants (n=12) and wildtypes (+/+) littermate controls (n=14) had no significant difference in the frequency of visits (P= 0.3707) and time spent moving (P= 0.9880). Values are plotted as mean ±SEM.

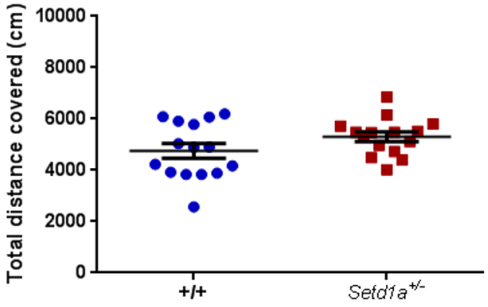
Ai



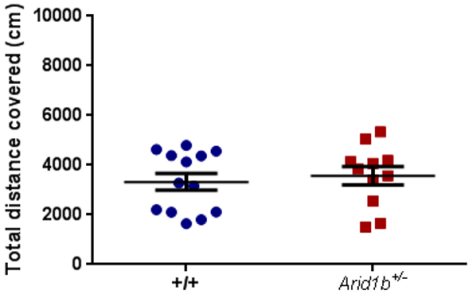
Bi



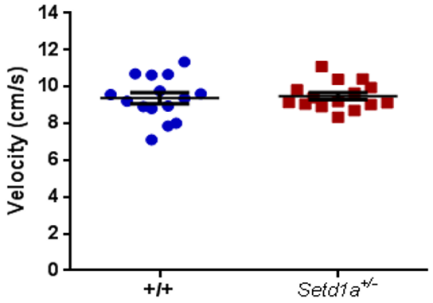
Aii



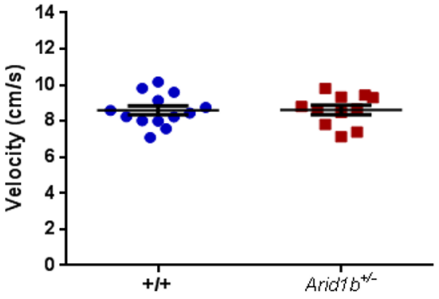
Bii



Aiii



Biii



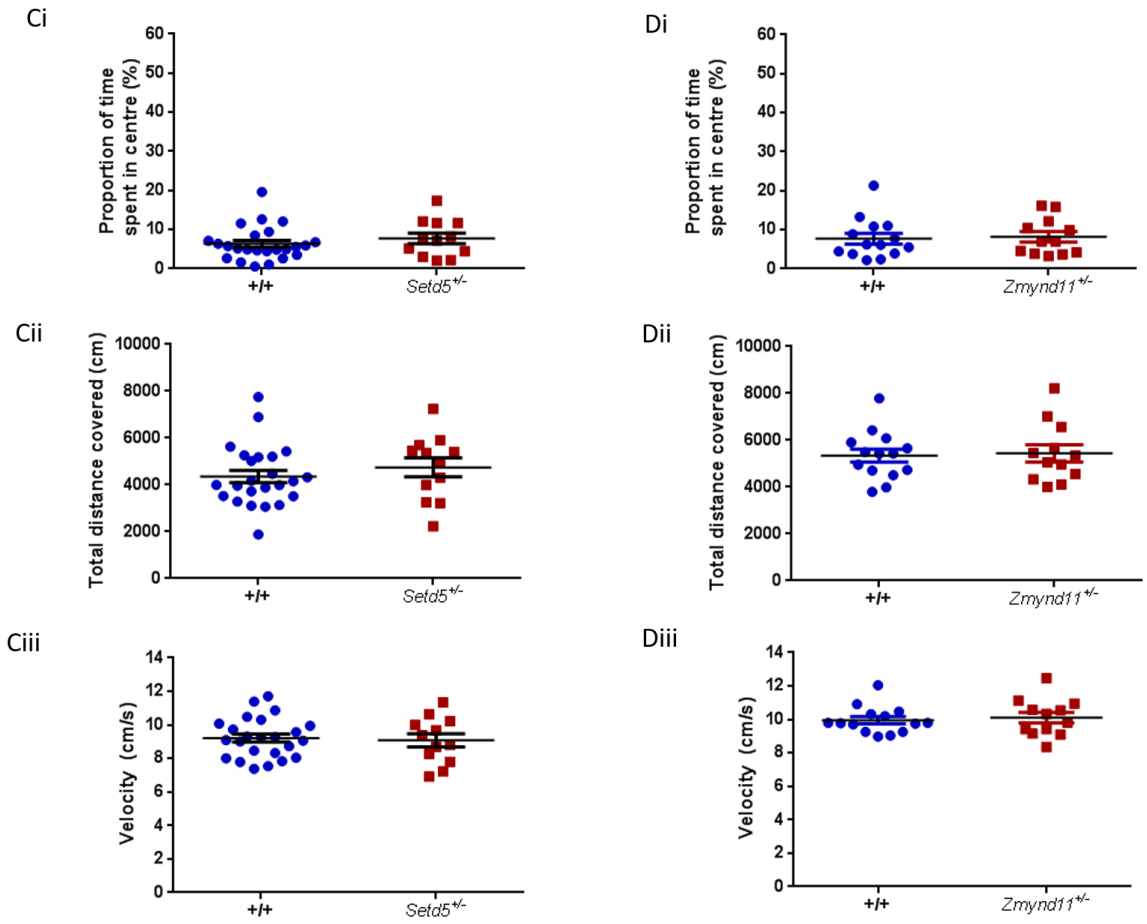


Figure 5.4 BCND open field results, part 2. Non-significant results in (i) time spent in the centre as proportion of total investigation time (%), (ii) total distance covered, and (iii) speed while moving (velocity) between mutants and littermate controls for the four lines is plotted. **A**. *Setd1a* line. Centre: P= 0.0555; distance: P=0.12; velocity: P=0.76. **B.** *Arid1b* line. Centre: P=0.44; velocity: P=0.95. **C.** *Setd5.* Centre: P=0.42; distance: P=0.41; velocity: P= 0.76. **D**. *Zmynd11* line. Centre: P=0.77; distance: P= 0.84; velocity: P=0.55. Values are plotted as mean ± SEM.

To assess a possible anxiety-like phenotype, *Setd1a*$^{+/-}$ mice were tested in the light/dark assay (Chapter 2, section 2.3.2). The transition frequencies to the light zone, as well as time spent in the light zone, were measured in mutants and controls. The light/dark assay was established in our laboratory midway through the testing of the lines; I was therefore only able to run this assay on naïve mice from two lines - *Setd1a and Zmynd11* (Fig.5.5). There was no significant difference, in mutants compared to controls, detected in both lines of transitions from dark zone to light zone (*Setd1a*: P=0.79, t=0.2740 df=23; *Zmynd11:* P=0.22, t=1.253 df=21; Two-tailed Student's t-test) or in time spent in the light zone (*Setd1a*: P=0.59, t=0.5475 df=23; *Zmynd11*: P=0.43, t=0.7963 df=21; two-tailed Student's t-test), implying a lack of anxiety-like behaviour (Fig.5.5). This also suggests that the increased frequency of visits observed in *Setd1a*$^{+/-}$ mice may be a consequence of increased exploration rather than increased anxiety.

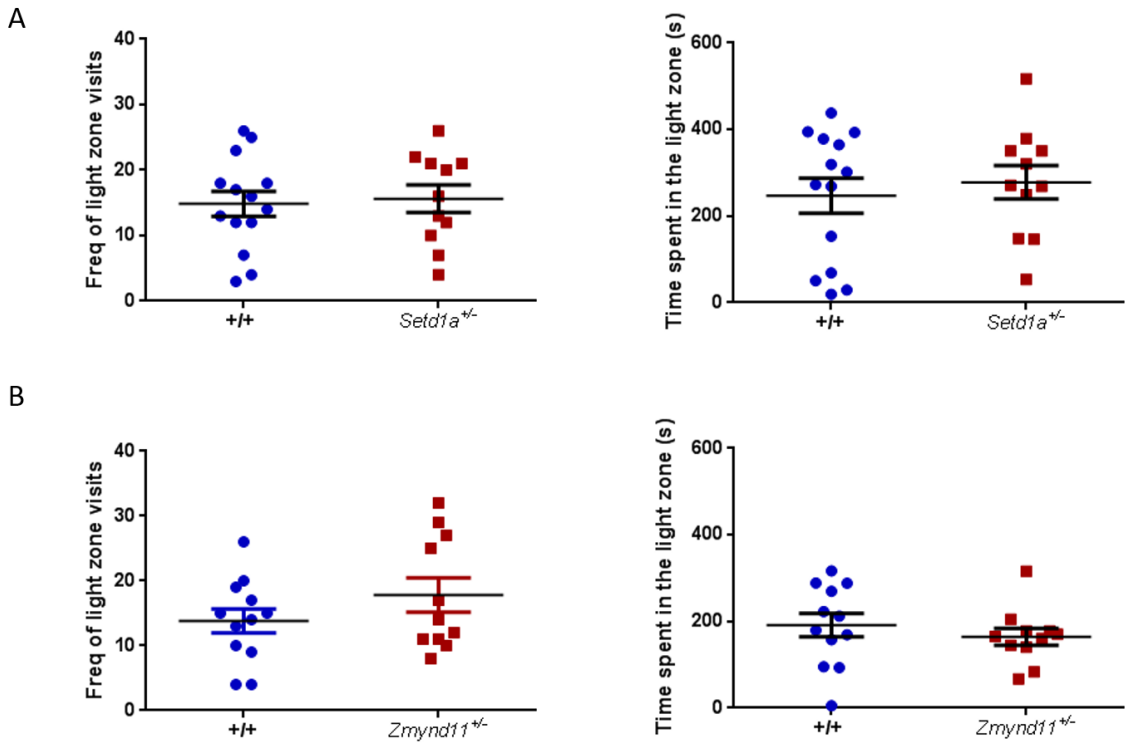
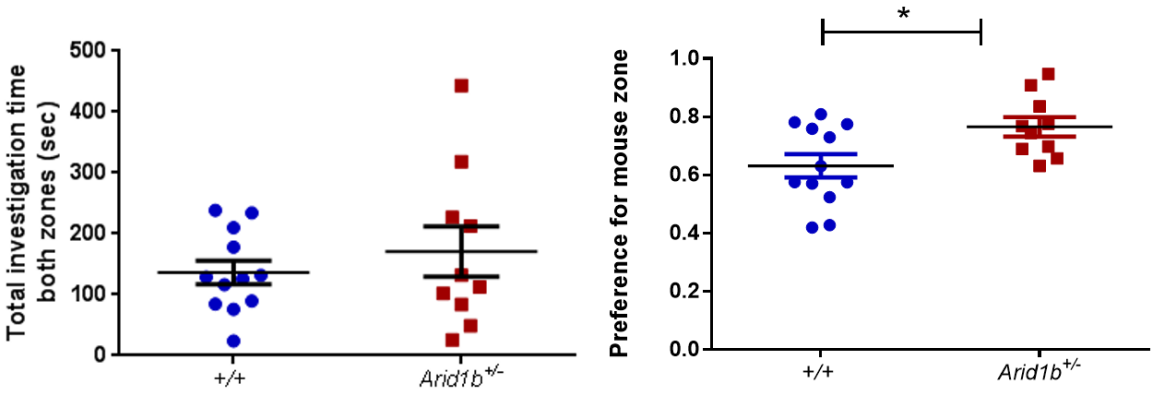


Figure 5.5 BCND light/dark assay results. Frequency of visits to the light zone and time spent in the light zone were plotted for mutants and their littermate controls (+/+). **A.** *Setd1a*: Freq visits: P=0.79; light zone: P=0.59. **B.** *Zmynd11*: Freq visits P=0.22; light zone: P=0.43. Values are plotted as mean ± SEM.

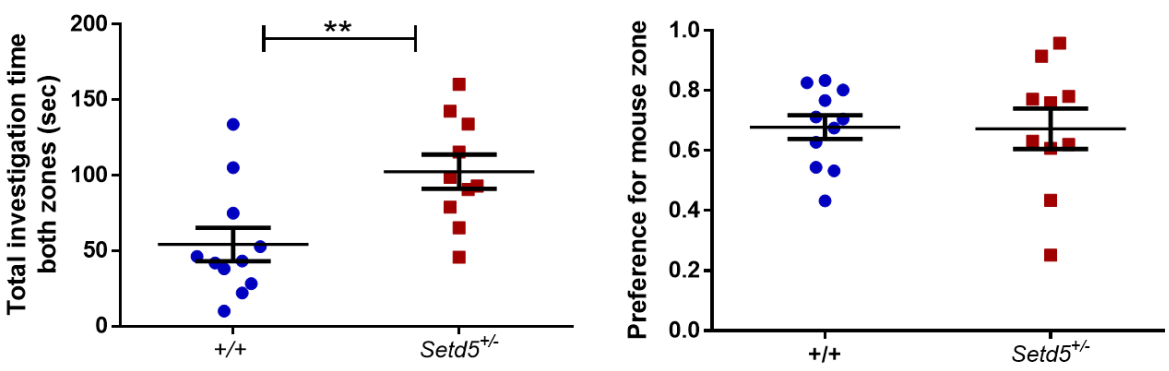
Because of the prevalence of autistic phenotypes in DDD patients overall and *ARID1B* (10%), *SETD5* (12%), *SETD1A* (9%) - I used the three-chamber sociability assay (Chapter 2, section 2.3.5; Fig.5.1). The time spent in the mouse and object zone and the preference for the mouse zone (%) was recorded (Fig.5.6). To ensure the investigation times are comparable between genotypes, overall investigation time was assessed (investigating both mouse zone and object zone). *Setd5*$^{+/-}$ mice spent significantly more time investigating both zones than their littermate controls (mean difference: 48.2sec, P=0.0068, t=3.036 df=19, two-tailed Student's t-test), while in *Arid1b* mice the variance of the two genotypes differed significantly (F test P=0.0392), but the means were not significantly different (mean difference: -0.13 secs, P=0.64, t=0.7909 df=20, two-tailed Student's t-test) (Fig.5.6A-B). There was no significant difference in overall investigation time between mutant and wildtypes in *Setd1a* (P=0.902, t=0.1234 df=26, two-tailed Student's t-test) (Fig.5.6C).

In order to account for the differences in overall investigation, I calculated the time spent investigating the mouse zone as a proportion of total investigation time. In *Arid1b* mice, there was a significant genotype difference in the preference for mouse zone, with an increase of 13% investigation of mouse zone by *Arid1b*$^{+/-}$ compared to the controls (P=0.0198, t=2.532 df=20, two-tailed Student's t-test) (Fig.5.6A). There was no genotype difference in *Setd5* and *Setd1a* mice in the preference for the mouse zone (*Setd5*: P= 0.95, t=0.06875 df=19, two-tailed Student's t-test; *Setd1a*: P=0.51, t=0.6715 df=27, two-tailed Student's t-test) (Fig.5.6B-C).

A


Total investigation time both zones (sec)
500
400
300
200
100
0
+/+
Arid1b+/-
*
Preference for mouse zone
1.0
0.8
0.6
0.4
0.2
0.0
+/+
Arid1b+/-

B


Total investigation time both zones (sec)
200
150
100
50
0
+/+
Setd5+/-
**
Preference for mouse zone
1.0
0.8
0.6
0.4
0.2
0.0
+/+
Setd5+/-

C

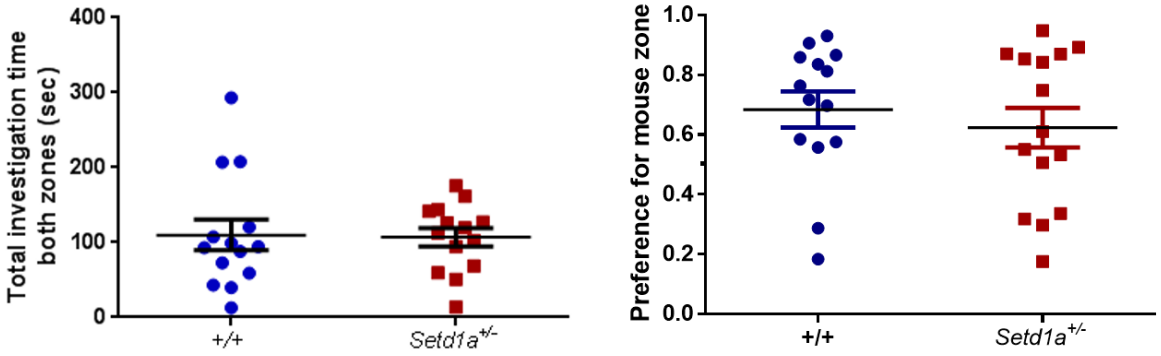

Total investigation time both zones (sec)
400
300
200
100
0
+/+
Setd1a+/-
Preference for mouse zone
1.0
0.8
0.6
0.4
0.2
0.0
+/+
Setd1a+/-

Figure 5.6 BCND sociability results. The overall duration spent near mouse and object zones for the mutants (red) and wildtype controls (blue) are displayed on the left in seconds and the preference for the mouse zone is displayed on the right as a preference score from 0-1, with 0.5 indicating no preference. **A.** *Arid1b* left graph: there was no genotype difference in the means of *Arid1b* overall investigation time but there was a significant difference in variance (F test: P=0.039). *Arid1b* right graph: there was a significant difference between genotypes (P=0.0198*), with an increase of 13% investigation of mouse zone in *Arid1b*$^{+/-}$ compared to littermate controls. **B.** *Setd5* left graph: the *Setd5* mutants had a significantly greater total investigation time of both holders (P=0.0068**). *Setd5* right graph: There was no significant genotype difference in preference for the mouse zone (P= 0.95). **C.** *Setd1a* left graph: There was no significant difference in overall investigation times between genotypes (P= 0.9027). *Setd1a* right graph: There was no significant genotype difference in the preference for the mouse zone (P=0.51).

5.2.4 Cognitive impairment

Next, I assessed cognitive impairment in the mice, by using two cognitive assays – social recognition (Chapter 2, section 2.3.4) and the Barnes maze (Chapter 2, section 2.3.7), both testing overlapping cognitive domains associated with hippocampus-dependent learning.

5.2.4.1 Olfactory-mediated memory

Social recognition (SOR) assay (Chapter 2, section 2.3.4), is reliant on the inherent preference of mice to investigate an unfamiliar conspecific longer than a familiar mouse. To ensure mice do not have a bias for either of the stimuli before acquisition of memory, I tested naïve wildtype mice (n=12) of the same background strain for the presence of inherent preferences and did not observe any stimulus bias. The mice did not spend a significantly different time investigating either of the stimuli (mean difference: 5.1sec, P= 0.57, t=0.5924 df=11, paired two-tailed Student's t-test) (Fig.5.7).

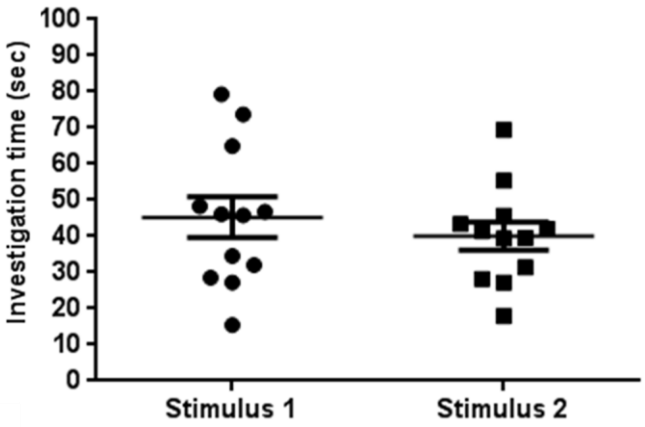


Figure 5.7 No inherent stimulus bias was detected. Mice naïve to the assay did not show a preference for either of the two stimulus animals presented (P=0.57).

I therefore proceeded with SOR assessment. On Day 1, there was an overall difference in investigation time across trials and no genotype difference for *Setd1a* line (trials: F (4, 68) = 12.13, P<0.0001; genotype: F (1, 17) = 1.165, P=0.2955; two-way ANOVA), *Arid1b* line: trials: F (4, 44) = 6.494, P=0.0003; genotype: F (1, 11) = 2.784, P=0.1234; two-way ANOVA), *Zmynd11* line: F (4, 106) = 11.76, P<0.0001; genotype: F (1, 106) = 1.269, P= 0.2624; two-way ANOVA) (Fig5.8). *Setd5* mice displayed a trial difference on Day 1, but also a genotype difference, with an increase in overall investigation time (trials: F (4, 112) = 14.81, P<0.0001; genotype: F (1, 112) = 10.28, P=0.0018; two-way ANOVA) (Fig.5.8B). On trial 5, for *Setd1a, Setd5, Zmynd11* lines, both genotypes showed increased investigation of the novel stimulus when comparing the investigation during trial 4 and trial 5 (*Setd1a*: wildtypes: P<0.001, mutants: P<0.05; *Setd5*: wildtypes: P<0.001, mutant: P<0.01; Zmynd11: wildtypes: P<0.001, mutants: P<0.05; post-hoc analysis after two-way ANOVA). *Arid1b* controls (n=6) did not display a significant difference in the *post hoc* analysis between trial 4 and 5, but did show a reduction in investigation between trial 3 and 5 (P<0.05) and trial 2 and 5 (P<0.05), while *Arid1b*$^{+/-}$ (n=7) did not display an overall difference across trials (two-way ANOVA), but there was a significant increase in trial 5 vs trial 4 when these were compared separately (wildtype: P= 0.0208, t=3.330 df=5; Mutant:P= 0.0157, t=3.334 df=6; two-tailed Student's t-test). I therefore progressed the *Arid1b* mice to Day 2 of testing.

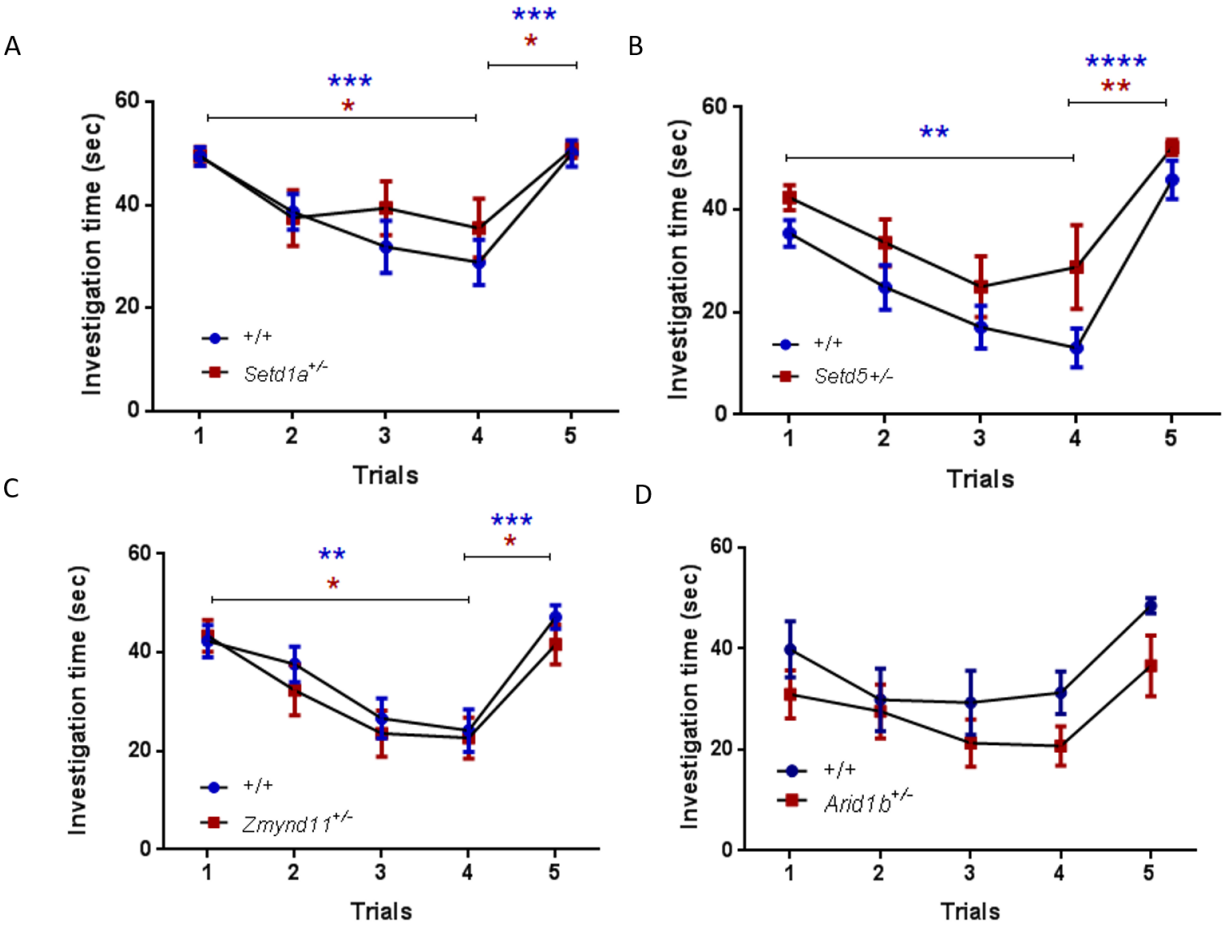


Figure 5.8 BCND results for Day 1 of SOR. Investigation time is plotted for wildtypes (+/+; blue) and mutants (red) for each of the four mouse lines. **A**. *Setd1a* line. There was a significant difference between trial 1 and 4 (wildtype: P<0.001***, Mutant: P<0.05*) and trial 4 and 5 (wildtype: P<0.001***, Mutant: P<0.05*). **B**. *Setd5* line. There was a significant difference between trial 1 and 4 only in the wildtypes (P<0.001***), but trial 4 and 5 in both genotypes (wildtype: P<0.001***, Mutant: P<0.001**). C. There was a significant difference between trial 1 and 4 (wildtype: P<0.01**, Mutant: P<0.05*) and trial 4 and 5 (wildtype: P<0.001***, Mutant: P<0.01**). **C.** *Zmynd11* line. There was a significant difference between trial 1 and 4 (wildtype: P<0.01**, Mutant: P<0.05*) and trial 4 and 5 (wildtype: P<0.001***, Mutant: P<0.05*). **D.** *Arid1b* line. There was a significant difference between trial 4 and 5 (with Student's t-test) (wildtype: P= 0.0208; Mutant: P= 0.0157). Values are plotted as mean ± SEM.

