A hPSCs-based model to study the biology of the schizophrenia and autism risk gene CYFIP1

Thesis submitted for the degree of Doctor of Philosophy

School of Medicine, Cardiff University

2018

Claudia Tamburini
DECLARATION

This work has not been submitted in substance for any other degree or award at this or any other
university or place of learning, nor is being submitted concurrently in candidature for any degree
or other award.

Signed ............................................ Date ............................................................

STATEMENT 1

This thesis is being submitted in partial fulfilment of the requirements for the degree of Doctor of
Philosophy, PhD

Signed ............................................ Date ............................................................

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated,
and the thesis has not been edited by a third party beyond what is permitted by Cardiff University's
Policy on the Use of Third Party Editors by Research Degree Students. Other sources are
acknowledged by explicit references. The views expressed are my own.

Signed ............................................ Date ............................................................

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available online in the University