On Day 2, there was no difference in overall investigation time of the two stimuli between genotypes in any of the lines (*Setd1a*: P=0.09, t=1.785 df=19, Two-tailed Student's t-test; *Setd5*: P=0.6405, Two-tailed Mann Whitney test; *Zmynd11*: P=0.88, t=0.1554 df=20, Two-tailed Student's t-test; *Arid1b*: P=0.58 , t=0.5664 df=11, Two-tailed Student's t-test). (Fig.5.9A-D). The *Setd5* data failed the normality test (wildtype P<0.0001, mutants P=0.0003) and therefore the difference in investigation time was analysed with a Mann-Whitney test.

The wildtypes in all the lines were able to distinguish between unfamiliar and familiar stimuli and therefore spent significantly longer with the unfamiliar mouse (*Setd1a*: P= 0.0078, t= 4.37891 df= 18.0; *Setd5*: P= 0.00022, t= 4.28034 df= 26.0; Mutant P= 0.033; *Zmynd11*: P= 0.014, t= 2.6401 df= 26.0; *Arid1b*: P= 0.016, t=2.87621 df=10.0, two-tailed multiple t-test with Holm-Sidak method corrections) (Fig.5.10A-D). The performance in the memory test varied in the mutants of the four lines. *Setd1a*$^{+/-}$ and *Setd5*$^{+/-}$ mice were able to distinguish between familiar and unfamiliar, and therefore spent significantly longer with the unfamiliar (*Setd1a*$^{+/-}$: mean difference: 15.1sec, P= 0.00036, t= 2.95575 df= 20.0; *Setd5*$^{+/-}$ *:*mean difference: 15.8sec, P= 0.033; , t= 2.29697 df= 20.0; two-tailed multiple t-test with Holm-Sidak method corrections), while *Arid1b*$^{+/-}$ and *Zmynd11*$^{+/-}$ did not spend significantly longer with either of the stimuli (*Zmynd11*$^{+/-}$: P= 0.86, t= 0.182443 df= 16; *Arid1b*$^{+/-}$ P= 0.399698, t= 0.873185 df= 12; two-tailed multiple t-test with multiple comparison corrections), which is indicative of 24h memory impairment in these mutants (Fig.5.10).

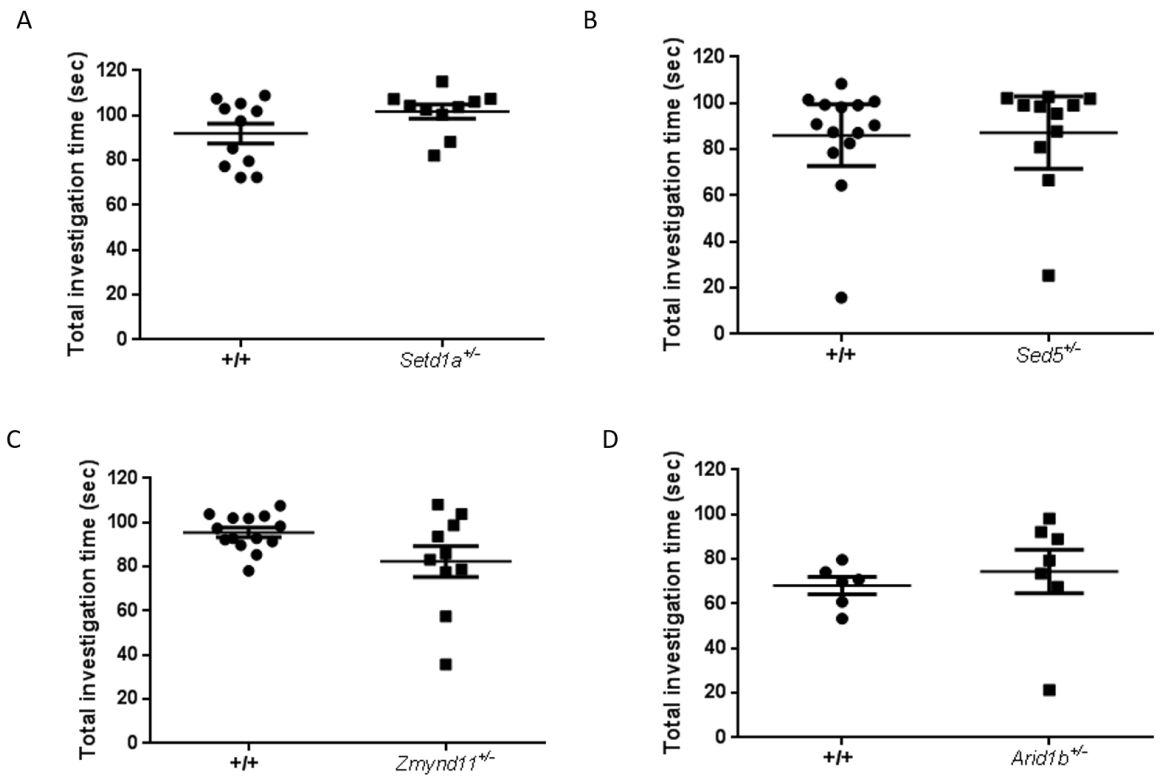


Figure 5.9 Total investigation time on Day 2 of social recognition. Total investigation time of familiar and unfamiliar stimuli by wildtypes (+/+) and mutants in four mouse lines is shown. There was no genotype difference between total investigation times in all four lines. **A.** *Setd1a* line (P=0.09), **B.** *Setd5* line (P=0.6405). **C.** *Zmynd11* line (P=0.88). D. *Arid1b* line (P=0.58). Values are plotted as mean ± SEM.

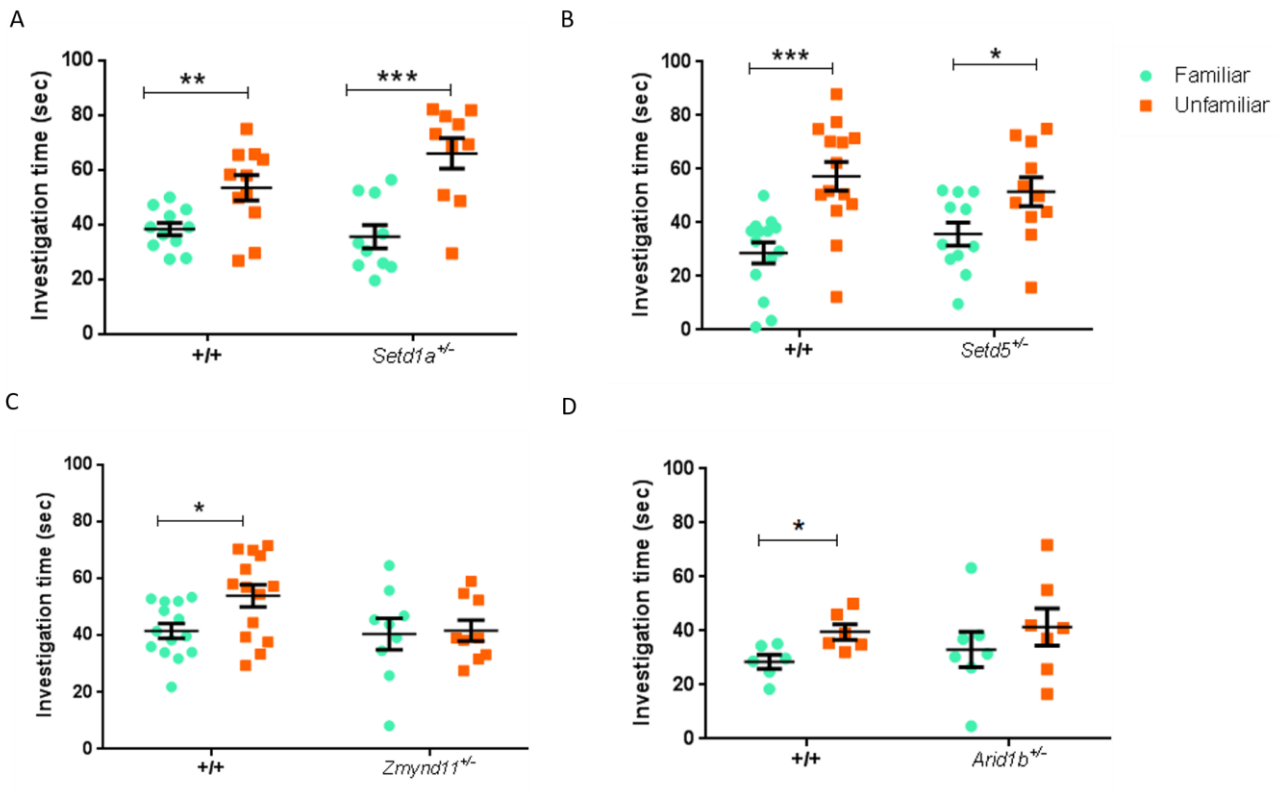


Figure 5.10 Day 2 SOR 24h memory test results. The investigation time of familiar (turquoise) and the unfamiliar (orange) stimuli by wildtype (+/+) and mutant mice from four mouse lines is shown. **A.** *Setd1a* line. Both genotypes spent longer investigating the unfamiliar mouse (wildtype: P= 0.0078**; Mutant P= 0.00036***). **B.** *Setd5* line. Both genotypes spent longer investigating the unfamiliar mouse (wildtype P= 0.00022***; Mutant P= 0.033*). **C**. *Zmynd11* line. The wildtypes spent longer with the unfamiliar stimuli than the familiar (P= 0.014*), whereas the mutants did not show a significant difference in time spent next to either stimulus. **D**. *Arid1b* line. The wiltdypes spent longer with the unfamiliar stimuli than the familiar (P= 0.016*), whereas the mutants did not show a significant difference in time spent next to either stimulus. Values are plotted as mean ± SEM.

5.2.4.2 Spatial memory

I then assessed spatial memory of the mice using the Barnes maze (Chapter 2, section 2.3.7; and as described in Chapter 4). Learning during training was assessed by measuring primary latency to reach the target and a total number of errors recorded before the mice went inside the escape box. Training was plotted either as an average performance per training day (Fig.5.11(i-ii)) or as a breakdown of all trials for each training day (Fig.5.11(iii)). Memory was tested during 24h and 72h probe trials, as the duration of time the mice spent at the target.

In *Zmynd11* and *Setd1a* lines, there was no genotype difference in the primary latency (Training 1: *Zmynd11*: F (1, 44) = 1.388, P=0.2450; *Setd1a*: F (1, 58) = 0.009076, P=0.9244; two-way ANOVA. Training 2: *Setd1a*: F (1, 89) = 1.123, P=0.2921, two-way ANOVA) and total errors made (Training 1: *Zmynd11*: F (1, 44) = 1.264, P=0.2670; *Setd1a*: F (1, 58) = 0.4114, P=0.5238; two-way ANOVA. Training 2: *Setd1a*: F (1, 89) = 0.3736, P=0.5426, two-way ANOVA) (Fig.5.11A,C).

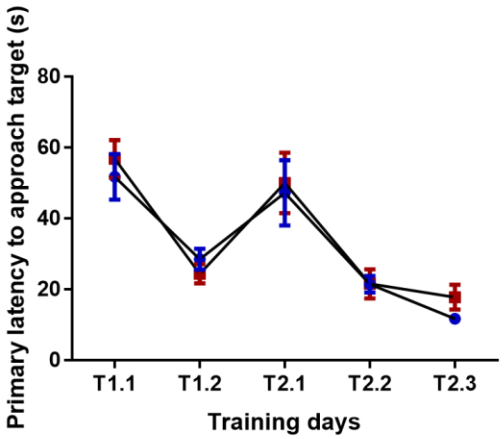
There was no significant difference in total errors made by *Arid1b* or *Setd5* mutants when compared to the littermate wildtype controls during training (*Arid1b*: training N1: F (1, 62) = 2.11, P=0.1514; training N2: F (1, 93) = 2.55, P=0.1137, two-way ANOVA. *Setd5* training N1: F (1, 42) = 0.6573, P=0.4221, two-way ANOVA) (Fig.5.11B,D). However, *Arid1b*$^{+/-}$ and *Setd5*$^{+/-}$ mice were slower during Training N1, as measured by the primary latency to approach the target zone. There was an overall genotype difference in primary latency in *Arid1b* line (F (1, 26) = 14.68, P=0.0007, two-way ANOVA) (Fig.5.11Di). Even though *Arid1b*$^{+/-}$ mice were not different to controls during the first trial (T1.1_1), by the last trial (T1.1_4) of T1.1 (Day 1 of TrainingN1) the mutants took significantly longer to approach the target hole than the controls as measured by primary latency (P<0.001, *post-hoc* analysis after two-way ANOVA) (Fig.5.11Diii). However, on the next day of training (T1.2) the mutants were able to perform as well as the controls throughout T1.2, as well as during training when the location of the target was reversed (T2.1-T2.3). There was no genotype difference in primary latency (F (1, 26) = 0.8018, P=0.3788, two-way ANOVA) and total errors (F (1, 93) = 2.55, P=0.1137, two-

way ANOVA) in the *Arid1b* cohort during Training N2 (Fig.5.11D). For *Setd5* line, there was no overall genotype difference in the average primary latency training (F (1, 42) = 3.424, P=0.0713, two-way ANOVA; Fig.5.11Bi). There was an overall genotype difference in primary latency when the breakdown of trials was analysed (F (1, 168) = 4.751, P=0.0307, two-way ANOVA), but no significant genotype difference in primary latency between individual trials (P>0.05, *post-hoc* analysis of trial comparisons after two-way ANOVA) (Fig.5.11Biii).

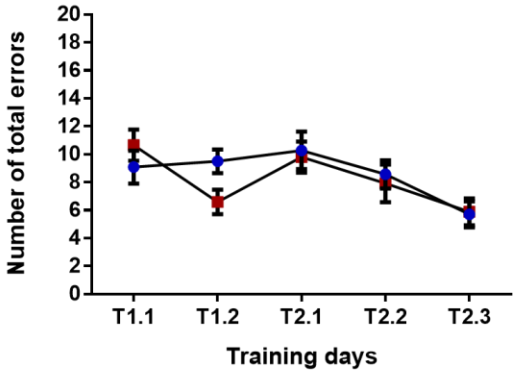
There was a significant overall difference in the investigation (%) of the 20 holes on 24h probe trial (*Setd1a* line: F (19, 580) = 22.13, P<0.0001; *Setd5*: F (19, 420) = 2.127, P=0.004; *Arid1b*: F (19, 570) = 12.94, P<0.0001; *Zmynd11*: F (19, 418) = 10.72, P<0.0001; two-way ANOVA), but no genotype difference in investigation of the target (hole N1) (P >0.9999 for all four mouse lines; post-*hoc* analysis after two-way ANOVA) (Fig.5.12A-D).

During the 72h probe, *Setd1a* mutant mice spent the same amount of time near the target as the controls, however *Setd1a*$^{+/-}$ also spent significantly longer around the holes on either side of the target when compared to the controls (hole 10: P=0.0171, hole 12: 0.0399, target: P= 0.9998, *post-hoc* analysis after two-way ANOVA), with a greater standard error around the mean time spent near those holes (Fig.5.13A). This is indicative of a reduced specificity of memory, with mutants remembering the overall area where the zone is but show uncertainty as to the exact whereabouts by investigating the neighbouring holes significantly longer than wildtype mice. For the other two lines, *Arid1b* and *Zmynd11*, there was a significant overall difference between time (%) spent at the 20 holes during the 72h probe trial (*Arid1b*: F (19, 570) = 12.94, P<0.0001; *Zmynd11:* F (19, 420) = 5.012, P<0.0001), but no genotype difference in investigation of the target (*post-hoc* analysis after two way ANOVA). (Fig.5.13B-C)

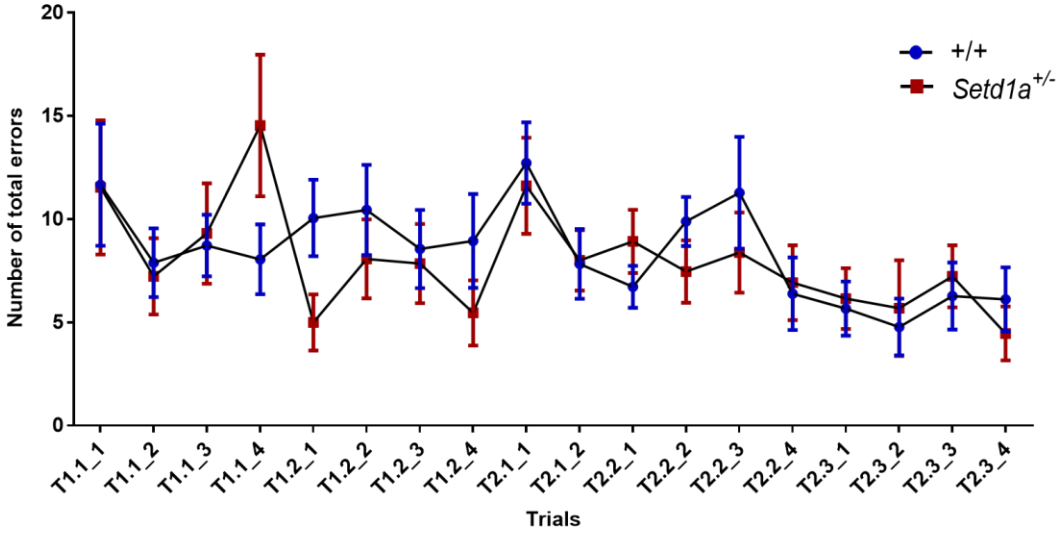
Ai


Primary latency to approach target (s)
80
60
40
20
0
T1.1
T1.2
T2.1
T2.2
T2.3
Training days

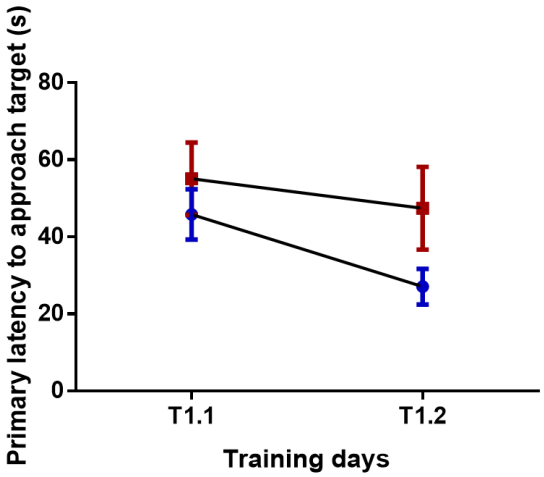
Aii


Number of total errors
20
18
16
14
12
10
8
6
4
2
0
T1.1
T1.2
T2.1
T2.2
T2.3
Training days

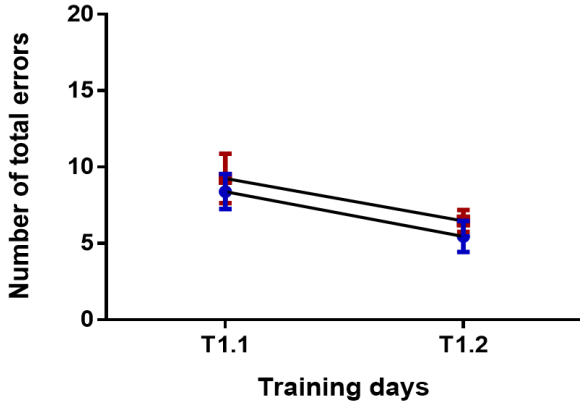
Aiii


Number of total errors
20
15
10
5
0
+/+
Setd1a+/-
T1.1_1
T1.1_2
T1.1_3
T1.1_4
T1.2_1
T1.2_2
T1.2_3
T1.2_4
T2.1_1
T2.1_2
T2.2_1
T2.2_2
T2.2_3
T2.2_4
T2.3_1
T2.3_2
T2.3_3
T2.3_4
Trials

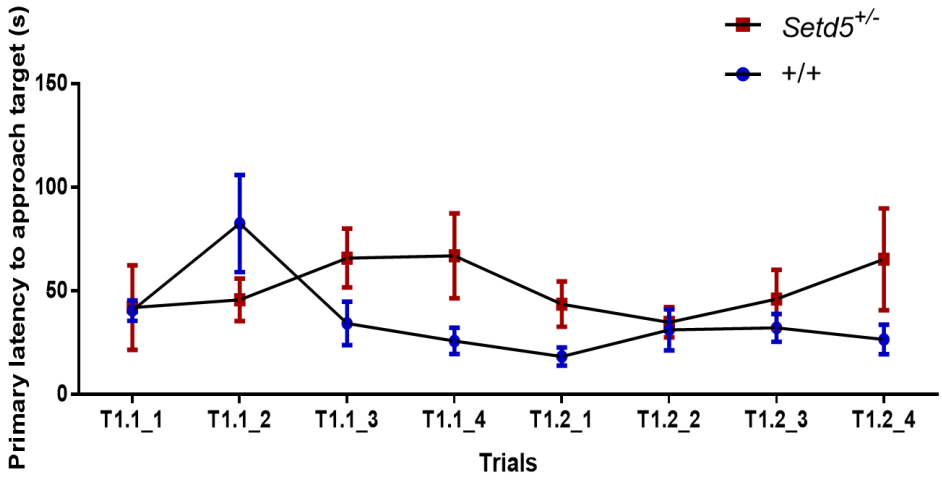
Bi


Primary latency to approach target (s)
80
60
40
20
0
T1.1
T1.2
Training days

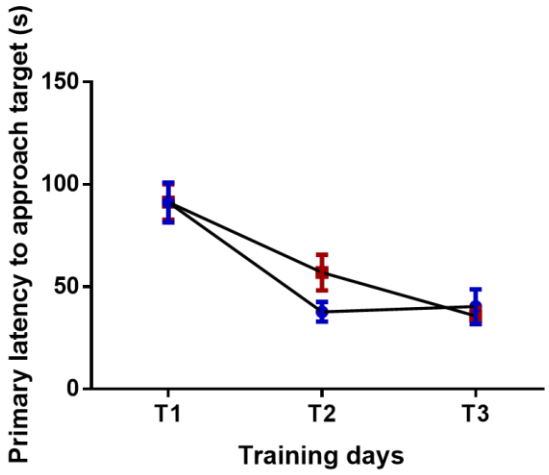
Bii


Number of total errors
20
15
10
5
0
T1.1
T1.2
Training days

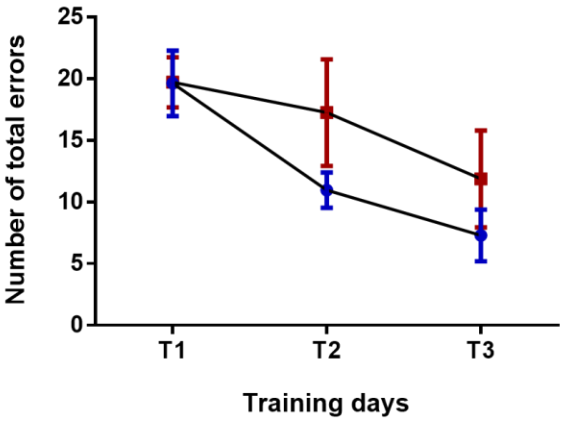
Biii


Setd5+/-
+/+
Primary latency to approach target (s)
150
100
50
0
T1.1_1
T1.1_2
T1.1_3
T1.1_4
T1.2_1
T1.2_2
T1.2_3
T1.2_4
Trials

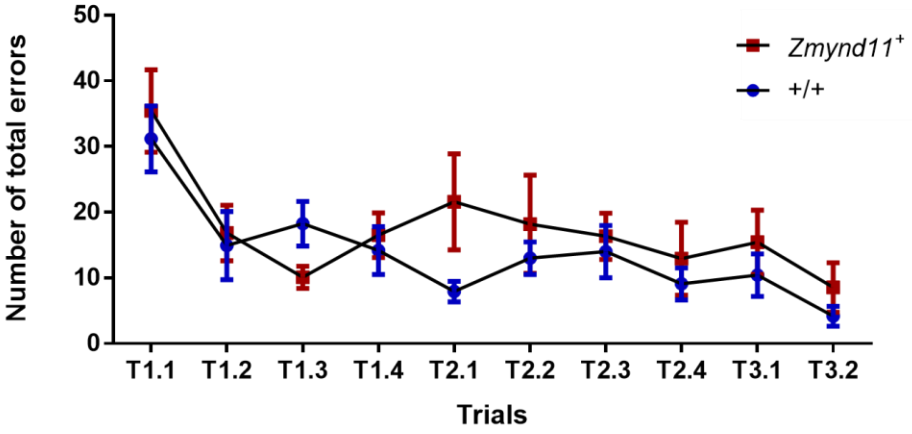
Ci


Primary latency to approach target (s)
150
100
50
0
T1
T2
T3
Training days

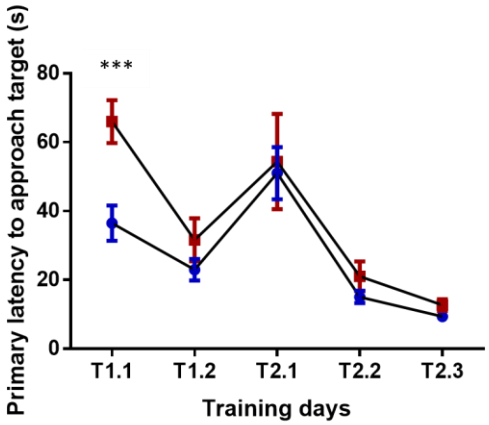
Cii


Number of total errors
25
20
15
10
5
0
T1
T2
T3
Training days

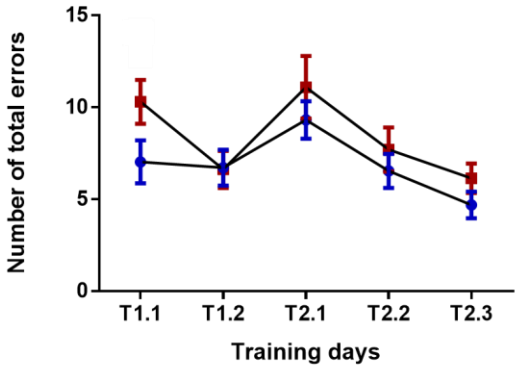
Ciii


Number of total errors
50
40
30
20
10
0
T1.1
T1.2
T1.3
T1.4
T2.1
T2.2
T2.3
T2.4
T3.1
T3.2
Trials
Zmynd11+
+/+

Di


***
Primary latency to approach target (s)
80
60
40
20
0
T1.1
T1.2
T2.1
T2.2
T2.3
Training days

Dii


Number of total errors
15
10
5
0
T1.1
T1.2
T2.1
T2.2
T2.3
Training days

Diii

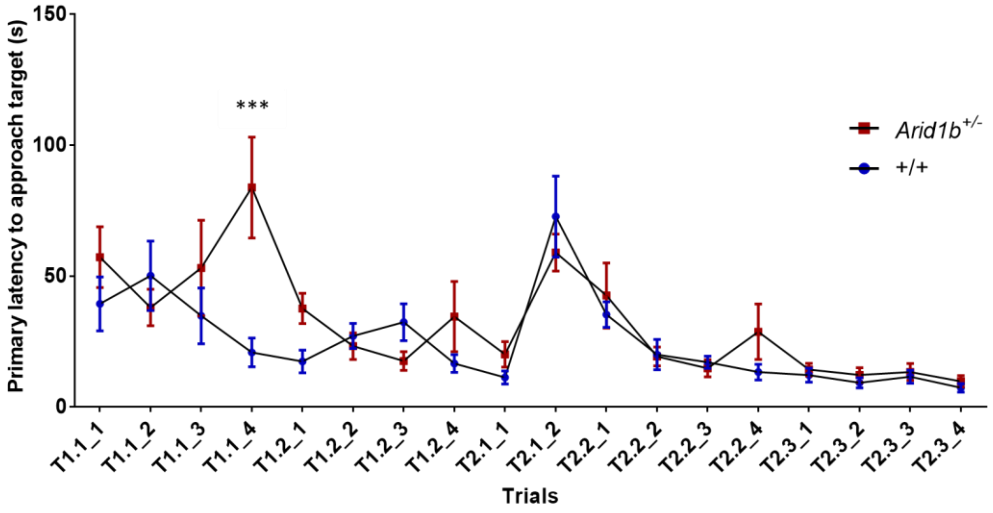

***
Primary latency to approach target (s)
150
100
50
0
Arid1b+/-
+/+
T1.1_1
T1.1_2
T1.1_3
T1.1_4
T1.2_1
T1.2_2
T1.2_3
T1.2_4
T2.1_1
T2.1_2
T2.2_1
T2.2_2
T2.2_3
T2.2_4
T2.3_1
T2.3_2
T2.3_3
T2.3_4
Trials

Figure 5.11 BCND Barnes maze training. A. *Setd1a* line. Ai) Average primary latency per training day during training N1 (T1.1-1.2) and training N2 (T2.1-2.3), Aii) Average errors per training day during training N1 (T1.1-1.2) and training N2 (T2.1-2.3), Aiii) Total errors made per trial over all training days. There was no significant genotype difference in primary latency of approach to target and total errors made. **B.** *Setd5* line. Bi) Average primary latency per training day during raining N1 (T1.1-1.2), Bii) Average errors per training day during raining N1 (T1.1-1.2), Biii) Primary latency to approach target zone per trial over all training days. There was an overall genotype difference in primary latency per trial (P=0.0307). **C.** *Zmynd11* line. Ci) Average primary latency per training day during training N1 (T1.1-1.2) and training N2 (T2.1-2.3), Cii) Average errors per training day during training N1 (T1.1-1.2) and training N2 (T2.1-2.3), Ciii) Total errors made per trial over all the training days. There was no significant genotype difference in primary latency of approach to target and total errors made. **D.** *Arid1b* line. Di) Average primary latency per training day during training N1 (T1.1-1.2) and training N2 (T2.1-2.3). There was an overall difference in the primary latency between genotypes and training days (genotype: P=0.0007, two way ANOVA), with a significant genotype difference on day 1 (T1) (P<0.001***). Dii) There was no genotype difference in the average errors per training day during training N1 (T1.1-1.2) and training N2 (T2.1-2.3), Diii) There was a significant genotype difference in primary latency in trial 4 of day one (T1.1_4) (P<0.001***). Values are plotted as mean ± SEM.

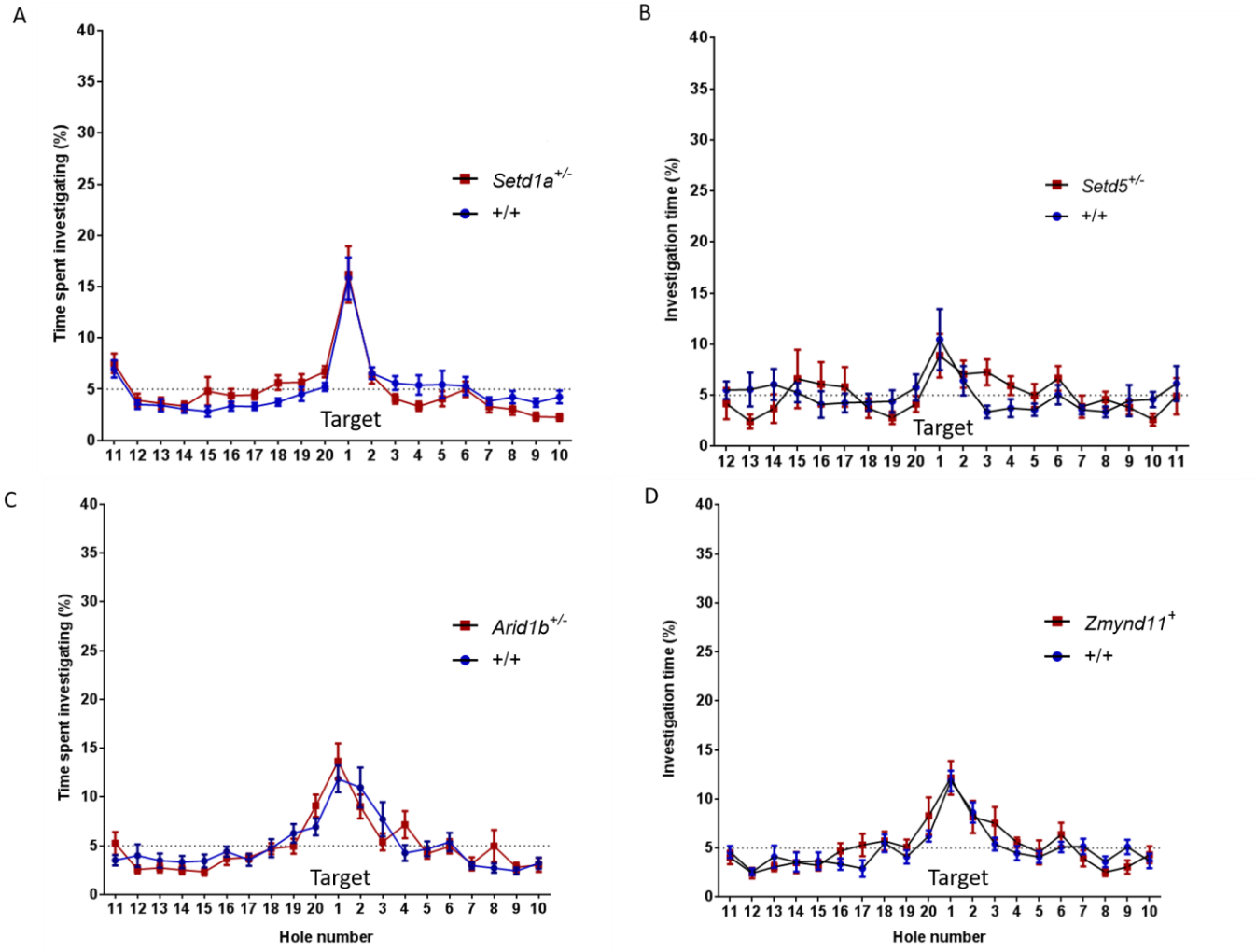
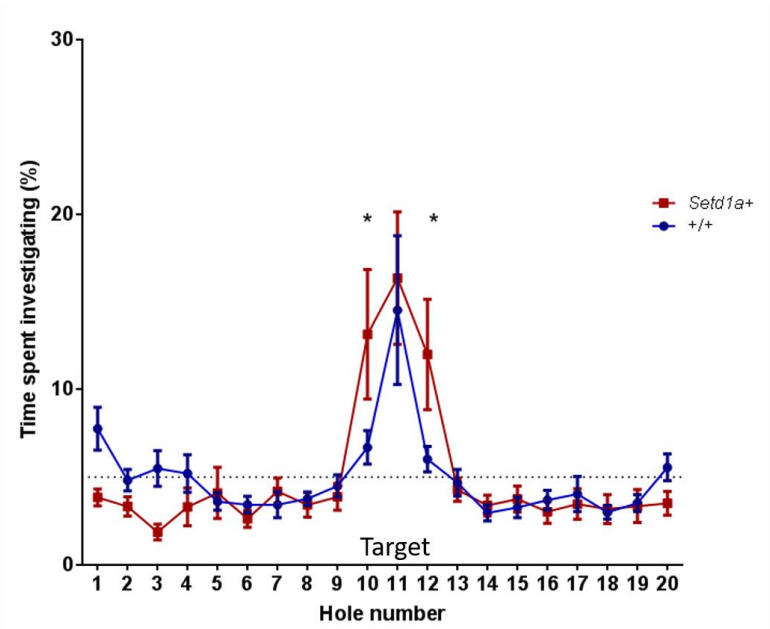
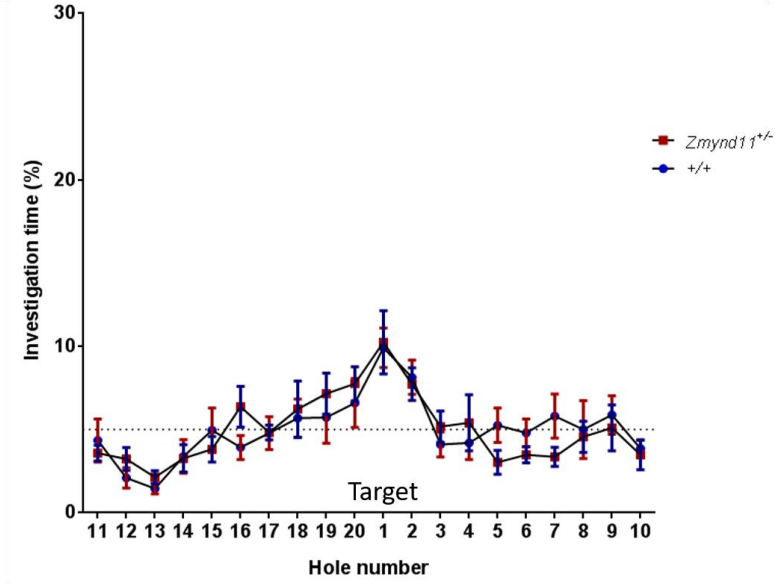


Figure 5.12 BCND 24h probe results. The percentage of time spent at each hole for controls (+/+; blue) and mutants (red) is plotted. For all four mouse lines there was a significant overall difference between % of time spent at the 20 holes, but no genotype difference in the investigation of the target (hole N1). **A.** *Setd1a* line, **B.** *Setd5* line, **C.** *Arid1b* line, **D**. *Zmynd11* line. Values are plotted as mean ± SEM.

A


Time spent investigating (%)
30
20
10
0
*
*
Setd1a+
+/+
Target
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Hole number

B


Investigation time (%)
30
20
10
0
Zmynd11+/-
+/+
Target
11 12 13 14 15 16 17 18 19 20 1 2 3 4 5 6 7 8 9 10
Hole number

C

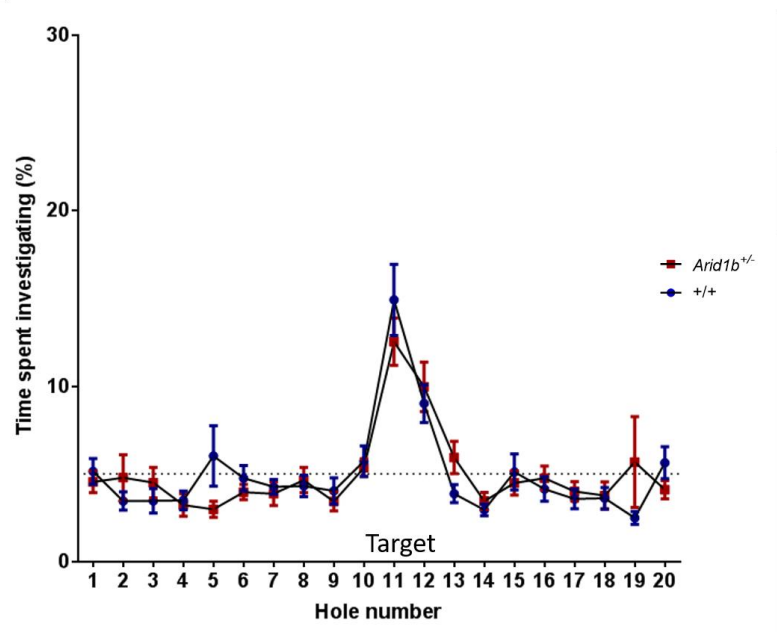

Time spent investigating (%)
30
20
10
0
Arid1b+/-
+/+
Target
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Hole number

Figure 5.13 BCND Barnes maze 72h probe results. The percentage of time spent at each hole for controls (+/+; blue) and mutants (red) is plotted, with the target at hole N11. **A.** *Setd1a* line. There was a significant genotype difference in the percentage of time spent at hole 10 and 12 (neighbouring holes to target), but no significant genotype difference at the target (hole 10: P=0.0171, hole 12: 0.0399, target: P= 0.9998). **B.** *Zmynd11* line. There was a significant overall difference between the percentage of time spent at the 20 holes, but no genotype difference in investigation. **C.** *Arid1b* line. There was a significant overall difference between the percentage of time spent at the 20 holes, but no genotype difference in investigation. Values are plotted as mean ± SEM.

5.2.5 Summary BCND results

The combined results across all four lines, as well as results from the previous testing of *Kptn* mice, are outlined in Table 9 (orange cells represent no phenotype, green cells represent the presence of phenotype, and white cells signify that the particular test was not done (ND) on that mouse line). By testing the same set of mice across a battery of tests, assessing several cognitive and behavioural domains, and screening multiple mouse models of novel distinct neurodevelopmental disorders in the same comparable manner using the BCDN screening paradigm, I was able to detect unique behavioural phenotypes for each loss-of-function mouse model investigated. However, recapitulating the full spectrum of phenotypes, as achieved in the *Kptn* mouse model, was challenging. This could be due to the variability of human phenotypes and mutations within each disorder.

Table 9 Summary of BCND results compared to *Kptn* work.

Assay	*Setd1a*$^{+/-}$	*Setd5*$^{+/-}$	*Arid1b*$^{+/-}$	*Zmynd11*$^{+/-}$	*Kptn*$^{-/-}$
Light/dark box (anxiety)		ND			Increased anxiety-like behaviour
Open Field (locomotion/activity)	Increased exploration and activity				Increased exploration and activity
Barnes Maze (spatial memory)	Specificity of memory retention (72h probe)	Slower acquisition learning during training	Slower acquisition learning during training		Memory impairment (72h probe)
Social recognition (olfactory memory)			24h memory impairment	24h memory impairment	24h memory impairment
Sociability (autistic-like traits)				ND	ND

5.3 Discussion and future perspectives

In this chapter, I describe the design of a behavioural and cognitive screening paradigm, BCND screen, which was used to characterize four loss-of-function mouse lines, each with a mutation in a gene, recently identified by Decipher Developmental Disorders (DDD) project, encoding a chromatin remodelling factor. Each of the four conditions modelled is associated with a variety of phenotypes that are variable between and within each disorder. By controlling environmental and genetic backgrounds in the mice, I can dissect the function of the candidate genes in a more focused manner than could be done in the patients. Because the patients in all four disorders are haploinsufficient, the four DDD mouse lines characterised here are dosage reduced for the genes under investigation. However, unlike with homozygous null models as *Kptn*, modelling haploinsufficient conditions is challenging as dosage sensitivity may vary between species. It is nevertheless preferable to use heterozygous mice when modelling heterozygous patients, in order to better recapitulate the effects of the mutations (Cuthbert et al., 2007).

5.3.1 Behavioural impairments

It is important to perform a careful assessment of all sensory function to all mouse lines entering the Behavioural and Cognitive Neurodevelopmental Disorder (BCND) screen, as sensory impairments may interfere with the performance of the mice and bias the results. All mouse lines underwent a systematic primary screening, assessing a broad range of physiological and neurological traits, as part of the Wellcome Trust Sanger Institute's Mouse Genetics Project (MGP), prior to being tested in the BCND screen (White *et al.,* 2013). In addition, many assays used in the BCND screen, such as social recognition, have specifically inbuilt criteria that would flag relevant impairments in sensory domains. For example, in social recognition, mice that had olfactory or social deficits would not reach relevant criteria on Day 1 and would be removed from testing. Nonetheless, subtle sensory defects could have been missed out.

Tasks that assess the overall behaviour of the mouse provide important information on the locomotor abilities and activity levels, which should be considered in the subsequent cognitive tests, as they may act as confounding factors. *Setd1a*$^{+/-}$ mice were the only mutant line to exhibit phenotypes on the first two initial behavioural tests – open field and light/dark assay. *Setd1a*$^{+/-}$ mice spent a long time moving in an open field and visited the centre more times, but did not spend a long time in the centre, when compared to their littermate controls. Besides, *Setd1a*$^{+/-}$ did not display behaviours in the light/dark assay that are associated with a reduced anxiety-like state, such as an increased transition to the light zone and longer time spent in the centre. Taken together the results are suggestive of an increased exploration and activity phenotype in *Setd1a*$^{+/-}$. No anxiety was detected in any of the lines, which is consistent with the lack of anxiety phenotypes in the patients.

Three mouse lines (*Setd1a*, *Setd5*, and *Arid1b*) were assessed for autistic-like phenotypes because the haploinsufficiency of *SETD1A*, *ARID1B*, and *SETD5* were associated on average with autistic features in 10% of the patients. Autism spectrum disorder (ASD) is a neurodevelopmental spectrum disorder diagnosed by social and communication deficits, as well as repetitive behaviours (American Psychiatric Association, 2013). ASD varies in severity and behavioural outputs, and therefore modelling autistic-like phenotypes is a

complex task (Moy et al., 2004; Shibutani et al., 2017). Mice are social species and their social behaviours have been extensively studied (Silverman et al., 2010; Winslow and Insel, 2002). The sociability test assesses social deficits in mice and is reliant on the innate preference of mice to spend time in investigating a novel conspecific rather an inanimate object (Moy et al., 2004; Silverman et al., 2010). Mice with social deficits will have lower levels of social approach behaviour than the controls. Heterozygous *Setd1a* and *Setd5* mice showed a preference for the mouse and not the object, suggestive of no detectable social deficit in these mouse lines. When the preference for the mouse zone was analysed in *Arid1b* line, the *Arid1b*$^{+/-}$ displayed an increased preference score for the mouse zone when compared to their littermate controls, implying no autistic-like phenotype in *Arid1b*$^{+/-}$. In fact, there was an increase in preference for the mouse zone in the *Arid1b*$^{+/-}$ mice. Social deficit phenotypes were reported in two other recent studies testing heterozygous *Arid1b* mice. However, there were differences between behaviours reported by the two studies (Celen et al., 2017; Shibutani et al., 2017). Shibutani et al. found milder behavioural abnormalities, and reduced social interaction only in a home-cage environment and not when tested using the sociability assay (Shibutani et al., 2017). The reported differences between mouse models with mutations in *Arid1b* may be reflective of the range of mutations and phenotypes associated with *ARID1B* pathologies (Celen et al., 2017; Yu et al., 2015; Shibutani et al., 2017).

5.3.2 Cognitive impairments

5.3.2.1 Olfactory-mediated memory for conspecific

In the social recognition (SOR) assay, the experimental design relies on the fact that the behaviour of test mouse towards a conspecific will change depending on whether the test mouse has the memory of prior interaction with the conspecific. Naïve mice presented with two stimuli investigate both mice without showing a particular preference for either. Conversely, as the test mice increase their exposure to a particular stimulus (Day 1, trial 1-4) they display a reduced interest in that stimulus over time. This is therefore a way to confirm the acquisition of memory for the stimulus over a short timeframe, as well as ensure that

the test animals have an intact olfactory system and do not have severe short-term memory impairment of up to 10 min. The mutant mice in each line displayed habituation to the stimulus over four trials and an increased investigation on trial 5 when a novel mouse was introduced. On Day 2 of social recognition, *Setd1a*$^{+/-}$ and *Setd5*$^{+/-}$ mice were able to distinguish between familiar stimulus (from Day 1) and an unfamiliar stimulus, displaying a preference for the unfamiliar, indicating the haploinsufficiency in these mice is not resulting in a deficit in 24h memory retention in this assay. *Zmynd11*$^{+/-}$ and *Arid1b*$^{+/-}$ mice spent on average an equivalent amount of time investigating both familiar and unfamiliar stimuli, without showing a significant preference for either of the stimuli, which implies a social recognition memory deficit. However, due to the lower number of *Arid1b* mice tested in social recognition than the estimated sample size (power analysis, Chapter 2, section 2.3.4.4), care must be taken when interpreting the results. Further studies with a larger sample size (n=9) would be necessary to confidently assess that *Arid1b* mutants show no evidence of learning.

Moreover, since *Arid1b*$^{+/-}$ mice had an indication of a mild social deficit in the sociability assay, this could affect the overall preference of the mutants to investigate novel conspecifics, which SOR assay is reliant upon (Dias et al., 2016; Moy et al., 2004; Sanchez-Andrade et al., 2005). Because diminished interest in novelty and avoidance of unfamiliar social entities has been described in some autistic individuals, social novelty preference test is widely used in autistic-like mouse models (American Psychiatric Association, 2013.; Moy et al, 2004; Moy et al, 2009; Shibutani et al., 2017; Silverman et al., 2010). Therefore, it is possible that the lack of preference for the unfamiliar over familiar is not due to lack of memory, but rather due to a deficit in the normal novelty-seeking behaviour of wildtype mice. In order to dissociate memory impairment from a social deficit, additional tests would need to be run, such as social novelty preference test, which can be run immediately after sociability, ensuring mice would still have the memory of the conspecific they were exposed to in sociability assay (Moy et al., 2004).

At first, the discrepancy between the results of SOR and Barnes may appear contradictory, as both assays are known to be dependent on hippocampal function. However, the two assays also test distinct cognitive domains, which may explain the phenotypic discrepancies.

The Barnes maze is a specialised spatial memory test and is primarily reliant on a functional hippocampus, while social recognition is an olfactory-mediated memory test and is additionally reliant on accessory olfactory systems, amygdala, and hypothalamus (Kogan et al., 2000; O'Leary et al., 2011). Furthermore, Barnes maze assay has the power to detect a larger spectrum of behaviours than the SOR assay, allowing uncovering subtle deficits which would not be possible by the SOR alone.

5.3.2.2 Spatial memory

Using the spatial memory test (Barnes maze), I found deficits in learning during training and 72h memory retention in individual mutant lines. *Arid1b* and *Setd5* mutants were slower during initial training (training N1), but did not display deficits in 24h memory retention, when subsequently tested during a 4-min probe trial, suggestive of a subtle learning deficit in *Arid1b*$^{+/-}$ and *Setd5*$^{+/-}$ mice. However, once *Arid1b* mutants acquire a strategy to find the target near the end of initial training, they no longer show learning deficits, during training N2, when the target location was moved. *Setd1a* and *Zmynd11* mutant mice performed as well as the controls in training and neither displayed deficits in the performance during the 24h probe. Of note, *Zmynd11* mice did not undergo training N2, but rather were re-trained briefly with the same target location after the 24h probe and then tested for 72h memory retention during the 72h probe. Although this made *Zmynd11* data not directly comparable to the other lines tested, the approach was useful in assessing whether a shorter protocol for memory retention in the Barnes maze could be used in future work, with the larger project in mind. Removing training N2 (the change in target location) allows for a cleaner comparison between 24h and 72h probe trials, without the potentially confounding effect of a change in training between the two memory tests (assessing different length of recall). It may, therefore, make sense, in future work, to schedule one day of training after the 24h probe and then test the mice 72h later. Other studies have addressed this by keeping both initial and reversal training the same length, and either kept a consistent interval between each probe trial or tested memory only after reversal training (O'Leary and Brown, 2009; O'Leary et al., 2011).

Out of three lines tested on the 72h probe, *Setd1a* mutants were the only ones to display a phenotype. The mutant mice did not spend any less time around the target zone, but they did spend longer investigating the adjacent holes on either side of the target. This may indicate a component of uncertainty for the specific location of the escape box, but an intact memory of a larger region, including the target zone. The results are unlikely to be due to non-cognitive factors, such as hyperactivity, because during the 24h probe both genotypes spent equivalent time near the target hole without an observed difference near the adjacent holes.

5.3.3 Concluding remarks and future perspectives

In summary, I designed a specialised behavioural and cognitive screen, BCND, and used it to test four DDD mouse models. The approach taken with the BCND screen allowed for the identification of specific phenotypic repertoires for each mouse line, allowing for comparisons to be made between several novel developmental disorders (Table 9). For example, in the Barnes maze I detected slower acquisition learning in *Arid1b*$^{+/-}$ and *Setd5*$^{+/-}$, but no memory deficit during probe sessions, while *Setd1a*$^{+/-}$ had a reduced specificity of spatial memory retention in the 72h probe which is suggestive of a mild memory deficit (Table 9). Moreover, the lessons learned about the practicalities and challenges of running such a screen helped shape the larger five-year project in Hurles group at the Sanger Institute.

In order to develop therapies for rare diseases it is important to focus on those disorders where postnatal intervention can have a reasonable prospect of improving the patient's condition, and where the pathogenicity of the disorder is not confined to a developmental time point. Although the four disorders modelled here have a broad array of different manifestations, if well targeted, elevating even a small subset of these could have a potentially significant benefit on the patient's quality of life. The conditional reversion approach in mice offers a good model for assessing whether reactivating postnatal gene expression can rescue some of the functional burden associated with the disorder. For this, it important to have robust, well-characterised phenotypes in which to assess rescue. This

approach has been successfully applied to the treatment of Rett syndrome (Buchovecky et al., 2013; Cobb et al., 2010; Guy et al., 2007). Moreover, postnatal reversal of a mutated *Shank3* gene, associated with autism, successfully rescued the social deficits and neuronal function in the adult mice (Mei et al., 2016). Encouragingly, in this thesis work, I was able to detect phenotypes in all of the mouse models characterised. The most promising for behavioural and cognitive rescue are S*etd1a*$^{+/-}$and *Arid1b*$^{+/-}$ mice, because they display the largest repertoire of cognitive and social deficits. A recent study has successfully rescued growth parameters in *Arid1b*$^{+/-}$ mutant mice and patients (Celen et al., 2017). Demonstrating a rescue of cognitive impairment would be a very powerful addition to the body of work implicating *Arid1b* as a promising therapeutic target for patients with *ARID1B* pathologies.

Chapter 6. Concluding discussion and future perspectives

6.1 Concluding remarks

Research focusing on the genetics of intellectual disability (ID) and other neurodevelopmental disorders has greatly expanded our understanding of the normal development of the brain and the fundamental mechanisms associated with cognition (Kaufman et al., 2010). In addition, advances in sequencing technology have facilitated an increase in diagnostic yield and enabled a higher discovery rate of rare and novel causes of ID. With a large number of novel neurodevelopmental disorders identified by genome-wide genotype-driven diagnostic efforts, there is an increasing demand for understanding the causal mechanisms underpinning these newly identified pathologies. In my thesis work, I focused on studying five such monogenic disorders associated with intellectual disability, by employing a multi-phase multi-model approach.

To characterise the five mouse models, I have established and refined a series of behavioural and cognitive assays, described in Chapter 3, aimed at detecting phenotypic abnormalities in the mice. By developing a robust phenotyping strategy, I then successfully characterised all five mouse models, capturing a spectrum of phenotypes reflective of those observed in patients. Combining this with morphometric, developmental, and molecular analyses I successfully characterised the *Kptn* mouse model in greater depth, described in Chapter 4, recapitulating all the main endophenotypes associated with a *KPTN*-related syndrome. Moreover, I gained further insight on the underlying pathogenic mechanisms associated with the disorder, opening the possibility of a treatment that could rescue some aspects of behavioural and cognitive impairments identified in the adults.

Adapting the testing paradigm used with *Kptn* mice, I designed a specialised behavioural and cognitive neurodevelopmental screen (BCND), described in Chapter 5, which I used to broadly characterize *Arid1b, Setd1a, Setd5, and Zmynd11* mutant mouse lines, modelling the four haploinsufficiency disorders. This approach allowed for identification of specific phenotypes for each of the four mouse lines, providing a platform for comparison between several novel developmental disorders, and acting as a refinement platform for a larger five-

year project, which is starting in Hurles group at the Sanger Institute, aimed at characterising a larger number of human neurodevelopmental disorders.

6.2 Characterizing intellectual disability in mice

The strength of an animal model of disease can be measured by the degree to which it parallels the human condition, enabling the study of pathology in greater depth. This is of particular importance for newly identified disorders, with unknown associated pathogenic mechanisms. Cognition involves multiple components, including various forms of learning, attention, decision-making, cognitive flexibility, and future planning (Nithianantharajah and Grant, 2013). These cognitive domains are homologous in mice and humans, with conserved underlying genetic and neuroanatomical mechanisms (Nithianantharajah et al., 2013). Therefore, in my thesis work, I have used inbred laboratory mice to dissect gene function associated with the human condition, while controlling for other environmental and genetic factors. The mouse models thus facilitate the study of the primary causative role of the genes under investigation and their specific contribution to the phenotypes associated with the human condition, without the confounding effects of other factors. This is especially useful for disorders that have multiple associated mutations and variable phenotypes across patients, such as the ones described in Chapter 5 (*ARID1B, SETD1A, SETD5, and ZMYND11* haploinsufficiencies), as the mouse models provide a more simplified framework for investigating the direct gene-phenotype relationships than is possible in the humans.

Many disease-causing mutations identified in the human genome cause haploinsufficiencies, yet heterozygous mice are less frequently characterised in knockout models compared to homozygous mice (Peça et al., 2011; Schmeisser et al., 2012; White et al., 2013). Although homozygous null alleles typically present the strongest phenotypic output and provide insight into the absolute requirement for a protein, they are not always reflective of the consequences of the human genotypes being modelled. It is therefore important to use the appropriate genotype of the mouse when modelling specific human disorders (Bakker et al., 1994; Cuthbert et al., 2007). In my thesis work, I have used homozygous mice (*Kptn*$^{-/-}$) for modelling homozygous or compound-heterozygous patients

with loss-of-function mutations in *KPTN,* as well as using heterozygous mice (*Arid1b*$^{+/-}$, *Setd1a*$^{+/-}$, *Setd5*$^{+/-}$, *Zmynd11*$^{+/-}$) for modelling disorders in which patients are heterozygous. Although the genotypes under investigation were chosen to best parallel the human condition and minimize animals used, testing heterozygous alongside the homozygous *Kptn* mutants can shed light on the dosage effect *Kptn* may have on the phenotype, as well as assess if there are any abnormalities detectable in the carriers, and should be explored in future studies. Likewise, testing the homozygous mutants for the latter four disorders would enable the assessment of the essential functions of the proteins. However, due to the essential function of chromatin remodelers, most of these disorders are homozygous lethal. This was, in fact, the case for *Arid1b, Zmynd11, Setd5,* and *Setd1a* mouse lines.

The range of mutations modelled in this thesis work, including inherited recessive and *de novo* dominant variants, facilitated the exploration of multiple causes of human neurodevelopmental disorders and intellectual disability in one study. For the heterozygous *de novo* disorders, I purposefully chose from a larger set for which clear evidence exists that the consequence of the mutations is very likely loss of function. This was often revealed by the presence of multiple affected individuals in the human cohorts carrying truncating premature stop codons. Because of this prior patient information, I felt that it was best to analyse the loss-of-function phenotypes in mouse models as a first approach to understanding these disorders. Further work linking different genotypes to a varied phenotypic outcome (so-called genotype-phenotype correlations) will entail engineering patient-specific missense mutations, with further phenotype testing. The advancements in CRISPR technology makes this a feasible proposition (Doudna and Charpentier, 2014). As a separate case, for the modelling of the *KPTN*-syndrome, further work will be required to determine whether the in-frame duplication variant identified in patients produces any functional protein, displaying hypomorphic, gain-of-function, or potentially neomorphic activity. These are valuable analyses that will further enrich the first-pass characterisation that we have undertaken. Moreover, additional considerations should be made to validate whether the loss-of-function mutations truly result in no protein product in the mouse lines. In the *Kptn* mouse model (Chapter 4, section 4.2.1), the *Kptn* expression levels were assessed in three different tissues confirming the complete loss of expression for *Kptn*.

Similar assessments should be carried out for other mutant lines. However, efforts were made to ensure true loss-of-function alleles were generated using sophisticated targeting strategies (Chapter 2, section 2.1) (Skarnes et al., 2011; Testa et al., 2004; White *et al.,* 2013).

The assessment of behavioural and cognitive deficits in rodent models of neurodevelopmental disorders is typically carried out on adult mice, after the maturation of the central nervous system (CNS) (Bakker et al., 1994; Guy et al., 2001). A limitation of characterising only adult mice, however, carries a risk of missing the effects of mutations on the development of the function of the CNS. Because developmental disorders have an onset during development, the assessment of neonatal, pre-weaning, and adolescent mice can identify age-dependent behavioural, cognitive, and motor deficits, as well as providing potential therapeutic targets in phases of development where an intervention can lead to a higher chance of recovery (Altafaj et al., 2001). Multiple tests have been established to assess mice at various postnatal stages (Branchi and Ricceri, 2002). For example, testing homing and passive avoidance learning allow for the detection of cognitive deficits in pups in the first two weeks of postnatal life, while more complex tasks such as spatial learning have been applied to post-weaning age animals (Alleva & Calamandrei 1986; Chapillonet al. 1995; Ricceri et al. 2000). It would be informative to explore *Kptn* deficiency on disease progression during development, using this approach. As discussed in Chapter 4, *Kptn* deficiency causes progressive macrocephaly (not present at birth) and cognitive deficits in the adult *Kptn*$^{-/-}$ mice. Once we determine the timing of the onset of macrocephaly in the *Kptn*$^{-/-}$ mice, the cognitive abilities of the *Kptn* mutant mice could be tested before and after this period, in order to clarify the relationship between the onsets of macrocephaly and cognitive development, and the contribution of the macrocephaly phenotype, if any, to the cognitive impairment observed in adults.

Another domain that would be a valuable addition to the testing repertoire for modelling neurodevelopmental disorders is ultrasonic vocalisation, characterised by frequencies ranging between 30 and 90 kHz (Branchi et al., 2001; Scattoni et al., 2009). Ultrasonic vocalization rate follows an ontogenetic profile, increasing in the first postnatal week (postnatal day (P) P5-6), peaking at around P6-8, decreasing to practically zero by the end of

second postnatal week and can quantitatively analysed with minimal handling of the pup (Branchi et al., 2001; Scattoni et al., 2009). Ultrasonic vocalization pattern analysis has therefore been applied to study several neurodevelopmental models, including Rett syndrome and Down syndrome mouse models (Holtzman et al., 1996; Picker et al., 2006). *Mecp2* null male and heterozygous female mutants (Rett syndrome model) exhibited an increase in ultrasonic vocalization as a response to social isolation, beginning at P5, while Ts65Dn mice (Down syndrome model) displayed delayed ultrasonic profile by four days (Holtzman et al., 1996; Picker et al., 2006). Moreover, it has been shown that mouse ultrasonic vocalization shares a common molecular mechanism with human speech and has been used to model speech-language disorders (Fujita et al., 2008). Ultrasonic vocalization can, therefore, be applied to model the speech and language delay detected in the DDD patients (Chapter 5). Speech and language delay is the second most frequent set of phenotypes in the four disorders under investigation in this thesis (Chapter 5), but has not been assessed in this thesis work and can, therefore, be an interesting future avenue to explore. Of note, care must be taken when different strains are tested, as considerable differences have been reported between different mouse strains (Hennessy et al., 1980; Roubertoux et al., 1996; Thornton et al., 2005).

6.3 Translational power of mouse modelling

An advantage of the current project design comes from access to clinical data and patients, through clinical collaborations, allowing for a two-way information flow, in which the mouse modelling is informed by clinical work and equally clinical work can be shaped by the research outcomes of mouse modelling. The initial clinical characterization of the undiagnosed patients is reliant on a broad set of clinical phenotype-driven observations. The patients then receive a genetic diagnosis, which clusters them into distinct conditions based on their genetic architecture. The phenotypes are then compiled for each condition and these are the set of phenotypes used to model these newly identified conditions in mice. Due to ease of genetic and environmental manipulations in mice, it is possible to perform in-depth initial characterization of the mice that is not always possible during the initial

clinical characterisation. For example, the patients with loss-of-function mutations in *KPTN,* described in Chapter 4, were diagnosed with global developmental delay and intellectual disability. I identified hippocampal-dependent memory deficits in *Kptn*$^{-/-}$ mice, but no difference in the performance of *Kptn*$^{-/-}$ when compared to wildtype controls in hippocampus-independent tasks. This provides an important insight into the effect of loss-of-function of *Kptn* on cognitive output, implicating the hippocampus as an affected brain region. These results can then guide further hypothesis-driven examination of the patients. Translational research of such nature, shaped by the work in rodents, includes the Pavlovian fear conditioning paradigm that facilitated mapping of fear circuitry in rodents being adapted for human studies to test fear inhibition (Jovanovic et al., 2005; Myers and Davis, 2004).

To draw meaningful comparisons between species it is important to ensure that equivalent cognitive domains are tested. The ideal way to achieve this is to test both mice and humans in the same manner. This can be achieved using automated touchscreen methods for rodents and the Cambridge Neuropsychological Test Automated Battery (CANTAB) touchscreen-based technology for patients (Bussey et al., 2012; Fray et al, 1996; Gur et al., 2001). Equivalent tasks include delayed matching to sample task, testing working memory, and five-choice serial-reaction task, testing cognitive flexibility (Mar et al., 2013; Aggleton et al., 1988; Bussey et al., 2002; Égerházi et al., 2007; Falconer et al., 2010; Semenova, 2012). *Kptn*$^{-/-}$ mice did not display deficits in a visual pairwise discrimination task, run on the touchscreen. It would be of value to test the patients in the CANTAB equivalent task, intra-extradimensional set-shifting, to assess whether they too perform normally. Similarly, determining whether patients have a spatial memory deficit, as was identified in the *Kptn* mutant mice, would allow for a more in-depth understanding of the impaired cognitive processes in the patients and allow for potentially better management of their symptoms. This could be assessed using a visuospatial learning and a touchscreen-based learning task in both mice (object-location paired associates learning task) and humans (paired-association learning task). The results outlined from the *Kptn* mouse model have been relayed to the clinicians working with the *KPTN*-related syndrome patients to encourage additional testing of the patients. Additional data on the concordance of our mouse model and the human

patients would heighten the relevance of this animal model, justifying further work. A similar approach was taken by Nithianantharajah et al. (2013), who tested *Dlg2*$^{-/-}$ mice, using visual discrimination acquisition and cognitive flexibility, and *DLG2* deficient patients on the equivalent task, intra-extradimensional set-shifting, using the touchscreen platform. Those authors showed that the patients displayed significantly more errors than healthy control subjects in the task, consistent with the phenotype of *Dlg2*$^{-/-}$ mice, highlighting the conserved gene-phenotype relationship between the two species (Nithianantharajah et al., 2013).

6.4 The use of phenotyping screens as a modelling approach

Although touchscreen technology proved to be a valuable tool when modelling the *Kptn*$^{-/-}$ mice (Chapter 4), the duration of the assays using this platform was incompatible with the requirements of screening paradigm, termed in my thesis as the Behavioural and Cognitive Neurodevelopmental Disorder (BCND) screen, in which each mouse line had to be tested in 4-6 weeks (Chapter 5). However, touchscreen technologies can be a powerful method for hypothesis-driven secondary phenotyping for those models. Moreover, other technologies, such as the automated IntelliCage system (NewBehaviour), can be used as part of the initial screening paradigm instead (Krackow et al., 2010). This technology allows for automated cognitive and behavioural testing of individual mice in an undisturbed group environment, and contains operant testing areas in the corners of the cage, enabling testing in the home-cage environment (Endo et al., 2011; Krackow et al., 2010; Mechan et al., 2009). This technology, therefore, maximises animal welfare, minimizes the human disturbance during testing, and allows for a high level of standardization. Furthermore, the testing can be done during active hours of the mouse, and not those artificially imposed by the working hours of the investigator. This technology can, therefore, be a powerful addition to the repertoire of tests in BCND. Since IntelliCage requires minimal human input, a second cohort can be run in parallel to the one being tested in the battery of tests outlined in Chapter 5, enriching the overall data collected. This method is less translatable to human testing than the touchscreen technologies but has the potential to capture deficits in the mice that would

otherwise be missed or confounded by other factors arising from a more stressful and artificial testing environment.

The power of using inbred strains of mice allows for isogenic animals of completely identical genetic makeup to be analysed, leading to a strong reduction in variance compared to outbred populations. This also allows to determine the degree of phenotypic variability in the absence of any contribution from genetic background, a common confounder in human studies. However, it is important to recognise that any choice of inbred mouse strain for performing phenotypic tests may lead to the identification of strain-specific effects of gene mutations (Homanics et al., 1999; Keane et al., 2011). This must be taken into account when interpreting the findings in our studies, and in any comparisons between animal models and their human patient counterparts. Establishing and testing disease models on alternative inbred strains may help explain the contribution of genetic background to phenotypic presentation, but given the effort, expense, and additional use of animals required, may not be justified in most cases.

The advent of human induced pluripotent stem cell (iPCS) in 2006, which are reprogrammed somatic cells that are induced into pluripotent state with embryonic stem cell-like properties, can provide a complementary approach for studying human disease and the associated biochemical and cellular processes (Takahashi and Yamanaka, 2006; Takahashi et al, 2007; Yu et al., 2007). hiPSC can be differentiated to various cell types, retaining the unique genetic signatures of the person they are derived from, allowing to study idiosyncratic patient-derived human neurons and their development. Therefore, a patient-derived iPSC strategy can be used not only to study the pathogenic pathways of all the conditions discussed in my thesis work but also to generate and characterize multiple iPSC lines reflective of the mutational spectrum in the patients in parallel. This could be particularly beneficial for frequently mutated genes, such as *ARID1B* that is associated with 11 independent loss-of-function mutations and over 200 phenotypes in 60 patients described here (DDD study, unpublished work).

Many neurodevelopmental diseases have been modelled using the iPSC technology, including Rett syndrome and fragile X syndrome, recapitulating relevant cellular and

molecular processes identified previously by mouse modelling and post-mortem human brain studies (Ananiev et al., 2011; Kim et al., 2011; Sheridan et al., 2011). However, these models had a one cell type approach, which poorly recapitulates the cellular complexity of the human condition. Subsequently, techniques have been established for co-culturing of more than one cell type derived from iPSCs, and more recently, 3-dimensional iPSC-derived organoids have been developed that enable the study of cell-cell interactions in more detail (Lancaster and Knoblich, 2014; Marchetto et al., 2008; McCracken et al., 2014). Although these technologies cannot recapitulate all the complexity of the brain, they provide a unique opportunity to dissect relevant processes associated with human condition using a human-derived tissue. A combinational approach, including animal modelling, genetics, and hypothesis-driven iPSC work can provide a comprehensive perspective on the disease being modelled. iPSC work will be conducted for the disorders discussed in this thesis, as part of the wider study.

Behaviour-centric phenotyping, which has been extensively used in my thesis work, is one way of identifying the effect of underlying mutations on the final output of the central nervous system. However, this can be complemented with other non-behaviour based approaches. The central nervous system (CNS) contains a large variety of cell types, each with a unique gene expression profile. Gene expression profiles are reflective of the cellular diversity in the brain and can be used to detect changes in cellular states (Lein et al., 2007). Whole genome expression profiling, therefore, can be used as a genome-wide functional analysis tool and has been used successfully to classify tumours by their expression profiles (Golub et al., 1999; Perou et al., 1999). Obtaining expression repertoires for multiple different mutant models should allow for the identification of co-regulated transcript groups and potential functional connections between different disorders. This concept of functional discovery using a compendium of gene expression data has been used in mammalian cells and has provided functional annotation of small molecules and genes in yeast (Hughes et al., 2000; Lamb et al., 2006). Such an approach could be extended to molecular phenotyping of neurodevelopmental disorders, which could facilitate the identification of shared pathogenic pathways and modules in these disorders and identify targets for therapy development. There are several limitations and challenges of using this

transcriptome-driven approach. Firstly, the heterogeneity of cell populations within and between brain regions and their interconnectivity may require single cell and system level expression profiling, to resolve the complexity. Secondly, although transcriptomic provides great insight into disease states, the link between gene expression and protein levels is not always straightforward (Anderson and Seilhamer, 1997; Greenbaum et al., 2003). Therefore, an ideal strategy would involve both transcriptomic and proteomic profiling of several brain regions across multiple mouse and human cellular models.

6.5 Treatment possibilities for intellectual disability

One of the research areas lacking sufficient research is in the area of treatment of neurodevelopmental disorders and intellectual disability. Treatment for ID falls into three main categories: (1) treatment centred on elevating the comorbidities associated with the disorder with the aim of improving the patient's functioning and life skills, such as pharmacologic treatment of behavioural disorders associated with fragile X syndrome (Lozano et al., 2015; Hagerman and Polussa, 2015); (2) treatment focused on early behavioural and cognitive intervention, appropriate education, and psychological support, which can improve the condition especially in milder cases of ID; (3) treatment that targets the underlying causes of ID, with the aim of mitigating the consequences of the disease, as has been done with dietary restriction of phenylalanine in phenylketonuria patients (Potocnik and Widhalm, 2013).

The latter category of treatment is most desirable as it has the potential to 'cure' the condition or prevent the disease from manifesting itself in the first place, but is also much harder to achieve due to the genetic complexity of the disorders, often limited knowledge of the underlying mechanisms, and confounding effects of comorbidities. A promising development in this area comes from the notion that it is possible to harness plasticity in the adult brain in order to improve cognitive function associated with developmental disorders in adulthood. Many mutations that cause developmental disorders affect genes that are expressed throughout life, and not only restricted to development. High level of plasticity that has been classically thought of being restricted to critical periods, time points

in development where the structure and function of the brain is most malleable to change (Hubel and Wiesel, 1970). For example, the acquisition of language in humans has been shown to be constrained by specific ontological windows (Lenneberg, 1967). However, subsequent studies have shown that the adult brain retains features of developmental plasticity, for example, the structural changes in spines, axons, and dendrites observed in the adult cortex, and this plasticity can be manipulated by altering levels of cortical inhibition (Hensch et al., 1998; Florence et al., 1998; Trachtenberg et al., 2002; Tailby et al., 2005; Sawtell et al., 2003; Sale et al., 2007) Increased levels of intracortical inhibition have been shown to alter the plasticity in the adult brain to the levels observed during development (Spolidoro et al., 2008). Moreover, McGraw et al. have demonstrated that an inducible model of Rett syndrome recapitulates the constitutive mutant allele phenotypes in the adult mice, implying the pathogenicity of the mutation is not confined to a critical period in development (McGraw et al., 2011). Taken together, these findings highlight the potential possibility of rescuing the cognitive impairments associated with the disease, by reversing the causative molecular deficits of the disorder in the adult brain.

The behavioural results from BCND provide an important initial characterization of these models. However, I would argue that on their own, the behavioural and cognitive phenotypes identified are not as informative as those from the *Kptn* mouse model. The patients with loss-of-function mutations in *KPTN* had a consistent set of phenotypes, facilitating testing for similar deficits in the *Kptn* mice, which is one of the reasons why this line was characterised more thoroughly. The other four disorders, with haploinsufficiencies in *ARID1B, SETD1A, SETD5, ZMYND11,* had variable phenotypes within each disorder. Therefore, because this variability in the four disorders made it challenging to parallel all the human phenotypes in the four mouse lines, the focus of the testing was to identify robust phenotypes in the mice which could be used when designing reversal experiments.

Conditional mouse models offer a unique opportunity to test whether the phenotypic consequences of mutations could be rescued postnatally. However, most studies have focused on conditionally inducing a mutant allele, rather than reverting a mutant allele to wildtype (McGraw et al., 2011). The latter approach has been shown to successfully rescue many of the neurological symptoms and increased longevity in a conditional loss-of-function

MECP2 model of Rett syndrome and has led to the identification of novel therapeutic targets for treatment of Rett syndrome (Buchovecky et al., 2013; McGraw et al., 2011). Therefore, the repertoire of behavioural phenotypes in the four mouse lines *Arid1b, Setd1a, Setd5, Zmynd11*, detected by the BCND screen, and the potential of all the mouse lines under investigation to be conditionally reversed to wildtype, provides a strong foundation for future reversion work and facilitate treatment development.

As discussed in Chapter 4, *KPTN* has recently been associated with a role as a negative regulator in mechanistic target of rapamycin (mTOR) signalling by Wolfson et al., (2017). This finding, in addition to the detected upregulation of differentially expressed genes associated with the mTOR pathway in the brain of *Kptn*$^{-/-}$ mice, and the hippocampus-dependent memory deficit observed in *Kptn*$^{-/-}$ from behavioural testing, together imply a potential mTOR signalling-induced brain defect due to loss of *Kptn* function, which could be targeted with rapamycin, a clinically approved drug. Studies targeting other negative regulators of mTOR that are involved in brain development, such as *Pten, Tsc1, Tsc2,* have shown that postnatal rapamycin treatment can rescue the neurological symptoms caused by the loss-of-function of these regulators (Ehninger et al., 2008; Harrison et al., 2009; Kwon et al., 2003; Shin et al., 2017). Moreover, pharmacological mTOR inhibitors have shown clinical benefits in neurological disorders, such as tuberous sclerosis, and there are ongoing clinical trials in epilepsy, autism, and dementia, making mTOR related treatment for *KPTN*-related syndrome a strong candidate for a successful bench to clinic translation.

An interesting possible link between the five genes is their associated roles in carcinogenesis and brain development. Mutations in chromatin remodelling factors (CRFs) that happen later in life have been extensively associated with multiple types of cancer (Helin et al., 2013; Vissers et al., 2016). Aberrant mTOR signalling has also been linked to many cancer types, and the recently identified role of *KPTN* as a mTOR negative regulator points to the possible involvement of *KPTN* in cancer, as has been shown with other negative mTOR regulators discussed above, such as *PTEN*, *TC1*, and *TC2* (Johannessen et al., 2005; Johannessen et al., 2008; Shuch et al., 2013). It remains to be seen whether the children with *KPTN*-related syndrome develop any of these comorbidities. However, the potential overlap between the biological processes involved in intellectual disability and

carcinogenesis, which have been extensively studied over decades, could be used to facilitate the development of more effective therapies for both.

References

1. Abramov, U., Puussaar, T., Raud, S., Kurrikoff, K. & Vasar, E. (2008) Behavioural differences between C57BL/6 and 129S6/SvEv strains are reinforced by environmental enrichment. *Neuroscience Letters,* **443**, 223-227.
2. Aggleton, J. P., Nicol, R. M., Huston, A. E. & Fairbairn, A. F. (1988) The performance of amnesic subjects on tests of experimental amnesia in animals: delayed matching-to-sample and concurrent learning. *Neuropsychologia,* **26**, 265-272.
3. Aimone, J. B., Li, Y., Lee, S. W., Clemenson, G. D., Deng, W. & Gage, F. H. (2014) Regulation and function of adult neurogenesis: from genes to cognition. *Physiological reviews,* **94**, 991-1026.
4. Albertson, D. G. & Pinkel, D. (2003) Genomic microarrays in human genetic disease and cancer. *Human Molecular Genetics,* **12**, R145-R152.
5. Alleva, E. & Calamandrei, G. (1986) Odor-aversion learning and retention span in neonatal mouse pups. *Behavioral and Neural Biology,* **46**, 348-357.
6. Allison, D. W., Gelfand, V. I., Spector, I. & Craig, A. M. (1998) Role of actin in anchoring postsynaptic receptors in cultured hippocampal neurons: differential attachment of NMDA versus AMPA receptors. *The Journal of neuroscience : the official journal of the Society for Neuroscience,* **18**, 2423-2436.
7. Altafaj, X., Dierssen, M., Baamonde, C., Martí, E., Visa, J., Guimerà, J.*, et al.* (2001) Neurodevelopmental delay, motor abnormalities and cognitive deficits in transgenic mice overexpressing Dyrk1A (minibrain), a murine model of Down's syndrome. *Human molecular genetics,* **10**, 1915-1923.
8. Amaral, D. G., Schumann, C. M. & Nordahl, C. W. (2008) Neuroanatomy of autism. *Trends in Neurosciences,* **31**, 137-145.
9. American Psychiatric, A. (2000) American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision. *American Psychiatric Association*.
10. Amir, R. E., Van den Veyver, I. B., Wan, M., Tran, C. Q., Francke, U. & Zoghbi, H. Y. (1999) Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nature Genetics,* **23**, 185-188.
11. Ananiev, G., Williams, E. C., Li, H. & Chang, Q. (2011) Isogenic Pairs of Wild Type and Mutant Induced Pluripotent Stem Cell (iPSC) Lines from Rett Syndrome Patients as In Vitro Disease Model. *PLoS ONE,* **6**, e25255-e25255.
12. Anderson, L. & Seilhamer, J. (1997) A comparison of selected mRNA and protein abundances in human liver. *Electrophoresis,* **18**, 533-537.
13. Anderson, M. I. & O'Mara, S. M. (2004) Responses of dorsal subicular neurons of rats during object exploration in an extended environment. *Experimental Brain Research,* **159**, 519-529.
14. O'Leary, T. P. & Brown, R. E. (2012) The effects of apparatus design and test procedure on learning and memory performance of C57BL/6J mice on the Barnes maze. *Journal of Neuroscience Methods,* **203**, 315-324.
15. Arqué, G., Fotaki, V., Fernández, D., de Lagrán, M. M., Arbonés, M. L. & Dierssen, M. (2008) Impaired Spatial Learning Strategies and Novel Object Recognition in Mice Haploinsufficient for the Dual Specificity Tyrosine-Regulated Kinase-1A (Dyrk1A). *PLoS ONE,* **3**, e2575-e2575.
16. Ashburner, J. & Friston, K. J. (2000) Voxel-Based Morphometry—The Methods. *NeuroImage,* **11**, 805-821.
17. Ayadi, A., Birling, M.-C., Bottomley, J., Bussell, J., Fuchs, H., Fray, M.*, et al.* (2012) Mouse large-scale phenotyping initiatives: overview of the European Mouse Disease Clinic (EUMODIC) and of the Wellcome Trust Sanger Institute Mouse Genetics Project. *Mammalian Genome,* **23**, 600-610.
18. Azevedo, F. A. C., Carvalho, L. R. B., Grinberg, L. T., Farfel, J. M., Ferretti, R. E. L., Leite, R. E. P.*, et al.* (2009) Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *The Journal of Comparative Neurology,* **513**, 532-541.
19. Babikian, T., Prins, M. L., Cai, Y., Barkhoudarian, G., Hartonian, I., Hovda, D. A.*, et al.* (2010) Molecular and physiological responses to juvenile traumatic brain injury: focus on growth and metabolism. *Developmental neuroscience,* **32**, 431-441.
20. Backx, L., Seuntjens, E., Devriendt, K., Vermeesch, J. & Van Esch, H. (2011) A balanced translocation t(6;14)(q25.3;q13.2) leading to reciprocal fusion transcripts in a patient with intellectual disability and agenesis of corpus callosum. *Cytogenetic and genome research,* **132**, 135-143.

21. Balderas, I., Rodriguez-Ortiz, C. J., Salgado-Tonda, P., Chavez-Hurtado, J., McGaugh, J. L. & Bermudez-Rattoni, F. (2008) The consolidation of object and context recognition memory involve different regions of the temporal lobe. *Learning & memory (Cold Spring Harbor, N.Y.),* **15**, 618-624.
22. Banks, P. J., Bashir, Z. I. & Brown, M. W. (2012) Recognition memory and synaptic plasticity in the perirhinal and prefrontal cortices. *Hippocampus,* **22**, 2012-2031.
23. Bansal, R., Gerber, A. J. & Peterson, B. S. (2008) Brain morphometry using anatomical magnetic resonance imaging. *Journal of the American Academy of Child and Adolescent Psychiatry,* **47**, 619-621.
24. Baple, E. L., Maroofian, R., Chioza, B. A., Izadi, M., Cross, H. E., Al-Turki, S.*, et al.* (2014) Mutations in KPTN cause macrocephaly, neurodevelopmental delay, and seizures. *American Journal of Human Genetics,* **94**, 87-94.
25. Basel-Vanagaite, L., Hershkovitz, T., Heyman, E., Raspall-Chaure, M., Kakar, N., Smirin-Yosef, P.*, et al.* (2013) Biallelic SZT2 mutations cause infantile encephalopathy with epilepsy and dysmorphic corpus callosum. *American Journal of Human Genetics,* **93**, 524-529.
26. Bearer, E. L. & Abraham, M. T. (1999) 2E4 (kaptin): a novel actin-associated protein from human blood platelets found in lamellipodia and the tips of the stereocilia of the inner ear. *European journal of cell biology,* **78**, 117-126.
27. Bearer, E. L., Chen, A. F., Chen, A. H., Li, Z., Mark, H. F., Smith, R. J.*, et al.* (2000) 2E4/Kaptin (KPTN)--a candidate gene for the hearing loss locus, DFNA4. *Annals of human genetics,* **64**, 189-196.
28. Bernier, P. J., Bedard, A., Vinet, J., Levesque, M. & Parent, A. (2002) Newly generated neurons in the amygdala and adjoining cortex of adult primates. *Proceedings of the National Academy of Sciences of the United States of America,* **99**, 11464-11469.
29. Beutler, E. (2006) Hemochromatosis: Genetics and Pathophysiology. *Annual Review of Medicine,* **57**, 331-347.
30. Bick, D., Fraser, P. C., Gutzeit, M. F., Harris, J. M., Hambuch, T. M., Helbling, D. C.*, et al.* (2017) Successful Application of Whole Genome Sequencing in a Medical Genetics Clinic. *Journal of pediatric genetics,* **6**, 61-76.
31. Biegel, J. A., Zhou, J. Y., Rorke, L. B., Stenstrom, C., Wainwright, L. M. & Fogelgren, B. (1999) Germ-line and acquired mutations of INI1 in atypical teratoid and rhabdoid tumors. *Cancer research,* **59**, 74-79.
32. Bishop, N. A., Lu, T. & Yankner, B. A. (2010) Neural mechanisms of ageing and cognitive decline. *Nature,* **464**, 529-535.
33. Blundell, J., Tabuchi, K., Bolliger, M. F., Blaiss, C. A., Brose, N., Liu, X.*, et al.* (2009) Increased anxiety-like behavior in mice lacking the inhibitory synapse cell adhesion molecule neuroligin 2. *Genes, Brain and Behavior,* **8**, 114-126.
34. Bockhorst, K. H., Narayana, P. A., Liu, R., Ahobila-Vijjula, P., Ramu, J., Kamel, M.*, et al.* (2008) Early postnatal development of rat brain: In vivo diffusion tensor imaging. *Journal of Neuroscience Research,* **86**, 1520-1528.
35. Bothe, G. W. M., Bolivar, V. J., Vedder, M. J. & Geistfeld, J. G. (2005) Behavioral differences among fourteen inbred mouse strains commonly used as disease models. *Comparative medicine,* **55**, 326-334.
36. Boyle, C. A., Boulet, S., Schieve, L. A., Cohen, R. A., Blumberg, S. J., Yeargin-Allsopp, M.*, et al.* (2011) Trends in the Prevalence of Developmental Disabilities in US Children, 1997-2008. *PEDIATRICS,* **127**, 1034-1042.
37. Branchi, I. & Ricceri, L. (2002) Transgenic and knock-out mouse pups: the growing need for behavioral analysis. *Genes, Brain and Behavior,* **1**, 135-141.
38. Branchi, I., Santucci, D. & Alleva, E. (2001) Ultrasonic vocalisation emitted by infant rodents: a tool for assessment of neurobehavioural development. *Behavioural Brain Research,* **125**, 49-56.
39. Brigman, G., Webb, L. & Campbell, C. (2007) Building Skills for School Success: Improving the Academic and Social Competence of Students. *Professional School Counseling,* **10**, 279-288.
40. Brigman, J. L., Daut, R. A., Wright, T., Gunduz-Cinar, O., Graybeal, C., Davis, M. I.*, et al.* (2013) GluN2B in corticostriatal circuits governs choice learning and choice shifting. *Nature Neuroscience,* **16**, 1101-1110.
41. Broadbent, N. J., Squire, L. R. & Clark, R. E. (2004) Spatial memory, recognition memory, and the hippocampus. *Proceedings of the National Academy of Sciences of the United States of America,* **101**, 14515-14520.
42. Brown, M. W. & Aggleton, J. P. (2001) Recognition memory: What are the roles of the perirhinal cortex and hippocampus? *Nature Reviews Neuroscience,* **2**, 51-61.
43. Bubser, M., Bridges, T. M., Dencker, D., Gould, R. W., Grannan, M., Noetzel, M. J.*, et al.* (2014) Selective Activation of M ₄ Muscarinic Acetylcholine Receptors Reverses MK-801-Induced Behavioral Impairments and Enhances Associative Learning in Rodents. *ACS Chemical Neuroscience,* **5**, 920-942.
44. Bućan, M. & Abel, T. (2002) The mouse: genetics meets behaviour. *Nature Reviews Genetics,* **3**, 114-123.

45. Buchovecky, C. M., Turley, S. D., Brown, H. M., Kyle, S. M., McDonald, J. G., Liu, B.*, et al.* (2013) A suppressor screen in Mecp2 mutant mice implicates cholesterol metabolism in Rett syndrome. *Nature Genetics,* **45**, 1013-1020.
46. Bussey, T. J. J., Holmes, A., Lyon, L., Mar, A. C. C., McAllister, K. A. L. A. L., Nithianantharajah, J.*, et al.* (2012) New translational assays for preclinical modelling of cognition in schizophrenia: The touchscreen testing method for mice and rats. *Neuropharmacology,* **62**, 1191-1203.
47. Bussey, T. J., Padain, T. L., Skillings, E. A., Winters, B. D., Morton, A. J. & Saksida, L. M. (2008) The touchscreen cognitive testing method for rodents: how to get the best out of your rat. *Learning & memory (Cold Spring Harbor, N.Y.),* **15**, 516-523.
48. Butler, M. G., Meaney, F. J., Palmer, C. G., Opitz, J. M. & Reynolds, J. F. (1986) Clinical and cytogenetic survey of 39 individuals with Prader-Labhart-Willi syndrome. *American Journal of Medical Genetics,* **23**, 793-809.
49. Catalani, A., Sabbatini, M., Consoli, C., Cinque, C., Tomassoni, D., Azmitia, E.*, et al.* (2002) Glial fibrillary acidic protein immunoreactive astrocytes in developing rat hippocampus. *Mechanisms of Ageing and Development,* **123**, 481-490.
50. Celen, C., Chuang, J.-C., Luo, X., Nijem, N., Walker, A. K., Chen, F.*, et al.* (2017) Arid1b haploinsufficient mice reveal neuropsychiatric phenotypes and reversible causes of growth impairment. *eLife,* **6**.
51. Chapillon, P., Roullet, P. & Lassalle, J. M. (1995) Ontogeny of orientation and spatial learning on the radial maze in mices. *Developmental Psychobiology,* **28**, 429-442.
52. Chelly, J. & Mandel, J.-L. (2001) Monogenic causes of X-linked mental retardation. *Nature Reviews Genetics,* **2**, 669-680.
53. Chen, X., Ba, Y., Ma, L., Cai, X., Yin, Y., Wang, K.*, et al.* (2008) Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases. *Cell Research,* **18**, 997-1006.
54. Chesler, E. J., Wilson, S. G., Lariviere, W. R., Rodriguez-Zas, S. L. & Mogil, J. S. (2002) Identification and ranking of genetic and laboratory environment factors influencing a behavioral trait, thermal nociception, via computational analysis of a large data archive. *Neuroscience and biobehavioral reviews,* **26**, 907-923.
55. Chinwalla, A. T., Cook, L. L., Delehaunty, K. D., Fewell, G. A., Fulton, L. A., Fulton, R. S.*, et al.* (2002) Initial sequencing and comparative analysis of the mouse genome. *Nature,* **420**, 520-562.
56. Choi, Y. J., Di Nardo, A., Kramvis, I., Meikle, L., Kwiatkowski, D. J., Sahin, M.*, et al.* (2008) Tuberous sclerosis complex proteins control axon formation. *Genes & Development,* **22**, 2485-2495.
57. Chudasama, Y. & Robbins, T. W. (2003) Dissociable contributions of the orbitofrontal and infralimbic cortex to pavlovian autoshaping and discrimination reversal learning: further evidence for the functional heterogeneity of the rodent frontal cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience,* **23**, 8771-8780.
58. Chudasama, Y., Bussey, T. J. & Muir, J. L. (2001) Effects of selective thalamic and prelimbic cortex lesions on two types of visual discrimination and reversal learning. *European Journal of Neuroscience,* **14**, 1009-1020.
59. Clancy, B., Darlington, R. B. & Finlay, B. L. (2000) The course of human events: predicting the timing of primate neural development. *Developmental Science,* **3**, 57-66.
60. Clancy, B., Finlay, B. L., Darlington, R. B. & Anand, K. J. S. (2007) Extrapolating brain development from experimental species to humans. *NeuroToxicology,* **28**, 931-937.
61. Cobb, S., Guy, J. & Bird, A. (2010) Reversibility of functional deficits in experimental models of Rett syndrome. *Biochemical Society transactions,* **38**, 498-506.
62. Cobben, J. M., Weiss, M. M., van Dijk, F. S., De Reuver, R., de Kruiff, C., Pondaag, W.*, et al.* (2014) A de novo mutation in ZMYND11, a candidate gene for 10p15.3 deletion syndrome, is associated with syndromic intellectual disability. *European Journal of Medical Genetics,* **57**, 636-638.
63. Coe, B. P., Witherspoon, K., Rosenfeld, J. A., van Bon, B. W. M., Vulto-van Silfhout, A. T., Bosco, P.*, et al.* (2014) Refining analyses of copy number variation identifies specific genes associated with developmental delay. *Nature Genetics,* **46**, 1063-1071.
64. Coffee, B., Keith, K., Albizua, I., Malone, T., Mowrey, J., Sherman, S. L.*, et al.* (2009) Incidence of Fragile X Syndrome by Newborn Screening for Methylated FMR1 DNA. *The American Journal of Human Genetics,* **85**, 503-514.
65. Collins, F. S. & McKusick, V. A. (2001) Implications of the Human Genome Project for Medical Science. *JAMA,* **285**, 540-540.

66. Consorthium, T. D.-B. F. X., Bakker, C. E., Verheij, C., Willemsen, R., Helm, R. v. d., Oerlemans, F.*, et al.* (1994) Fmr1 knockout mice: a model to study fragile X mental retardation. The Dutch-Belgian Fragile X Consortium. *Cell,* **78**, 23-33.
67. Contet, C., Rawlins, J. N. P. & Deacon, R. M. J. (2001) A comparison of 129S2/SvHsd and C57BL/6JOlaHsd mice on a test battery assessing sensorimotor, affective and cognitive behaviours: Implications for the study of genetically modified mice. *Behavioural Brain Research,* **124**, 33-46.
68. Cooper, S. A., Smiley, E., Morrison, J., Williamson, A. & Allan, L. (2007) Mental ill-health in adults with intellectual disabilities: prevalence and associated factors. *The British Journal of Psychiatry,* **190**, 27-35.
69. Copping, N. A., Berg, E. L., Foley, G. M., Schaffler, M. D., Onaga, B. L., Buscher, N.*, et al.* (2017) Touchscreen learning deficits and normal social approach behavior in the Shank3B model of Phelan–McDermid Syndrome and autism. *Neuroscience,* **345**, 155-165.
70. Courchesne, E. (1997) Brainstem, cerebellar and limbic neuroanatomical abnormalities in autism, pp. 269-278.
71. Crawley, J. & Goodwin, F. K. (1980) Preliminary report of a simple animal behavior model for the anxiolytic effects of benzodiazepines. *Pharmacology, biochemistry, and behavior,* **13**, 167-170.
72. Crawley, J. N. & Paylor, R. (1997) A proposed test battery and constellations of specific behavioral paradigms to investigate the behavioral phenotypes of transgenic and knockout mice. *Hormones and behavior,* **31**, 197-211.
73. Crino, P. B. (2011) MTOR: A pathogenic signaling pathway in developmental brain malformations, pp. 734-742. Elsevier Current Trends.
74. Cuthbert, P. C., Stanford, L. E., Coba, M. P., Ainge, J. A., Fink, A. E., Opazo, P.*, et al.* (2007) Synapse-associated protein 102/dlgh3 couples the NMDA receptor to specific plasticity pathways and learning strategies. *The Journal of neuroscience : the official journal of the Society for Neuroscience,* **27**, 2673-2682.
75. de Ligt, J., Willemsen, M. H., van Bon, B. W. M., Kleefstra, T., Yntema, H. G., Kroes, T.*, et al.* (2012) Diagnostic Exome Sequencing in Persons with Severe Intellectual Disability. *New England Journal of Medicine,* **367**, 1921-1929.
76. de Vries, B. B. A. A., Pfundt, R., Leisink, M., Koolen, D. A., Vissers, L. E. L. M. L. M., Janssen, I. M.*, et al.* (2005) Diagnostic genome profiling in mental retardation. *American journal of human genetics,* **77**, 606-616.
77. Deciphering Developmental Disorders Study, Fitzgerald, T. W., Gerety, S. S., Jones, W. D., van Kogelenberg, M., King, D. A., McRae, J.*, et al.* (2015) Large-scale discovery of novel genetic causes of developmental disorders. *Nature,* **519**, 223-228.
78. Dekaban, A. S. & Sadowsky, D. (1978) Changes in brain weights during the span of human life: Relation of brain weights to body heights and body weights. *Annals of Neurology,* **4**, 345-356.
79. Dere, E., Huston, J. P. & De Souza Silva, M. A. (2005) Integrated memory for objects, places, and temporal order: Evidence for episodic-like memory in mice. *Neurobiology of Learning and Memory,* **84**, 214-221.
80. DeScipio, C., Conlin, L., Rosenfeld, J., Tepperberg, J., Pasion, R., Patel, A.*, et al.* (2012) Subtelomeric deletion of chromosome 10p15.3: clinical findings and molecular cytogenetic characterization. *American journal of medical genetics. Part A,* **158A**, 2152-2161.
81. Dias, C., Estruch, Sara B., Graham, Sarah A., McRae, J., Sawiak, Stephen J., Hurst, Jane A.*, et al.* (2016) BCL11A Haploinsufficiency Causes an Intellectual Disability Syndrome and Dysregulates Transcription. *The American Journal of Human Genetics,* **99**, 253-274.
82. Dias, R., Robbins, T. W. & Roberts, A. C. (1996) Dissociation in prefrontal cortex of affective and attentional shifts. *Nature,* **380**, 69-72.
83. Diesen, C., Saarinen, A., Pihko, H., Rosenlew, C., Cormand, B., Dobyns, W. B.*, et al.* (2004) POMGnT1 mutation and phenotypic spectrum in muscle-eye-brain disease. *Journal of Medical Genetics,* **41**, e115-e115.
84. Dillon, S. C., Zhang, X., Trievel, R. C. & Cheng, X. (2005) The SET-domain protein superfamily: protein lysine methyltransferases. *Genome biology,* **6**, 227-227.
85. Dobbing, J. & Sands, J. (1973) Quantitative growth and development of human brain. *Archives of disease in childhood,* **48**, 757-767.
86. Doudna, J. A. & Charpentier, E. (2014) Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science (New York, N.Y.),* **346**, 1258096-1258096.
87. Du, A. T., Schuv, N., Amend, D., Laakso, M. P., Hsu, Y. Y., Jagust, W. J.*, et al.* (2001) Magnetic resonance imaging of the entorhinal cortex and hippocampus in mild cognitive impairment and Alzheimer's disease. *J Neurol Neurosurg Psychiatry,* **71**, 441-447.

88. Duan, J.-H., Wang, H.-Q., Xu, J., Lin, X., Chen, S.-Q., Kang, Z.*, et al.* (2006) White matter damage of patients with Alzheimer's disease correlated with the decreased cognitive function. *Surgical and Radiologic Anatomy,* **28**, 150-156.
89. Dudley, J. T., Deshpande, T. & Butte, A. J. (2011) Exploiting drug-disease relationships for computational drug repositioning. *Briefings in Bioinformatics,* **12**, 303-311.
90. Durkin, M. (2002) The epidemiology of developmental disabilities in low-income countries. *Mental Retardation and Developmental Disabilities Research Reviews,* **8**, 206-211.
91. Égerházi, A., Berecz, R., Bartók, E. & Degrell, I. (2007) Automated Neuropsychological Test Battery (CANTAB) in mild cognitive impairment and in Alzheimer's disease. *Progress in Neuro-Psychopharmacology and Biological Psychiatry,* **31**, 746-751.
92. Ehninger, D., Li, W., Fox, K., Stryker, M.P., Silva, A.J. (2008) Reversing Neurodevelopmental Disorders in Adults. *Neuron*, **60**, 950-960.
93. Einfeld, S. L. & Tonge, B. J. (2007) Population prevalence of psychopathology in children and adolescents with intellectual disability: II epidemiological findings. *Journal of Intellectual Disability Research,* **40**, 99-109.
94. Emerson, E. (2007) Poverty and people with intellectual disabilities. *Mental Retardation and Developmental Disabilities Research Reviews,* **13**, 107-113.
95. Endo, T., Maekawa, F., Võikar, V., Haijima, A., Uemura, Y., Zhang, Y.*, et al.* (2011) Automated test of behavioral flexibility in mice using a behavioral sequencing task in IntelliCage. *Behavioural Brain Research,* **221**, 172-181.
96. Ennaceur, A. & Delacour, J. (1988) A new one-trial test for neurobiological studies of memory in rats. 1: Behavioral data. *Behavioural brain research,* **31**, 47-59.
97. Ennaceur, A., Neave, N. & Aggleton, J. P. (1997) Spontaneous object recognition and object location memory in rats: the effects of lesions in the cingulate cortices, the medial prefrontal cortex, the cingulum bundle and the fornix. *Experimental Brain Research,* **113**, 509-519.
98. Epp, J. R., Silva Mera, R., Köhler, S., Josselyn, S. A. & Frankland, P. W. (2016) Neurogenesis-mediated forgetting minimizes proactive interference. *Nature Communications,* **7**, 10838-10838.
99. Fahrner, J. A. & Bjornsson, H. T. (2014) Mendelian Disorders of the Epigenetic Machinery: Tipping the Balance of Chromatin States. *Annual Review of Genomics and Human Genetics,* **15**, 269-293.
100. Falconer, D. W., Cleland, J., Fielding, S. & Reid, I. C. (2010) Using the Cambridge Neuropsychological Test Automated Battery (CANTAB) to assess the cognitive impact of electroconvulsive therapy on visual and visuospatial memory. *Psychological medicine,* **40**, 1017-1025.
101. Ferguson, J. N., Aldag, J. M., Insel, T. R. & Young, L. J. (2001) Oxytocin in the medial amygdala is essential for social recognition in the mouse. *The Journal of neuroscience : the official journal of the Society for Neuroscience,* **21**, 8278-8285.
102. Fernandez, F. & Garner, C. C. (2007) Object recognition memory is conserved in Ts1Cje, a mouse model of Down syndrome. *Neuroscience Letters,* **421**, 137-141.
103. Filippakopoulos, P., Qi, J., Picaud, S., Shen, Y., Smith, W. B., Fedorov, O.*, et al.* (2010) Selective inhibition of BET bromodomains. *Nature,* **468**, 1067-1073.
104. Firth, H. V. & Wright, C. F. (2011) The Deciphering Developmental Disorders (DDD) study. *Developmental Medicine & Child Neurology,* **53**, 702-703.
105. Florence, S. L., Taub, H. B. & Kaas, J. H. (1998) Large-Scale Sprouting of Cortical Connections After Peripheral Injury in Adult Macaque Monkeys. *Science,* **282**, 1117-1121.
106. Flores-Alcantar, A., Gonzalez-Sandoval, A., Escalante-Alcalde, D. & Lomelí, H. (2011) Dynamics of expression of ARID1A and ARID1B subunits in mouse embryos and in cells during the cell cycle. *Cell and Tissue Research,* **345**, 137-148.
107. Forman, M. S., Squier, W., Dobyns, W. B. & Golden, J. A. (2005) Genotypically Defined Lissencephalies Show Distinct Pathologies. *Journal of Neuropathology and Experimental Neurology,* **64**, 847-857.
108. Forwood, S. E., Winters, B. D. & Bussey, T. J. (2005) Hippocampal lesions that abolish spatial maze performance spare object recognition memory at delays of up to 48 hours. *Hippocampus,* **15**, 347-355.
109. Frankel, W. N., Yang, Y., Mahaffey, C. L., Beyer, B. J. & O'Brien, T. P. (2009) Szt2, a novel gene for seizure threshold in mice. *Genes, Brain and Behavior,* **8**, 568-576.
110. Frey, J. W., Jacobs, B. L., Goodman, C. A. & Hornberger, T. A. (2014) A role for Raptor phosphorylation in the mechanical activation of mTOR signaling. *Cellular Signalling,* **26**, 313-322.
111. Frias, M. A., Thoreen, C. C., Jaffe, J. D., Schroder, W., Sculley, T., Carr, S. A.*, et al.* (2006) mSin1 Is Necessary for Akt/PKB Phosphorylation, and Its Isoforms Define Three Distinct mTORC2s. *Current Biology,* **16**, 1865-1870.

112. Fuchs, H., Gailus-Durner, V., Neschen, S., Adler, T., Afonso, L. C., Aguilar-Pimentel, J. A., *et al.* (2012) Innovations in phenotyping of mouse models in the German Mouse Clinic. *Mammalian Genome,* **23**, 611-622.
113. Fujita, E., Tanabe, Y., Shiota, A., Ueda, M., Suwa, K., Momoi, M. Y., *et al.* (2008) Ultrasonic vocalization impairment of Foxp2 (R552H) knockin mice related to speech-language disorder and abnormality of Purkinje cells. *Proceedings of the National Academy of Sciences of the United States of America,* **105**, 3117-3122.
114. Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., *et al.* (1999) Brain development during childhood and adolescence: a longitudinal MRI study. *Nature Neuroscience,* **2**, 861-863.
115. Gilbert, P. E. & Kesner, R. P. (2003) Recognition memory for complex visual discriminations is influenced by stimulus interference in rodents with perirhinal cortex damage. *Learning & memory (Cold Spring Harbor, N.Y.),* **10**, 525-530.
116. Gilissen, C., Hehir-Kwa, J. Y., Thung, D. T., van de Vorst, M., van Bon, B. W. M., Willemsen, M. H., *et al.* (2014) No Title. *Nature,* **511**, 344-347.
117. Gogliotti, R. G., Senter, R. K., Fisher, N. M., Adams, J., Zamorano, R., Walker, A. G., *et al.* (2017) mGlu7 potentiation rescues cognitive, social, and respiratory phenotypes in a mouse model of Rett syndrome. *Science translational medicine,* **9**, eaai7459-eaai7459.
118. Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J. P., *et al.* (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science (New York, N.Y.),* **286**, 531-537.
119. Gould, E. (2007) How widespread is adult neurogenesis in mammals? *Nature Reviews Neuroscience,* **8**, 481-488.
120. Graybeal, C., Bachu, M., Mozhui, K., Saksida, L. M., Bussey, T. J., Sagalyn, E., *et al.* (2014) Strains and Stressors: An Analysis of Touchscreen Learning in Genetically Diverse Mouse Strains. *PLoS ONE,* **9**, e87745-e87745.
121. Grayton, H. M., Fernandes, C., Rujescu, D. & Collier, D. A. (2012) Copy number variations in neurodevelopmental disorders, pp. 81-91. Pergamon.
122. Green, R. E., Krause, J., Briggs, A. W., Maricic, T., Stenzel, U., Kircher, M., *et al.* (2010) A draft sequence of the Neandertal genome. *Science (New York, N.Y.),* **328**, 710-722.
123. Greenbaum, D., Colangelo, C., Williams, K. & Gerstein, M. (2003) Comparing protein abundance and mRNA expression levels on a genomic scale. *Genome Biology,* **4**, 117-117.
124. Grozeva, D., Carss, K., Spasic-Boskovic, O., Parker, Michael J., Archer, H., Firth, Helen V., *et al.* (2014) De Novo Loss-of-Function Mutations in SETD5, Encoding a Methyltransferase in a 3p25 Microdeletion Syndrome Critical Region, Cause Intellectual Disability. *The American Journal of Human Genetics,* **94**, 618-624.
125. Gu, Y., Lindner, J., Kumar, A., Yuan, W. & Magnuson, M. A. (2011) Rictor/mTORC2 is essential for maintaining a balance between beta-cell proliferation and cell size. *Diabetes,* **60**, 827-837.
126. Guan, J., Gunn, A. J., Sirimanne, E. S., Tuffin, J., Gunning, M. I., Clark, R., *et al.* (2000) The Window of Opportunity for Neuronal Rescue with Insulin-Like Growth Factor-1 after Hypoxia—Ischemia in Rats is Critically Modulated by Cerebral Temperature during Recovery. *Journal of Cerebral Blood Flow & Metabolism,* **20**, 513-519.
127. Guertin, D. A., Stevens, D. M., Thoreen, C. C., Burds, A. A., Kalaany, N. Y., Moffat, J., *et al.* (2006) Ablation in Mice of the mTORC Components raptor, rictor, or mLST8 Reveals that mTORC2 Is Required for Signaling to Akt-FOXO and PKCα, but Not S6K1. *Developmental Cell,* **11**, 859-871.
128. Gur, R. C., Ragland, J. D., Moberg, P. J., Turner, T. H., Bilker, W. B., Kohler, C., *et al.* (2001) Computerized neurocognitive scanning: I. Methodology and validation in healthy people. *Neuropsychopharmacology,* **25**, 766-776.
129. Guy, J., Gan, J., Selfridge, J., Cobb, S. & Bird, A. (2007) Reversal of neurological defects in a mouse model of Rett syndrome. *Science (New York, N.Y.),* **315**, 1143-1147.
130. Guy, J., Hendrich, B., Holmes, M., Martin, J. E. & Bird, A. (2001) A mouse Mecp2-null mutation causes neurological symptoms that mimic Rett syndrome. *Nature Genetics,* **27**, 322-326.
131. Hagerman, R. J. & Polussa, J. (2015) Treatment of the psychiatric problems associated with fragile X syndrome. *Current opinion in psychiatry,* **28**, 107-107.
132. Halgren, C., Kjaergaard, S., Bak, M., Hansen, C., El-Schich, Z., Anderson, C. M., *et al.* (2012) Corpus callosum abnormalities, intellectual disability, speech impairment, and autism in patients with haploinsufficiency of ARID1B. *Clinical Genetics,* **82**, 248-255.

133. Hamamy, H. (2012) Consanguineous marriages : Preconception consultation in primary health care settings. *Journal of community genetics,* **3**, 185-192.
134. Hamamy, H., Antonarakis, S. E., Cavalli-Sforza, L. L., Temtamy, S., Romeo, G., Kate, L. P. T.*, et al.* (2011) Consanguineous marriages, pearls and perils: Geneva International Consanguinity Workshop Report. *Genetics in Medicine,* **13**, 841-847.
135. Hamdan, F. F., Srour, M., Capo-Chichi, J.-M., Daoud, H., Nassif, C., Patry, L.*, et al.* (2014) De Novo Mutations in Moderate or Severe Intellectual Disability. *PLoS Genetics,* **10**, e1004772-e1004772.
136. Hamilton, D. A. & Brigman, J. L. (2015) Behavioral flexibility in rats and mice: contributions of distinct frontocortical regions. *Genes, Brain and Behavior,* **14**, 4-21.
137. Hammond, R., Tull, L. E. & Stackman, R. W. (2004) On the delay-dependent involvement of the hippocampus in object recognition memory. *Neurobiology of Learning and Memory,* **82**, 26-34.
138. Handy, D. E., Castro, R. & Loscalzo, J. (2011) Epigenetic modifications: basic mechanisms and role in cardiovascular disease. *Circulation,* **123**, 2145-2156.
139. Harrison, F. E., Hosseini, A. H. & McDonald, M. P. (2009) Endogenous anxiety and stress responses in water maze and Barnes maze spatial memory tasks. *Behavioural Brain Research,* **198**, 247-251.
140. Harrison, F. E., Reiserer, R. S., Tomarken, A. J. & McDonald, M. P. (2006) Spatial and nonspatial escape strategies in the Barnes maze. *Learning & memory (Cold Spring Harbor, N.Y.),* **13**, 809-819.
141. Hascoët, M. & Bourin, M. (2009) The mouse light-dark box test, pp. 197-223.
142. Hateboer, G., Gennissen, A., Ramos, Y. F., Kerkhoven, R. M., Sonntag-Buck, V., Stunnenberg, H. G.*, et al.* (1995) BS69, a novel adenovirus E1A-associated protein that inhibits E1A transactivation. *The EMBO journal,* **14**, 3159-3169.
143. Helin, K. & Dhanak, D. (2013) Chromatin proteins and modifications as drug targets. *Nature,* **502**, 480-488.
144. Hennessy, M. B., Li, J., Lowe, E. L. & Levine, S. (1980) Maternal behavior, pup vocalizations, and pup temperature changes following handling in mice of 2 inbred strains. *Developmental Psychobiology,* **13**, 573-584.
145. Hensch, T. K., Fagiolini, M., Mataga, N., Stryker, M. P., Baekkeskov, S. & Kash, S. F. (1998) Local GABA circuit control of experience-dependent plasticity in developing visual cortex. *Science (New York, N.Y.),* **282**, 1504-1508.
146. Hentges, K. E., Sirry, B., Gingeras, A. C., Sarbassov, D., Sonenberg, N., Sabatini, D.*, et al.* (2001) FRAP/mTOR is required for proliferation and patterning during embryonic development in the mouse. *Proceedings of the National Academy of Sciences,* **98**, 13796-13801.
147. Hill, A. S., Sahay, A. & Hen, R. (2015) Increasing Adult Hippocampal Neurogenesis is Sufficient to Reduce Anxiety and Depression-Like Behaviors. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology,* **40**, 2368-2378.
148. Hoeffer, C. A. & Klann, E. (2010) mTOR signaling: At the crossroads of plasticity, memory and disease, pp. 67-75. Elsevier Current Trends.
149. Hölscher, C. (1999) Stress impairs performance in spatial water maze learning tasks. *Behavioural Brain Research,* **100**, 225-235.
150. Holtzman, D. M., Santucci, D., Kilbridge, J., Chua-Couzens, J., Fontana, D. J., Daniels, S. E.*, et al.* (1996) Developmental abnormalities and age-related neurodegeneration in a mouse model of Down syndrome. *Proceedings of the National Academy of Sciences of the United States of America,* **93**, 13333-13338.
151. Homanics, G. E., Quinlan, J. J. & Firestone, L. L. (1999) Pharmacologic and behavioral responses of inbred C57BL/6J and strain 129/SvJ mouse lines. *Pharmacology Biochemistry and Behavior,* **63**, 21-26.
152. Honeycutt, A. A., Grosse, S. D., Dunlap, L. J., Schendel, D. E., Chen, H., Brann, E.*, et al.* ECONOMIC COSTS OF MENTAL RETARDATION, CEREBRAL PALSY, HEARING LOSS, AND VISION IMPAIRMENT, pp. 207-228.
153. Hong, H.-K., Chakravarti, A. & Takahashi, J. S. (2004) The gene for soluble N-ethylmaleimide sensitive factor attachment protein alpha is mutated in hydrocephaly with hop gait (hyh) mice. *Proceedings of the National Academy of Sciences of the United States of America,* **101**, 1748-1753.
154. Horinouchi, K., Erlich, S., Perl, D. P., Ferlinz, K., Bisgaier, C. L., Sandhoff, K.*, et al.* (1995) Acid sphingomyelinase deficient mice: a model of types A and B Niemann–Pick disease. *Nature Genetics,* **10**, 288-293.
155. Horner, A. E., Heath, C. J., Hvoslef-Eide, M., Kent, B. A., Kim, C. H., Nilsson, S. R. O.*, et al.* (2013) The touchscreen operant platform for testing learning and memory in rats and mice. *Nature Protocols,* **8**, 1961-1984.

156. Hoyer, J., Ekici, A. B., Endele, S., Popp, B., Zweier, C., Wiesener, A.*, et al.* (2012) Haploinsufficiency of ARID1B, a member of the SWI/SNF-A chromatin-remodeling complex, is a frequent cause of intellectual disability. *American Journal of Human Genetics,* **90**, 565-572.
157. Hu, W. F., Chahrour, M. H. & Walsh, C. A. (2014) The Diverse Genetic Landscape of Neurodevelopmental Disorders. *Annual Review of Genomics and Human Genetics,* **15**, 195-213.
158. Hubel, D. H. & Wiesel, T. N. (1970) The period of susceptibility to the physiological effects of unilateral eye closure in kittens. *The Journal of Physiology,* **206**, 419-436.
159. Huckins, L. M., Logan, D. W. & Sánchez-Andrade, G. (2013) Olfaction and olfactory-mediated behaviour in psychiatric disease models. *Cell and Tissue Research,* **354**, 69-80.
160. Huether, R., Dong, L., Chen, X., Wu, G., Parker, M., Wei, L.*, et al.* (2014) The landscape of somatic mutations in epigenetic regulators across 1,000 paediatric cancer genomes. *Nature communications,* **5**, 3630-3630.
161. Hughes, T. R., Marton, M. J., Jones, A. R., Roberts, C. J., Stoughton, R., Armour, C. D.*, et al.* (2000) Functional discovery via a compendium of expression profiles. *Cell,* **102**, 109-126.
162. Hussain, R. & Bittles, A. H. (1998) The prevalence and demographic characteristics of consanguineous marriages in Pakistan. *Journal of biosocial science,* **30**, 261-275.
163. Huttenlocher, P. R. (1979) Synaptic Density In Human Frontal Cortex - Developmental Changes And Effects Of Aging. *Brain Research,* **163**, 195-205.
164. Iafrate, A. J., Feuk, L., Rivera, M. N., Listewnik, M. L., Donahoe, P. K., Qi, Y.*, et al.* (2004) Detection of large-scale variation in the human genome. *Nature Genetics,* **36**, 949-951.
165. Inoki, K., Zhu, T. & Guan, K.-L. (2003) TSC2 Mediates Cellular Energy Response to Control Cell Growth and Survival. *Cell,* **115**, 577-590.
166. Ishikawa, R., Fukushima, H., Frankland, P. W. & Kida, S. (2016) Hippocampal neurogenesis enhancers promote forgetting of remote fear memory after hippocampal reactivation by retrieval. *eLife,* **5**, e17464-e17464.
167. Itsara, A., Cooper, G. M., Baker, C., Girirajan, S., Li, J., Absher, D.*, et al.* (2009) Population Analysis of Large Copy Number Variants and Hotspots of Human Genetic Disease. *The American Journal of Human Genetics,* **84**, 148-161.
168. Izquierdo, A., Darling, C., Manos, N., Pozos, H., Kim, C., Ostrander, S.*, et al.* (2013) Basolateral amygdala lesions facilitate reward choices after negative feedback in rats. *The Journal of neuroscience : the official journal of the Society for Neuroscience,* **33**, 4105-4109.
169. Fray, P.J., Robbins, T, W. & Sahakian, B, J. (1996) Neuorpsychiatyric applications of CANTAB. *International Journal of Geriatric Psychiatry,* **11**, 329-336.
170. Jacinto, E., Loewith, R., Schmidt, A., Lin, S., Rüegg, M. A., Hall, A.*, et al.* (2004) Mammalian TOR complex 2 controls the actin cytoskeleton and is rapamycin insensitive. *Nature Cell Biology,* **6**, 1122-1128.
171. Jeste, S. S. (2015) Neurodevelopmental Behavioral and Cognitive Disorders. *CONTINUUM: Lifelong Learning in Neurology,* **21**, 690-714.
172. Johannessen, C. M., Johnson, B. W., Williams, S. M. G., Chan, A. W., Reczek, E. E., Lynch, R. C.*, et al.* (2008) TORC1 is essential for NF1-associated malignancies. *Current biology : CB,* **18**, 56-62.
173. Johannessen, C. M., Reczek, E. E., James, M. F., Brems, H., Legius, E. & Cichowski, K. (2005) The NF1 tumor suppressor critically regulates TSC2 and mTOR. *Proceedings of the National Academy of Sciences of the United States of America,* **102**, 8573-8578.
174. Johnson, S. C., Yanos, M. E., Kayser, E.-B., Quintana, A., Sangesland, M., Castanza, A.*, et al.* (2013) mTOR inhibition alleviates mitochondrial disease in a mouse model of Leigh syndrome. *Science (New York, N.Y.),* **342**, 1524-1528.
175. Jovanovic, T., Keyes, M., Fiallos, A., Myers, K. M., Davis, M. & Duncan, E. J. (2005) Fear potentiation and fear inhibition in a human fear-potentiated startle paradigm. *Biological Psychiatry,* **57**, 1559-1564.
176. Kadam, P. & Bhalerao, S. (2010) Sample size calculation. *International journal of Ayurveda research,* **1**, 55-57.
177. Kaufman, L., Ayub, M. & Vincent, J. B. (2010) The genetic basis of non-syndromic intellectual disability: a review. *Journal of Neurodevelopmental Disorders,* **2**, 182-209.
178. Keane, T. M., Goodstadt, L., Danecek, P., White, M. A., Wong, K., Yalcin, B.*, et al.* (2011) Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature,* **477**, 289-294.
179. Keeney, S. E., Adcock, E. W., McArdle, C. B., Mercuri, E., Cowan, F. M., Dubowitz, L. M. S.*, et al.* (1991) Prospective observations of 100 high-risk neonates by high-field (1.5 Tesla) magnetic resonance imaging of the central nervous system. II. Lesions associated with hypoxic-ischemic encephalopathy. *Pediatrics,* **87**, 431-438.

180. Kelavkar, U. P., Kelavkar & P, U. (2001) Human Genome Project. John Wiley & Sons, Ltd, Chichester.
181. Kennedy, M. B. (1997) The postsynaptic density at glutamatergic synapses, pp. 264-268.
182. Keshavan, M. S., Diwadkar, V. A., DeBellis, M., Dick, E., Kotwal, R., Rosenberg, D. R.*, et al.* (2002) Development of the corpus callosum in childhood, adolescence and early adulthood. *Life Sciences,* **70**, 1909-1922.
183. Kessels, M. M., Schwintzer, L., Schlobinski, D. & Qualmann, B. (2011) Controlling actin cytoskeletal organization and dynamics during neuronal morphogenesis. *European Journal of Cell Biology,* **90**, 926-933.
184. Kim, D.-H., Sarbassov, D. D., Ali, S. M., King, J. E., Latek, R. R., Erdjument-Bromage, H.*, et al.* (2002) mTOR interacts with raptor to form a nutrient-sensitive complex that signals to the cell growth machinery. *Cell,* **110**, 163-175.
185. Kim, E., Goraksha-Hicks, P., Li, L., Neufeld, T. P. & Guan, K.-L. (2008) Regulation of TORC1 by Rag GTPases in nutrient response. *Nature Cell Biology,* **10**, 935-945.
186. Kim, J. M., Kim, D. H., Lee, Y., Park, S. J. & Ryu, J. H. (2014) Distinct roles of the hippocampus and perirhinal cortex in GABAA receptor blockade-induced enhancement of object recognition memory. *Brain Research,* **1552**, 17-25.
187. Kim, K.-Y., Hysolli, E. & Park, I.-H. (2011) Neuronal maturation defect in induced pluripotent stem cells from patients with Rett syndrome. *Proceedings of the National Academy of Sciences of the United States of America,* **108**, 14169-14174.
188. Kitamura, T. & Inokuchi, K. (2014) Role of adult neurogenesis in hippocampal-cortical memory consolidation. *Molecular Brain,* **7**, 13-13.
189. Kitamura, T., Saitoh, Y., Takashima, N., Murayama, A., Niibori, Y., Ageta, H.*, et al.* (2009) Adult Neurogenesis Modulates the Hippocampus-Dependent Period of Associative Fear Memory. *Cell,* **139**, 814-827.
190. Kleefstra, T., Kramer, J. M., Neveling, K., Willemsen, M. H., Koemans, T. S., Vissers, L. E. L. M.*, et al.* (2012) Disruption of an EHMT1-associated chromatin-modification module causes intellectual disability. *American Journal of Human Genetics,* **91**, 73-82.
191. Kleefstra, T., Schenck, A., Kramer, J. M. & Van Bokhoven, H. (2014) The genetics of cognitive epigenetics, pp. 83-94.
192. Klintsova, A., Hamilton, G. & Boschen, K. (2012) Long-Term Consequences of Developmental Alcohol Exposure on Brain Structure and Function: Therapeutic Benefits of Physical Activity. *Brain Sciences,* **3**, 1-38.
193. Kneussel, M. & Betz, H. (2000) Clustering of inhibitory neurotransmitter receptors at developing postsynaptic sites: The membrane activation model, pp. 429-435. Elsevier Current Trends.
194. Kogan, J. H., Frankland, P. W. & Silva, A. J. Long - Term Memory Underlying Hippocampus - Dependent Social Recognition in Mice.
195. Koopmans, G., Blokland, A., van Nieuwenhuijzen, P. & Prickaerts, J. (2003) Assessment of spatial learning abilities of mice in a new circular maze. *Physiology & Behavior,* **79**, 683-693.
196. Koopmans, G., Blokland, A., Van Nieuwenhuijzen, P. & Prickaerts, J. (2003) Assessment of spatial learning abilities of mice in a new circular maze. *Physiology and Behavior,* **79**, 683-693.
197. Krackow, S., Vannoni, E., Codita, A., Mohammed, A. H., Cirulli, F., Branchi, I.*, et al.* (2010) Consistent behavioral phenotype differences between inbred mouse strains in the IntelliCage. *Genes, Brain and Behavior,* **9**, 722-731.
198. Kriegstein, A. & Alvarez-Buylla, A. (2009) The Glial Nature of Embryonic and Adult Neural Stem Cells. *Annual Review of Neuroscience,* **32**, 149-184.
199. Krucker, T., Siggins, G. R. & Halpain, S. (2000) Dynamic actin filaments are required for stable long-term potentiation (LTP) in area CA1 of the hippocampus. *Proceedings of the National Academy of Sciences of the United States of America,* **97**, 6856-6861.
200. Kuechler, A., Zink, A. M., Wieland, T., Lüdecke, H.-J., Cremer, K., Salviati, L.*, et al.* (2015) Loss-of-function variants of SETD5 cause intellectual disability and the core phenotype of microdeletion 3p25.3 syndrome. *European journal of human genetics : EJHG,* **23**, 753-760.
201. Kulesskaya, N. & Voikar, V. (2014) Assessment of mouse anxiety-like behavior in the light–dark box and open-field arena: Role of equipment and procedure. *Physiology & Behavior,* **133**, 30-38.
202. Kwon, C.-H., Luikart, B. W., Powell, C. M., Zhou, J., Matheny, S. A., Zhang, W.*, et al.* (2006) Pten regulates neuronal arborization and social interaction in mice. *Neuron,* **50**, 377-388.
203. Kwon, C.-H., Zhu, X., Zhang, J. & Baker, S. J. (2003) mTor is required for hypertrophy of Pten-deficient neuronal soma in vivo. *Proceedings of the National Academy of Sciences of the United States of America,* **100**, 12923-12928.

204. Kwon, C.-H., Zhu, X., Zhang, J., Knoop, L. L., Tharp, R., Smeyne, R. J., *et al.* (2001) Pten regulates neuronal soma size: a mouse model of Lhermitte-Duclos disease. *Nature Genetics,* **29**, 404-411.
205. Lamb, J., Crawford, E. D., Peck, D., Modell, J. W., Blat, I. C., Wrobel, M. J., *et al.* (2006) The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science (New York, N.Y.),* **313**, 1929-1935.
206. Lambrechts, A., Van Troys, M. & Ampe, C. (2004) The actin cytoskeleton in normal and pathological cell motility, pp. 1890-1909. Pergamon.
207. Lamming, D. W., Ye, L., Katajisto, P., Goncalves, M. D., Saitoh, M., Stevens, D. M., *et al.* (2012) Rapamycin-Induced Insulin Resistance Is Mediated by mTORC2 Loss and Uncoupled from Longevity. *Science,* **335**, 1638-1643.
208. Lancaster, M. A. & Knoblich, J. A. (2014) Organogenesis in a dish: modeling development and disease using organoid technologies. *Science (New York, N.Y.),* **345**, 1247125-1247125.
209. Laplante, M. & Sabatini, David M. (2012) mTOR Signaling in Growth Control and Disease. *Cell,* **149**, 274-293.
210. Larkin, A. E., Fahey, B., Gobbo, O., Callaghan, C. K., Cahill, E., O'Mara, S. M., *et al.* (2008) Blockade of NMDA receptors pre-training, but not post-training, impairs object displacement learning in the rat. *Brain Research,* **1199**, 126-132.
211. Larson, S. A., Lakin, K. C. & Anderson, L. L. (2003) Definitions and findings on intellectual. *Research in Social Science and Disability*. **3**, 229-255.
212. Laughlin, M. R., Lloyd, K. C. K., Cline, G. W., Wasserman, D. H. & Mouse Metabolic Phenotyping Centers, C. (2012) NIH Mouse Metabolic Phenotyping Centers: the power of centralized phenotyping. *Mammalian Genome,* **23**, 623-631.
213. Lebel, C. & Beaulieu, C. (2011) Longitudinal development of human brain wiring continues from childhood into adulthood. *The Journal of neuroscience : the official journal of the Society for Neuroscience,* **31**, 10937-10947.
214. Lee, J.-H. & Skalnik, D. G. (2008) Wdr82 is a C-terminal domain-binding protein that recruits the Setd1A Histone H3-Lys4 methyltransferase complex to transcription start sites of transcribed human genes. *Molecular and cellular biology,* **28**, 609-618.
215. Leger, M., Quiedeville, A., Bouet, V., Haelewyn, B., Boulouard, M., Schumann-Bard, P., *et al.* (2013) Object recognition test in mice. *Nature Protocols,* **8**, 2531-2537.
216. Lein, E. S., Hawrylycz, M. J., Ao, N., Ayres, M., Bensinger, A., Bernard, A., *et al.* (2007) Genome-wide atlas of gene expression in the adult mouse brain. *Nature,* **445**, 168-176.
217. Lejeune, J., Gautier, M. & Turpin, R. (1959) Etude des chromosomes somatiques de neuf enfants mongoliens. *Comptes rendus des seances de l'Academie des Sciences,* **248**, 1721-1722.
218. Lenneberg, E. H. (1971) Of language knowledge, apes, and brains. *Journal of psycholinguistic research,* **1**, 1-29.
219. Levitt, P., Eagleson, K. L. & Powell, E. M. (2004) Regulation of neocortical interneuron development and the implications for neurodevelopmental disorders. *Trends in Neurosciences,* **27**, 400-406.
220. Lidow, M. S., Goldman-Rakic, P. S., Gallager, D. W. & Rakic, P. (1991) Distribution of dopaminergic receptors in the primate cerebral cortex: Quantitative autoradiographic analysis using [3H]raclopride, [3H]spiperone and [3H]SCH23390. *Neuroscience,* **40**, 657-671.
221. Lipton, J. O. & Sahin, M. (2014) The neurology of mTOR. *Neuron,* **84**, 275-291.
222. Lledo, P.-M., Alonso, M. & Grubb, M. S. (2006) Adult neurogenesis and functional plasticity in neuronal circuits. *Nature Reviews Neuroscience,* **7**, 179-193.
223. Lozano, R., Martinez-Cerdeno, V. & Hagerman, R. J. (2015) Advances in the Understanding of the Gabaergic Neurobiology of FMR1 Expanded Alleles Leading to Targeted Treatments for Fragile X Spectrum Disorder. *Current pharmaceutical design,* **21**, 4972-4979.
224. Lubs, H. A. & Ruddle, F. H. (1970) Chromosomal abnormalities in the human population: estimation of rates based on New Haven newborn study. *Science (New York, N.Y.),* **169**, 495-497.
225. Lubs, H. A., Stevenson, R. E. & Schwartz, C. E. (2012) Fragile X and X-linked intellectual disability: four decades of discovery. *American journal of human genetics,* **90**, 579-590.
226. Luciana, M. & Collins, P. F. (1997) Dopaminergic Modulation of Working Memory for Spatial but Not Object Cues in Normal Humans. *Journal of Cognitive Neuroscience,* **9**, 330-347.
227. Luine, V., Martinez, C., Villegas, M., Magari??os, A. M. a. & McEwen, B. S. (1996) Restraint stress reversibly enhances spatial memory performance. *Physiology and Behavior,* **59**, 27-32.

228. Lynch, M. (2010) Rate, molecular spectrum, and consequences of human mutation. *Proceedings of the National Academy of Sciences of the United States of America,* **107**, 961-968.
229. Mar, A. C., Horner, A. E., Nilsson, S. R. O., Alsiö, J., Kent, B. A., Kim, C. H.*, et al.* (2013) The touchscreen operant platform for assessing executive function in rats and mice. *Nature Protocols,* **8**, 1985-2005.
230. Marchetto, M. C. N., Muotri, A. R., Mu, Y., Smith, A. M., Cezar, G. G. & Gage, F. H. (2008) Non-Cell-Autonomous Effect of Human SOD1G37R Astrocytes on Motor Neurons Derived from Human Embryonic Stem Cells. *Cell Stem Cell,* **3**, 649-657.
231. Marston, H. M., Spratt, C. & Kelly, J. S. (2001) Phenotyping complex behaviours: assessment of circadian control and 5-choice serial reaction learning in the mouse. *Behavioural Brain Research,* **125**, 189-193.
232. Martin, C. & Zhang, Y. (2005) The diverse functions of histone lysine methylation. *Nature Reviews Molecular Cell Biology,* **6**, 838-849.
233. Masselink, H. & Bernards, R. (2000) The adenovirus E1A binding protein BS69 is a corepressor of transcription through recruitment of N-CoR. *Oncogene,* **19**, 1538-1546.
234. Maulik, P. K., Mascarenhas, M. N., Mathers, C. D., Dua, T. & Saxena, S. (2011) Prevalence of intellectual disability: A meta-analysis of population-based studies, pp. 419-436. Pergamon.
235. Maximov, A. V., Vedernikova, E. A., Hinssen, H., Khaitlina, S. Y. & Negulyaev, Y. A. (1997) Ca-dependent regulation of Na ⁺ -selective channels via actin cytoskeleton modification in leukemia cells. *FEBS Letters,* **412**, 94-96.
236. Mayer, C. & Grummt, I. (2006) Ribosome biogenesis and cell growth: mTOR coordinates transcription by all three classes of nuclear RNA polymerases. *Oncogene,* **25**, 6384-6391.
237. McCarthy, S. E., Gillis, J., Kramer, M., Lihm, J., Yoon, S., Berstein, Y.*, et al.* (2014) De novo mutations in schizophrenia implicate chromatin remodeling and support a genetic overlap with autism and intellectual disability. *Molecular psychiatry,* **19**, 652-658.
238. McCracken, K. W., Catá, E. M., Crawford, C. M., Sinagoga, K. L., Schumacher, M., Rockich, B. E.*, et al.* (2014) Modelling human development and disease in pluripotent stem-cell-derived gastric organoids. *Nature,* **516**, 400-404.
239. McGonigle, P. & Ruggeri, B. (2014) Animal models of human disease: Challenges in enabling translation. *Biochemical Pharmacology,* **87**, 162-171.
240. McGraw, C. M., Samaco, R. C. & Zoghbi, H. Y. (2011) Adult neural function requires MeCP2. *Science,* **333**, 186-186.
241. McIlwain, K. L., Merriweather, M. Y., Yuva-Paylor, L. A. & Paylor, R. (2001) The use of behavioral test batteries: effects of training history. Physiology & Behavior. **73**. 705-717
242. McRae, J. F., Clayton, S., Fitzgerald, T. W., Kaplanis, J., Prigmore, E., Rajan, D.*, et al.* (2017) Prevalence and architecture of de novo mutations in developmental disorders. *Nature,* **542**, 433-438.
243. Mechan, A. O., Wyss, A., Rieger, H. & Mohajeri, M. H. (2009) A comparison of learning and memory characteristics of young and middle-aged wild-type mice in the IntelliCage. *Journal of Neuroscience Methods,* **180**, 43-51.
244. Meehan, T. F., Conte, N., West, D. B., Jacobsen, J. O., Mason, J., Warren, J.*, et al.* (2017) Disease model discovery from 3,328 gene knockouts by The International Mouse Phenotyping Consortium. *Nature Genetics,* **49**, 1231-1238.
245. Mefford, H. C., Batshaw, M. L. & Hoffman, E. P. (2012) Genomics, Intellectual Disability, and Autism. *New England Journal of Medicine,* **366**, 733-743.
246. Mei, Y., Monteiro, P., Zhou, Y., Kim, J.-A., Gao, X., Fu, Z.*, et al.* (2016) Adult restoration of Shank3 expression rescues selective autistic-like phenotypes. *Nature,* **530**, 481-484.
247. Mellert, H. S. & McMahon, S. B. (2009) Biochemical pathways that regulate acetyltransferase and deacetylase activity in mammalian cells, pp. 571-578. Elsevier Current Trends.
248. Michelson, D. J., Shevell, M. I., Sherr, E. H., Moeschler, J. B., Gropman, A. L. & Ashwal, S. (2011) Evidence report: Genetic and metabolic testing on children with global developmental delay: report of the Quality Standards Subcommittee of the American Academy of Neurology and the Practice Committee of the Child Neurology Society. *Neurology,* **77**, 1629-1635.
249. Mikhaleva, A., Kannan, M., Wagner, C. & Yalcin, B. (2016) Histomorphological Phenotyping of the Adult Mouse Brain. *Current protocols in mouse biology,* **6**, 307-332.

250. Miller, D. T., Adam, M. P., Aradhya, S., Biesecker, L. G., Brothman, A. R., Carter, N. P., *et al.* (2010) Consensus Statement: Chromosomal Microarray Is a First-Tier Clinical Diagnostic Test for Individuals with Developmental Disabilities or Congenital Anomalies. *American Journal of Human Genetics,* **86**, 749-764.
251. Miller, E. K. (2000) The prefrontal cortex and cognitive control. *Nature Reviews Neuroscience,* **1**, 59-65.
252. Mircsof, D., Langouët, M., Rio, M., Moutton, S., Siquier-Pernet, K., Bole-Feysot, C., *et al.* (2015) Mutations in NONO lead to syndromic intellectual disability and inhibitory synaptic defects. *Nature neuroscience,* **18**, 1731-1736.
253. Miyakawa, T., Yared, E., Pak, J. H., Huang, F. L., Huang, K.-P. & Crawley, J. N. (2001) Neurogranin null mutant mice display performance deficits on spatial learning tasks with anxiety related components. *Hippocampus,* **11**, 763-775.
254. Modabbernia, A., Mollon, J., Boffetta, P. & Reichenberg, A. (2016) Impaired Gas Exchange at Birth and Risk of Intellectual Disability and Autism: A Meta-analysis. *Journal of Autism and Developmental Disorders,* **46**, 1847-1859.
255. Moldin, S. O. & Rubenstein, J. L. R. (2006) Understanding autism: from basic neuroscience to treatment. *The new england journal of medicine*, 1690-1691.
256. Molloy, A. M., Pangilinan, F. & Brody, L. C. (2017) Genetic Risk Factors for Folate-Responsive Neural Tube Defects. *Annual Review of Nutrition,* **37**, 269-291.
257. Monroe, G. R., Frederix, G. W., Savelberg, S. M. C., de Vries, T. I., Duran, K. J., van der Smagt, J. J., *et al.* (2016) Effectiveness of whole-exome sequencing and costs of the traditional diagnostic trajectory in children with intellectual disability. *Genetics in Medicine,* **18**, 949-956.
258. Morton, A. J., Skillings, E., Bussey, T. J. & Saksida, L. M. (2006) Measuring cognitive deficits in disabled mice using an automated interactive touchscreen system. *Nature Methods,* **3**, 767-767.
259. Moy, S. S., Nadler, J. J., Perez, A., Barbaro, R. P., Johns, J. M., Magnuson, T. R., *et al.* (2004) Sociability and preference for social novelty in five inbred strains: an approach to assess autistic-like behavior in mice. *Genes, Brain and Behavior,* **3**, 287-302.
260. Moy, S. S., Nadler, J. J., Young, N. B., Nonneman, R. J., Grossman, A. W., Murphy, D. L., *et al.* (2009) Social approach in genetically engineered mouse lines relevant to autism. *Genes, Brain and Behavior,* **8**, 129-142.
261. Murai, T., Okuda, S., Tanaka, T. & Ohta, H. (2007) Characteristics of object location memory in mice: Behavioral and pharmacological studies. *Physiology & Behavior,* **90**, 116-124.
262. Murray, E. A., Mishkin, M., Mishkin, M. & Murray, E. A. (1998) Object recognition and location memory in monkeys with excitotoxic lesions of the amygdala and hippocampus. *The Journal of neuroscience : the official journal of the Society for Neuroscience,* **18**, 6568-6582.
263. Murray, E. A., Mishkin, M., Mishkin, M. & Murray, E. A. (1998) Object recognition and location memory in monkeys with excitotoxic lesions of the amygdala and hippocampus. *The Journal of neuroscience : the official journal of the Society for Neuroscience,* **18**, 6568-6582.
264. Musante, L. & Ropers, H. H. (2014) Genetics of recessive cognitive disorders, pp. 32-39. Elsevier Current Trends.
265. Myers, K. M. & Davis, M. (2004) AX+, BX- Discrimination Learning in the Fear-Potentiated Startle Paradigm: Possible Relevance to Inhibitory Fear Learning in Extinction. *Learning & Memory,* **11**, 464-475.
266. Nagl, S., Schaeferling, M. & Wolfbeis, O. S. (2005) Fluorescence Analysis in Microarray Technology. *Microchimica Acta,* **151**, 1-21.
267. Najmabadi, H., Hu, H., Garshasbi, M., Zemojtel, T., Abedini, S. S., Chen, W., *et al.* (2011) Deep sequencing reveals 50 novel genes for recessive cognitive disorders. *Nature,* **478**, 57-63.
268. Nan, X., Ng, H. H., Johnson, C. A., Laherty, C. D., Turner, B. M., Eisenman, R. N., *et al.* (1998) Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature,* **393**, 386-389.
269. Nemanic, S., Alvarado, M. C. & Bachevalier, J. (2004) The hippocampal/parahippocampal regions and recognition memory: insights from visual paired comparison versus object-delayed nonmatching in monkeys. *The Journal of neuroscience : the official journal of the Society for Neuroscience,* **24**, 2013-2026.
270. Nie, D., Di Nardo, A., Han, J. M., Baharanyi, H., Kramvis, I., Huynh, T., *et al.* (2010) Tsc2-Rheb signaling regulates EphA-mediated axon guidance. *Nature Neuroscience,* **13**, 163-172.
271. Nithianantharajah, J. & Grant, S. G. N. (2013) Cognitive components in mice and humans: Combining genetics and touchscreens for medical translation. *Neurobiology of Learning and Memory,* **105**, 13-19.

272. Nithianantharajah, J., McKechanie, A. G., Stewart, T. J., Johnstone, M., Blackwood, D. H., St Clair, D.*, et al.* (2015) Bridging the translational divide: identical cognitive touchscreen testing in mice and humans carrying mutations in a disease-relevant homologous gene. *Scientific reports,* **5**, 14613-14613.
273. Numis, A. L. & Sankar, R. (2016) Neurodevelopmental disorders, pp. 641-644. John Wiley & Sons, Ltd, Chichester, UK.
274. O'Leary, T. P. & Brown, R. E. (2012) The effects of apparatus design and test procedure on learning and memory performance of C57BL/6J mice on the Barnes maze. *Journal of Neuroscience Methods,* **203**, 315-324.
275. O'Leary, T. P., Savoie, V. & Brown, R. E. (2011) Learning, memory and search strategies of inbred mouse strains with different visual abilities in the Barnes maze. *Behavioural Brain Research,* **216**, 531-542.
276. Oomen, C. A., Hvoslef-Eide, M., Heath, C. J., Mar, A. C., Horner, A. E., Bussey, T. J.*, et al.* (2013) The touchscreen operant platform for testing working memory and pattern separation in rats and mice. *Nature protocols,* **8**, 2006-2021.
277. O'Roak, B. J., Vives, L., Girirajan, S., Karakoc, E., Krumm, N., Coe, B. P.*, et al.* (2012) Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature,* **485**, 246-250.
278. Osipovich, A. B., Gangula, R., Vianna, P. G. & Magnuson, M. A. (2016) Setd5 is essential for mammalian development and the co-transcriptional regulation of histone acetylation. *Development (Cambridge, England),* **143**, 4595-4607.
279. Osipovich, A. B., Gangula, R., Vianna, P. G. & Magnuson, M. A. (2016) Setd5 is essential for mammalian development and the co-transcriptional regulation of histone acetylation. *Development (Cambridge, England),* **143**, 4595-4607.
280. Otterbach, B. & Stoffel, W. (1995) Acid sphingomyelinase-deficient mice mimic the neurovisceral form of human lysosomal storage disease (Niemann-Pick disease). *Cell,* **81**, 1053-1061.
281. Pajusalu, S., Reimand, T. & Õunap, K. (2015) No Title. *American Journal of Medical Genetics Part A,* **167**, 1913-1915.
282. Parker, C. C., Gopalakrishnan, S., Carbonetto, P., Gonzales, N. M., Leung, E., Park, Y. J.*, et al.* (2016) Genome-wide association study of behavioral, physiological and gene expression traits in outbred CFW mice. *Nature Genetics,* **48**, 919-926.
283. Passtoors, W. M., Beekman, M., Deelen, J., van der Breggen, R., Maier, A. B., Guigas, B.*, et al.* (2013) Gene expression analysis of mTOR pathway: association with human longevity. *Aging Cell,* **12**, 24-31.
284. Peça, J., Feliciano, C., Ting, J. T., Wang, W., Wells, M. F., Venkatraman, T. N.*, et al.* (2011) Shank3 mutant mice display autistic-like behaviours and striatal dysfunction. *Nature,* **472**, 437-442.
285. Peltekova, I. T., Macdonald, A. & Armour, C. M. (2012) Microdeletion on 3p25 in a patient with features of 3p deletion syndrome. *American Journal of Medical Genetics Part A,* **158A**, 2583-2586.
286. Peng, M., Yin, N. & Li, M. O. (2017) SZT2 dictates GATOR control of mTORC1 signalling. *Nature,* **543**, 433-437.
287. Perou, C. M., Jeffrey, S. S., van de Rijn, M., Rees, C. A., Eisen, M. B., Ross, D. T.*, et al.* (1999) Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. *Proceedings of the National Academy of Sciences of the United States of America,* **96**, 9212-9217.
288. Peterson, T. R., Laplante, M., Thoreen, C. C., Sancak, Y., Kang, S. A., Kuehl, W. M.*, et al.* (2009) DEPTOR Is an mTOR Inhibitor Frequently Overexpressed in Multiple Myeloma Cells and Required for Their Survival. *Cell,* **137**, 873-886.
289. Phelps, E. A., Delgado, M. R., Nearing, K. I. & Ledoux, J. E. (2004) Extinction learning in humans: Role of the amygdala and vmPFC. *Neuron,* **43**, 897-905.
290. Phung, T. L., Ziv, K., Dabydeen, D., Eyiah-Mensah, G., Riveros, M., Perruzzi, C.*, et al.* (2006) Pathological angiogenesis is induced by sustained Akt signaling and inhibited by rapamycin. *Cancer Cell,* **10**, 159-170.
291. Picker, J. D., Yang, R., Ricceri, L. & Berger-Sweeney, J. (2006) An altered neonatal behavioral phenotype in Mecp2 mutant mice. *Neuroreport,* **17**, 541-544.
292. Pieretti, M., Zhang, F., Fu, Y.-H., Warren, S. T., Oostra, B. A., Caskey, C. T.*, et al.* (1991) Absence of expression of the FMR-1 gene in fragile X syndrome. *Cell,* **66**, 817-822.
293. Piven, J., Arndt, S., Bailey, J., Havercamp, S., Andreasen, N. C. & Palmer, P. (1995) An MRI study of brain size in autism. *American Journal of Psychiatry,* **152**, 1145-1149.
294. Polder, J. J., Meerding, W. J., Bonneux, L. & Maas, P. J. v. d. (2002) Healthcare costs of intellectual disability in the Netherlands: a cost-of-illness perspective. *Journal of Intellectual Disability Research,* **46**, 168-178.

295. Pollard, T. D. & Cooper, J. A. (1986) Actin and Actin-Binding Proteins. A Critical Evaluation of Mechanisms and Functions. *Annual Review of Biochemistry,* **55**, 987-1035.
296. Potocnik, U. & Widhalm, K. (1994) Long-term follow-up of children with classical phenylketonuria after diet discontinuation: a review. *Journal of the American College of Nutrition,* **13**, 232-236.
297. Prins, M. L. & Hovda, D. A. (1998) Traumatic Brain Injury in the Developing Rat: Effects of Maturation on Morris Water Maze Acquisition. *JOURNAL OF NEUROTRAUMA,* **15**.
298. Radio, N. M. & Mundy, W. R. (2008) Developmental neurotoxicity testing in vitro: Models for assessing chemical effects on neurite outgrowth, pp. 361-376. Elsevier.
299. Ramos, A., Berton, O., Mormède, P. & Chaouloff, F. (1997) A multiple-test study of anxiety-related behaviours in six inbred rat strains. *Behavioural brain research,* **85**, 57-69.
300. Rauch, A., Hoyer, J., Guth, S., Zweier, C., Kraus, C., Becker, C.*, et al.* (2006) Diagnostic yield of various genetic approaches in patients with unexplained developmental delay or mental retardation. *American Journal of Medical Genetics Part A,* **140A**, 2063-2074.
301. Rauch, A., Wieczorek, D., Graf, E., Wieland, T., Endele, S., Schwarzmayr, T.*, et al.* (2012) Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: An exome sequencing study. *The Lancet,* **380**, 1674-1682.
302. Redon, R., Ishikawa, S., Fitch, K. R., Feuk, L., Perry, G. H., Andrews, T. D.*, et al.* (2006) Global variation in copy number in the human genome. *Nature,* **444**, 444-454.
303. Reversing Neurodevelopmental Disorders in Adults. (2008), pp. 950-960. Cell Press.
304. Ricceri, L., Colozza, C. & Calamandrei, G. (2000) Ontogeny of spatial discrimination in mice: a longitudinal analysis in the modified open-field with objects. *Developmental psychobiology,* **37**, 109-118.
305. Rice, D., Barone, S. & Jr. (2000) Critical periods of vulnerability for the developing nervous system: evidence from humans and animal models. *Environmental health perspectives,* **108 Suppl**, 511-533.
306. Rivière, J.-B., van Bon, B. W. M., Hoischen, A., Kholmanskikh, S. S., O'Roak, B. J., Gilissen, C.*, et al.* (2012) De novo mutations in the actin genes ACTB and ACTG1 cause Baraitser-Winter syndrome. *Nature Genetics,* **44**, 440-444.
307. Roach, J. C., Glusman, G., Smit, A. F. A., Huff, C. D., Hubley, R., Shannon, P. T.*, et al.* (2010) Analysis of genetic inheritance in a family quartet by whole-genome sequencing. *Science (New York, N.Y.),* **328**, 636-639.
308. Robbins, T. W., James, M., Owen, A. M., Sahakian, B. J., McInnes, L. & Rabbitt, P. (2010) Cambridge Neuropsychological Test Automated Battery (CANTAB): A Factor Analytic Study of a Large Sample of Normal Elderly Volunteers. *Dementia and Geriatric Cognitive Disorders,* **5**, 266-281.
309. Rogers, D. C., Fisher, E. M. C., Brown, S. D. M., Peters, J., Hunter, A. J. & Martin, J. E. (1997) Behavioral and functional analysis of mouse phenotype: SHIRPA, a proposed protocol for comprehensive phenotype assessment. *Mammalian Genome,* **8**, 711-713.
310. Rogers, D. C., Jones, D. N. C., Nelson, P. R., Jones, C. M., Quilter, C. A., Robinson, T. L.*, et al.* (1999) Use of SHIRPA and discriminant analysis to characterise marked differences in the behavioural phenotype of six inbred mouse strains. *Behavioural Brain Research,* **105**, 207-217.
311. Rollins, J. D., Collins, J. S. & Holden, K. R. (2010) United States Head Circumference Growth Reference Charts: Birth to 21 Years. *The Journal of Pediatrics,* **156**, 907-913.e902.
312. Rolls, E. T., O'Doherty, J., Kringelbach, M. L., Hornak, J. & Andrews, C. (2001) Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience,* **4**, 95-102.
313. Ronan, J. L., Wu, W. & Crabtree, G. R. (2013) From neural development to cognition: unexpected roles for chromatin. *Nature Reviews Genetics,* **14**, 347-359.
314. Ropers, H. H. (2010) Genetics of Early Onset Cognitive Impairment. *Annual Review of Genomics and Human Genetics,* **11**, 161-187.
315. Roubertoux, P. L., Martin, B., Le Roy, I., Beau, J., Perez-Diaz, F., Cohen-Salmon, C.*, et al.* (1996) Vocalizations in newborn mice: Genetic analysis. *Behavior Genetics,* **26**, 427-437.
316. Rousset, C. I., Chalon, S., Cantagrel, S., Bodard, S., Andres, C., Gressens, P.*, et al.* (2006) Maternal Exposure to LPS Induces Hypomyelination in the Internal Capsule and Programmed Cell Death in the Deep Gray Matter in Newborn Rats. *Pediatric Research,* **59**, 428-433.
317. Rubenstein, J. L. R. (2011) Annual Research Review: Development of the cerebral cortex: implications for neurodevelopmental disorders. *Journal of Child Psychology and Psychiatry,* **52**, 339-355.

318. Rutherford, M. A., Pennock, J. M., Counsell, S. J., Mercuri, E., Cowan, F. M., Dubowitz, L. M., *et al.* (1998) Abnormal magnetic resonance signal in the internal capsule predicts poor neurodevelopmental outcome in infants with hypoxic-ischemic encephalopathy. *Pediatrics,* **102**, 323-328.
319. Sahay, A., Scobie, K. N., Hill, A. S., O'Carroll, C. M., Kheirbek, M. A., Burghardt, N. S., *et al.* (2011) Increasing adult hippocampal neurogenesis is sufficient to improve pattern separation. *Nature,* **472**, 466-470.
320. Sale, A., Maya Vetencourt, J. F., Medini, P., Cenni, M. C., Baroncelli, L., De Pasquale, R., *et al.* (2007) Environmental enrichment in adulthood promotes amblyopia recovery through a reduction of intracortical inhibition. *Nature Neuroscience,* **10**, 679-681.
321. Salvador-Carulla, L. & Bertelli, M. (2008) 'Mental retardation' or 'intellectual disability': time for a conceptual change. *Psychopathology,* **41**, 10-16.
322. Sánchez-Andrade, G., James, B. M. & Kendrick, K. M. (2005) Neural Encoding of Olfactory Recognition Memory. *Journal of Reproduction and Development,* **51**.
323. Santen, G. W. E., Aten, E., Sun, Y., Almomani, R., Gilissen, C., Nielsen, M., *et al.* (2012) Mutations in SWI/SNF chromatin remodeling complex gene ARID1B cause Coffin-Siris syndrome. *Nature Genetics,* **44**, 379-380.
324. Sarbassov, D. D., Ali, S. M., Sengupta, S., Sheen, J. H., Hsu, P. P., Bagley, A. F., *et al.* (2006) Prolonged Rapamycin Treatment Inhibits mTORC2 Assembly and Akt/PKB. *Molecular Cell,* **22**, 159-168.
325. Sawiak, S. J., Wood, N. I., Williams, G. B., Morton, A. J. & Carpenter, T. A. (2013) Voxel-based morphometry with templates and validation in a mouse model of Huntington's disease. *Magnetic Resonance Imaging,* **31**, 1522-1531.
326. Sawtell, N. B., Frenkel, M. Y., Philpot, B. D., Nakazawa, K., Tonegawa, S. & Bear, M. F. (2003) NMDA receptor-dependent ocular dominance plasticity in adult visual cortex. *Neuron,* **38**, 977-985.
327. Saxton, R. A. & Sabatini, D. M. (2017) mTOR Signaling in Growth, Metabolism, and Disease. *Cell,* **168**, 960-976.
328. Scattoni, M. L., Crawley, J. & Ricceri, L. (2009) Ultrasonic vocalizations: A tool for behavioural phenotyping of mouse models of neurodevelopmental disorders. *Neuroscience & Biobehavioral Reviews,* **33**, 508-515.
329. Schmeisser, M. J., Ey, E., Wegener, S., Bockmann, J., Stempel, A. V., Kuebler, A., *et al.* (2012) Autistic-like behaviours and hyperactivity in mice lacking ProSAP1/Shank2. *Nature*.
330. Schneider, R., Bannister, A. J. & Kouzarides, T. (2002) Unsafe SETs: histone lysine methyltransferases and cancer. *Trends in Biochemical Sciences,* **27**, 396-402.
331. Schneider, R., Bannister, A. J., Myers, F. A., Thorne, A. W., Crane-Robinson, C. & Kouzarides, T. (2004) Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. *Nature Cell Biology,* **6**, 73-77.
332. Schoenbaum, G., Chiba, A. A. & Gallagher, M. (2000) Changes in functional connectivity in orbitofrontal cortex and basolateral amygdala during learning and reversal training. *The Journal of neuroscience : the official journal of the Society for Neuroscience,* **20**, 5179-5189.
333. Schuff, N., Neylan, T. C., Lenoci, M. A., Du, A.-T., Weiss, D. S., Marmar, C. R., *et al.* (2001) Decreased hippocampal N-acetylaspartate in the absence of atrophy in posttraumatic stress disorder. *Biological Psychiatry,* **50**, 952-959.
334. Schultz, W. & Tremblay, L. (1999) Relative reward preference in primate orbitofrontal cortex. *Nature,* **398**, 704-708.
335. Schuurs-Hoeijmakers, J. H. M., Geraghty, M. T., Kamsteeg, E. J., Ben-Salem, S., De Bot, S. T., Nijhof, B., *et al.* (2012) Mutations in DDHD2, encoding an intracellular phospholipase A1, cause a recessive form of complex hereditary spastic paraplegia. *American Journal of Human Genetics,* **91**, 1073-1081.
336. Schuurs-Hoeijmakers, J. H. M., Hehir-Kwa, J. Y., Pfundt, R., van Bon, B. W. M., de Leeuw, N., Kleefstra, T., *et al.* (2011) Homozygosity mapping in outbred families with mental retardation. *European Journal of Human Genetics,* **19**, 597-601.
337. Seibenhener, M. L. & Wooten, M. C. (2015) Use of the Open Field Maze to measure locomotor and anxiety-like behavior in mice. *Journal of visualized experiments : JoVE*, e52434-e52434.
338. Semenova, S. (2012) Attention, impulsivity, and cognitive flexibility in adult male rats exposed to ethanol binge during adolescence as measured in the five-choice serial reaction time task: the effects of task and ethanol challenges. *Psychopharmacology,* **219**, 433-442.
339. Semple, B. D., Blomgren, K., Gimlin, K., Ferriero, D. M. & Noble-Haeusslein, L. J. (2013) Brain development in rodents and humans: Identifying benchmarks of maturation and vulnerability to injury across species, pp. 1-16. Pergamon.

340. Sharma, A., Hoeffer, C. A., Takayasu, Y., Miyawaki, T., McBride, S. M., Klann, E., *et al.* (2010) Dysregulation of mTOR signaling in fragile X syndrome. *The Journal of neuroscience : the official journal of the Society for Neuroscience,* **30**, 694-702.
341. Sheridan, S. D., Theriault, K. M., Reis, S. A., Zhou, F., Madison, J. M., Daheron, L., *et al.* (2011) Epigenetic Characterization of the FMR1 Gene and Aberrant Neurodevelopment in Human Induced Pluripotent Stem Cell Models of Fragile X Syndrome. *PLOS ONE,* **6**, e26203-e26203.
342. Shibutani, M., Horii, T., Shoji, H., Morita, S., Kimura, M., Terawaki, N., *et al.* (2017) Arid1b Haploinsufficiency Causes Abnormal Brain Gene Expression and Autism-Related Behaviors in Mice. *International Journal of Molecular Sciences,* **18**, 1872-1872.
343. Shin, J., Kim, M., Jung, H.-J., Cha, H. L., Suh-Kim, H., Ahn, S., *et al.* (2017) Characterization of developmental defects in the forebrain resulting from hyperactivated mTOR signaling by integrative analysis of transcriptomic and proteomic data. *Scientific reports,* **7**, 2826-2826.
344. Shogren-Knaak, M., Ishii, H., Sun, J.-M., Pazin, M. J., Davie, J. R. & Peterson, C. L. (2006) Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science (New York, N.Y.),* **311**, 844-847.
345. Shuch, B., Ricketts, C. J., Vocke, C. D., Komiya, T., Middelton, L. A., Kauffman, E. C., *et al.* (2013) Germline PTEN Mutation Cowden Syndrome: An Underappreciated Form of Hereditary Kidney Cancer. *The Journal of Urology,* **190**, 1990-1998.
346. Shum, F. W. F., Ko, S. W., Lee, Y.-S., Kaang, B.-K. & Zhuo, M. (2005) Genetic Alteration of Anxiety and Stress-Like Behavior in Mice Lacking CaMKIV. *Molecular Pain,* **1**, 1744-8069-1741-1722.
347. Silverman, J. L., Yang, M., Lord, C. & Crawley, J. N. (2010) Behavioural phenotyping assays for mouse models of autism. *Nature Reviews Neuroscience,* **11**, 490-502.
348. Sim, J. C. H., White, S. M., Fitzpatrick, E., Wilson, G. R., Gillies, G., Pope, K., *et al.* (2014) Expanding the phenotypic spectrum of ARID1B-mediated disorders and identification of altered cell-cycle dynamics due to ARID1B haploinsufficiency. *Orphanet Journal of Rare Diseases,* **9**, 43-43.
349. Singh, T., Kurki, M. I., Curtis, D., Purcell, S. M., Crooks, L., McRae, J., *et al.* (2016) Rare loss-of-function variants in SETD1A are associated with schizophrenia and developmental disorders. *Nature Neuroscience,* **19**, 571-577.
350. Skarnes, W. C., Rosen, B., West, A. P., Koutsourakis, M., Bushell, W., Iyer, V., *et al.* (2011) A conditional knockout resource for the genome-wide study of mouse gene function. *Nature,* **474**, 337-342.
351. Snijders Blok, L., Madsen, E., Juusola, J., Gilissen, C., Baralle, D., Reijnders, Margot R. F., *et al.* (2015) Mutations in DDX3X Are a Common Cause of Unexplained Intellectual Disability with Gender-Specific Effects on Wnt Signaling. *The American Journal of Human Genetics,* **97**, 343-352.
352. Sowell, E. R., Thompson, P. M., Holmes, C. J., Jernigan, T. L. & Toga, A. W. (1999) In vivo evidence for post-adolescent brain maturation in frontal and striatal regions. *Nature Neuroscience,* **2**, 859-861.
353. Spencer, R. L., O'Steen, W. K. & McEwen, B. S. (1995) Water maze performance of aged Sprague-Dawley rats in relation to retinal morphologic measures. *Behavioural Brain Research,* **68**, 139-150.
354. Spilman, P., Podlutskaya, N., Hart, M. J., Debnath, J., Gorostiza, O., Bredesen, D., *et al.* (2010) Inhibition of mTOR by Rapamycin Abolishes Cognitive Deficits and Reduces Amyloid-$β$ Levels in a Mouse Model of Alzheimer's Disease. *PLoS ONE,* **5**, e9979-e9979.
355. Spolidoro, M., Sale, A., Berardi, N. & Maffei, L. (2009) Plasticity in the adult brain: lessons from the visual system. *Experimental Brain Research,* **192**, 335-341.
356. Squire, L. R., Wixted, J. T. & Clark, R. E. (2007) Recognition memory and the medial temporal lobe: a new perspective. *Nature reviews. Neuroscience,* **8**, 872-883.
357. Stelzer, G., Rosen, N., Plaschkes, I., Zimmerman, S., Twik, M., Fishilevich, S., *et al.* (2016) The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses, pp. 1.30.31-31.30.33. John Wiley & Sons, Inc., Hoboken, NJ, USA.
358. Strekalova, T., Spanagel, R., Dolgov, O. & Bartsch, D. (2005) Stress-induced hyperlocomotion as a confounding factor in anxiety and depression models in mice. *Behavioural Pharmacology,* **16**, 171-180.
359. Stur, E., Soares, L. & Louro, I. (2017) SETD5 gene variant associated with mild intellectual disability -a case report. *Genetics and Molecular Research Mol. Res. Genetics and Molecular Research,* **16**.
360. Sweatt, J. D. (2004) Hippocampal function in cognition. *Psychopharmacology,* **174**, 99-110.
361. Szczałuba, K., Brzezinska, M., Kot, J., Rydzanicz, M., Walczak, A., Stawiński, P., *et al.* (2016) SETD5 loss-of-function mutation as a likely cause of a familial syndromic intellectual disability with variable phenotypic expression. *American Journal of Medical Genetics Part A,* **170**, 2322-2327.

362. Szulc, K. U., Lerch, J. P., Nieman, B. J., Bartelle, B. B., Friedel, M., Suero-Abreu, G. A., *et al.* (2015) 4D MEMRI atlas of neonatal FVB/N mouse brain development. *NeuroImage,* **118**, 49-62.
363. Tailby, C., Wright, L. L., Metha, A. B. & Calford, M. B. (2005) Activity-dependent maintenance and growth of dendrites in adult cortex. *Proceedings of the National Academy of Sciences of the United States of America,* **102**, 4631-4636.
364. Takahashi, K. & Yamanaka, S. (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell,* **126**, 663-676.
365. Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., *et al.* (2007) Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell,* **107**, 861-872.
366. Takata, A., Xu, B., Ionita-Laza, I., Roos, J. L., Gogos, J. A. & Karayiorgou, M. (2014) Loss-of-Function Variants in Schizophrenia Risk and SETD1A as a Candidate Susceptibility Gene. *Neuron,* **82**, 773-780.
367. Tarantino, L. M., Gould, T. J., Druhan, J. P. & Bucan, M. (2000) Behavior and mutagenesis screens: the importance of baseline analysis of inbred strains. *Mammalian Genome,* **11**, 555-564.
368. Tarpey, P. S., Smith, R., Pleasance, E., Whibley, A., Edkins, S., Hardy, C., *et al.* (2009) A systematic, large-scale resequencing screen of X-chromosome coding exons in mental retardation. *Nature Genetics,* **41**, 535-543.
369. Tartaglia, M. & Gelb, B. D. (2005) NOONAN SYNDROME AND RELATED DISORDERS: Genetics and Pathogenesis. *Annual Review of Genomics and Human Genetics,* **6**, 45-68.
370. Tavazoie, S. F., Alvarez, V. A., Ridenour, D. A., Kwiatkowski, D. J. & Sabatini, B. L. (2005) Regulation of neuronal morphology and function by the tumor suppressors Tsc1 and Tsc2. *Nature Neuroscience,* **8**, 1727-1734.
371. Teber, S., Sezer, T., Kafalı, M., Chiara Manzini, M., Konuk Yüksel, B., Tekin, M., *et al.* (2008) Severe muscle–eye–brain disease is associated with a homozygous mutation in the POMGnT1 gene. *European Journal of Paediatric Neurology,* **12**, 133-136.
372. Tee, A. R., Sampson, J. R., Pal, D. K. & Bateman, J. M. (2016) The role of mTOR signalling in neurogenesis, insights from tuberous sclerosis complex, pp. 12-20. Academic Press.
373. Ten Kate, L. P., Teeuw, M., Henneman, L. & Cornel, M. C. (2010) Autosomal recessive disease in children of consanguineous parents: inferences from the proportion of compound heterozygotes. *Journal of community genetics,* **1**, 37-40.
374. Testa, G., Schaft, J., van der Hoeven, F., Glaser, S., Anastassiadis, K., Zhang, Y., *et al.* (2004) A reliable lacZ expression reporter cassette for multipurpose, knockout-first alleles. *genesis,* **38**, 151-158.
375. The Dutch-Belgian Fragile, X. C., Bakker, C. E., Verheij, C., Willemsen, R., van der Helm, R., Oerlemans, F., *et al.* (1994) Fmr1 knockout mice: A model to study fragile X mental retardation. *Cell,* **78**, 23-33.
376. Thomanetz, V., Angliker, N., Cloëtta, D., Lustenberger, R. M., Schweighauser, M., Oliveri, F., *et al.* (2013) Ablation of the mTORC2 component rictor in brain or Purkinje cells affects size and neuron morphology. *The Journal of Cell Biology,* **201**, 293-308.
377. Thomas, M. J., Watabe, A. M., Moody, T. D., Makhinson, M., O'Dell, T. J., Wenthold, R. J., *et al.* (1998) Postsynaptic complex spike bursting enables the induction of LTP by theta frequency synaptic stimulation. *The Journal of neuroscience : the official journal of the Society for Neuroscience,* **18**, 7118-7126.
378. Thomson, J. A., Itskovitz-Eldor, J., Shapiro, S. S., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S., *et al.* (1998) Embryonic stem cell lines derived from human blastocysts. *Science (New York, N.Y.),* **282**, 1145-1147.
379. Thornton, L. M., Hahn, M. E. & Schanz, N. (2005) Genetic and Developmental Influences on Infant Mouse Ultrasonic Calling. III. Patterns of Inheritance in the Calls of Mice 3?9 Days of Age. *Behavior Genetics,* **35**, 73-83.
380. Trachtenberg, J. T., Chen, B. E., Knott, G. W., Feng, G., Sanes, J. R., Welker, E., *et al.* (2002) Long-term in vivo imaging of experience-dependent synaptic plasticity in adult cortex. *Nature,* **420**, 788-794.
381. Tsuchida, N., Nakashima, M., Miyauchi, A., Yoshitomi, S., Kimizu, T., Ganesan, V., *et al.* (2017) Novel biallelic SZT2 mutations in 3 cases of early-onset epileptic encephalopathy. *Clinical Genetics*.
382. Tsurusaki, Y., Okamoto, N., Ohashi, H., Kosho, T., Imai, Y., Hibi-Ko, Y., *et al.* (2012) Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. *Nature Genetics,* **44**, 376-378.
383. Tuscher, J. J., Fortress, A. M., Kim, J. & Frick, K. M. (2015) Regulation of object recognition and object placement by ovarian sex steroid hormones, pp. 140-157. Elsevier.
384. Uher, R. (2009) The role of genetic variation in the causation of mental illness: an evolution-informed framework. *Molecular Psychiatry,* **14**, 1072-1082.
385. Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., *et al.* (2015) Proteomics. Tissue-based map of the human proteome. *Science (New York, N.Y.),* **347**, 1260419-1260419.

386. Voikar, V., K??ks, S., Vasar, E. & Rauvala, H. (2001) Strain and gender differences in the behavior of mouse lines commonly used in transgenic studies. *Physiology and Behavior,* **72**, 271-281.
387. Vals, M.-A., Õiglane-Shlik, E., Nõukas, M., Shor, R., Peet, A., Kals, M.*, et al.* (2014) Coffin-Siris Syndrome with obesity, macrocephaly, hepatomegaly and hyperinsulinism caused by a mutation in the ARID1B gene. *European journal of human genetics : EJHG,* **22**, 1327-1329.
388. van Bokhoven, H. (2011) Genetic and Epigenetic Networks in Intellectual Disabilities. *Annual Review of Genetics,* **45**, 81-104.
389. van der Weyden, L., White, J. K., Adams, D. J. & Logan, D. W. (2011) The mouse genetics toolkit: revealing function and mechanism. *Genome Biology,* **12**, 224-224.
390. Van Naarden Braun, K., Christensen, D., Doernberg, N., Schieve, L., Rice, C., Wiggins, L.*, et al.* (2015) Trends in the Prevalence of Autism Spectrum Disorder, Cerebral Palsy, Hearing Loss, Intellectual Disability, and Vision Impairment, Metropolitan Atlanta, 1991–2010. *PLOS ONE,* **10**, e0124120-e0124120.
391. Vann, S. D. & Albasser, M. M. (2011) Hippocampus and neocortex: recognition and spatial memory. *Current Opinion in Neurobiology,* **21**, 440-445.
392. Velasco, G., Grkovic, S. & Ansieau, S. (2006) New insights into BS69 functions. *The Journal of biological chemistry,* **281**, 16546-16550.
393. Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G.*, et al.* (2001) The sequence of the human genome. *Science (New York, N.Y.),* **291**, 1304-1351.
394. Verkerk, A. J. M. H., Pieretti, M., Sutcliffe, J. S., Fu, Y. H., Kuhl, D. P. A., Pizzuti, A.*, et al.* (1991) Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell,* **65**, 905-914.
395. Vissers, L. E. L. M., de Ligt, J., Gilissen, C., Janssen, I., Steehouwer, M., de Vries, P.*, et al.* (2010) A de novo paradigm for mental retardation. *Nature Genetics,* **42**, 1109-1112.
396. Vissers, L. E. L. M., Gilissen, C. & Veltman, J. A. (2015) Genetic studies in intellectual disability and related disorders. *Nature Reviews Genetics,* **17**, 9-18.
397. Vogel-Ciernia, A. & Wood, M. A. (2014) Neuron-specific chromatin remodeling: A missing link in epigenetic mechanisms underlying synaptic plasticity, memory, and intellectual disability disorders. *Neuropharmacology,* **80**, 18-27.
398. Voikar, V., Rossi, J., Rauvala, H. & Airaksinen, M. S. (2004) Impaired behavioural flexibility and memory in mice lacking GDNF family receptor alpha2. *European Journal of Neuroscience,* **20**, 308-312.
399. Voikar, V., Vasar, E. & Rauvala, H. (2004) Behavioral alterations induced by repeated testing in C57BL/6J and 129S2/Sv mice: implications for phenotyping screens. *Genes, Brain and Behavior,* **3**, 27-38.
400. Voikar, V., Vasar, E. & Rauvala, H. (2004) Behavioral alterations induced by repeated testing in C57BL/6J and 129S2/Sv mice: implications for phenotyping screens. *Genes, Brain and Behavior,* **3**, 27-38.
401. Völkel, P. & Angrand, P.-O. (2007) The control of histone lysine methylation in epigenetic regulation. *Biochimie,* **89**, 1-20.
402. Wagenstaller, J., Spranger, S., Lorenz-Depiereux, B., Kazmierczak, B., Nathrath, M., Wahl, D.*, et al.* (2007) Copy-Number Variations Measured by Single-Nucleotide–Polymorphism Oligonucleotide Arrays in Patients with Mental Retardation. *The American Journal of Human Genetics,* **81**, 768-779.
403. Wakana, S., Suzuki, T., Furuse, T., Kobayashi, K., Miura, I., Kaneda, H.*, et al.* (2009) Introduction to the Japan Mouse Clinic at the RIKEN BioResource Center. *Experimental animals,* **58**, 443-450.
404. Walsh, C. A. & Engle, E. C. (2010) Allelic Diversity in Human Developmental Neurogenetics: Insights into Biology and Disease. *Neuron,* **68**, 245-253.
405. Wang, B. T., Ducker, G. S., Barczak, A. J., Barbeau, R., Erle, D. J. & Shokat, K. M. (2011) The mammalian target of rapamycin regulates cholesterol biosynthetic gene expression and exhibits a rapamycin-resistant transcriptional profile. *Proceedings of the National Academy of Sciences of the United States of America,* **108**, 15201-15206.
406. Wang, L., Harris, T. E., Roth, R. A. & Lawrence, J. C. (2007) PRAS40 regulates mTORC1 kinase activity by functioning as a direct inhibitor of substrate binding. *The Journal of biological chemistry,* **282**, 20036-20044.
407. Watase, K. & Zoghbi, H. Y. (2003) Modelling brain diseases in mice: the challenges of design and analysis. *Nature Reviews Genetics,* **4**, 296-307.
408. Way, S. W., Rozas, N. S., Wu, H. C., McKenna, J., Reith, R. M., Hashmi, S. S.*, et al.* (2012) The differential effects of prenatal and/or postnatal rapamycin on neurodevelopmental defects and cognition in a neuroglial mouse model of tuberous sclerosis complex. *Human Molecular Genetics,* **21**, 3226-3236.

409. Weitzdoerfer, R., Fountoulakis, M. & Lubec, G. (2002) Reduction of actin-related protein complex 2/3 in fetal Down syndrome brain. *Biochemical and Biophysical Research Communications,* **293**, 836-841.
410. Wen, H., Li, Y., Li, H. & Shi, X. (2014) ZMYND11. *Cell Cycle,* **13**, 2153-2154.
411. Wen, H., Li, Y., Xi, Y., Jiang, S., Stratton, S., Peng, D.*, et al.* (2014) ZMYND11 links histone H3.3K36me3 to transcription elongation and tumour suppression. *Nature,* **508**, 263-268.
412. White, J. K., Gerdin, A.-K., Karp, N. A., Ryder, E., Buljan, M., Bussell, J. N.*, et al.* (2013) Genome-wide generation and systematic phenotyping of knockout mice reveals new roles for many genes. *Cell,* **154**, 452-464.
413. White, P., Chant, D., Edwards, N., Townsend, C. & Waghorn, G. (2005) Prevalence of intellectual disability and comorbid mental illness in an Australian community sample. *Australian and New Zealand Journal of Psychiatry,* **39**, 395-400.
414. Wilde, J. J., Petersen, J. R. & Niswander, L. (2014) Genetic, Epigenetic, and Environmental Contributions to Neural Tube Closure. *Annual Review of Genetics,* **48**, 583-611.
415. Williams, C. A., Dagli, A. & Battaglia, A. (2008) Genetic disorders associated with macrocephaly. *American Journal of Medical Genetics Part A,* **146A**, 2023-2037.
416. Williams, C. A., Dagli, A. & Battaglia, A. (2008) Genetic disorders associated with macrocephaly. *American Journal of Medical Genetics Part A,* **146A**, 2023-2037.
417. Williams, R. W. & Herrup, K. (1988) The Control of Neuron Number. *Annual Review of Neuroscience,* **11**, 423-453.
418. Winder, S. J. & Ayscough, K. R. (2005) Actin-binding proteins. *Journal of cell science,* **118**, 651-654.
419. Winslow, J. T. & Insel, T. R. (2004) Neuroendocrine basis of social recognition, pp. 248-253. Elsevier Current Trends.
420. Wolfer, D. P., Stagljar-Bozicevic, M., Errington, M. L. & Lipp, H.-P. (1998) Spatial Memory and Learning in Transgenic Mice: Fact or Artifact? *News in physiological sciences : an international journal of physiology produced jointly by the International Union of Physiological Sciences and the American Physiological Society,* **13**, 118-123.
421. Wolfson, R. L., Chantranupong, L., Wyant, G. A., Gu, X., Orozco, J. M., Shen, K.*, et al.* (2017) KICSTOR recruits GATOR1 to the lysosome and is necessary for nutrients to regulate mTORC1. *Nature,* **543**, 438-442.
422. Worthey, E. A., Mayer, A. N., Syverson, G. D., Helbling, D., Bonacci, B. B., Decker, B.*, et al.* (2011) Making a definitive diagnosis: Successful clinical application of whole exome sequencing in a child with intractable inflammatory bowel disease. *Genetics in Medicine,* **13**, 255-262.
423. Wright, C. F., Fitzgerald, T. W., Jones, W. D., Clayton, S., McRae, J. F., Van Kogelenberg, M.*, et al.* (2015) Genetic diagnosis of developmental disorders in the DDD study: A scalable analysis of genome-wide research data. *The Lancet,* **385**, 1305-1314.
424. Yan, Z., Wang, Z., Sharova, L., Sharov, A. A., Ling, C., Piao, Y.*, et al.* (2008) BAF250B-associated SWI/SNF chromatin-remodeling complex is required to maintain undifferentiated mouse embryonic stem cells. *Stem Cells (Dayton, Ohio),* **26**, 1155-1165.
425. Yang, M., Lewis, F. C., Sarvi, M. S., Foley, G. M. & Crawley, J. N. (2015) 16p11.2 Deletion mice display cognitive deficits in touchscreen learning and novelty recognition tasks. *Learning & memory (Cold Spring Harbor, N.Y.),* **22**, 622-632.
426. Yang, Y. J., Baltus, A. E., Mathew, R. S., Murphy, E. A., Evrony, G. D., Gonzalez, D. M.*, et al.* (2012) Microcephaly gene links trithorax and REST/NRSF to control neural stem cell proliferation and differentiation. *Cell,* **151**, 1097-1112.
427. Yang, Y., Muzny, D. M., Reid, J. G., Bainbridge, M. N., Willis, A., Ward, P. A.*, et al.* (2013) Clinical Whole-Exome Sequencing for the Diagnosis of Mendelian Disorders. *New England Journal of Medicine,* **369**, 1502-1511.
428. Yao, H., Han, X. & Han, X. (2014) The Cardioprotection of the Insulin-Mediated PI3K/Akt/mTOR Signaling Pathway. *American Journal of Cardiovascular Drugs,* **14**, 433-442.
429. Yi, J. H., Park, H. J., Kim, B. C., Kim, D. H. & Ryu, J. H. (2016) Evidences of the role of the rodent hippocampus in the non-spatial recognition memory. *Behavioural Brain Research,* **297**, 141-149.
430. Yokogami, K., Wakisaka, S., Avruch, J. & Reeves, S. A. (2000) Serine phosphorylation and maximal activation of STAT3 during CNTF signaling is mediated by the rapamycin target mTOR. *Current Biology,* **10**, 47-50.
431. Yu, B., Shao, Y., Zhang, C., Chen, Y., Zhong, Q., Zhang, J.*, et al.* (2009) BS69 undergoes SUMO modification and plays an inhibitory role in muscle and neuronal differentiation. *Experimental Cell Research,* **315**, 3543-3553.

432. Yu, B., Shao, Y., Zhang, C., Chen, Y., Zhong, Q., Zhang, J., *et al.* (2009) BS69 undergoes SUMO modification and plays an inhibitory role in muscle and neuronal differentiation. *Experimental cell research,* **315**, 3543-3553.
433. Yu, Y., Yao, R., Wang, L., Fan, Y., Huang, X., Hirschhorn, J., *et al.* (2015) De novo mutations in ARID1B associated with both syndromic and non-syndromic short stature. *BMC genomics,* **16**, 701-701.
434. Zariwala, M. A., Knowles, M. R. & Omran, H. (2007) Genetic Defects in Ciliary Structure and Function. *Annual Review of Physiology,* **69**, 423-450.
435. Zhang, H. H., Huang, J., Düvel, K., Boback, B., Wu, S., Squillace, R. M., *et al.* (2009) Insulin Stimulates Adipogenesis through the Akt-TSC2-mTORC1 Pathway. *PLoS ONE,* **4**, e6189-e6189.
436. Zhang, Y., Chen, K., Sloan, S. A., Bennett, M. L., Scholze, A. R., O'Keeffe, S., *et al.* (2014) An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience,* **34**, 11929-11947.
437. Zhang, Y., Chen, K., Sloan, S. A., Bennett, M. L., Scholze, A. R., O'Keeffe, S., *et al.* (2014) An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience,* **34**, 11929-11947.
438. Zhao, C., Deng, W. & Gage, F. H. (2008) Mechanisms and Functional Implications of Adult Neurogenesis, pp. 645-660. Cell Press.
439. Zoghbi, H. Y. (2003) Postnatal neurodevelopmental disorders: meeting at the synapse? *Science (New York, N.Y.),* **302**, 826-830.
440. Zola, S. M., Squire, L. R., Teng, E., Stefanacci, L., Buffalo, E. A. & Clark, R. E. (2000) Impaired recognition memory in monkeys after damage limited to the hippocampal region. *The Journal of neuroscience : the official journal of the Society for Neuroscience,* **20**, 451-463.
441. Zola-Morgan, S., Squire, L. R., Amaral, D. G., Amaral, D. G. & Holland, P. (1986) Human amnesia and the medial temporal region: enduring memory impairment following a bilateral lesion limited to field CA1 of the hippocampus. *The Journal of neuroscience : the official journal of the Society for Neuroscience,* **6**, 2950-2967.
442. Zong, L., Lu, C., Zhao, Y., Li, Q., Han, D., Yang, W., *et al.* (2012) Clue to a New Deafness Gene: A Large Chinese Nonsyndromic Hearing Loss Family Linked to DFNA4, pp. 653-657.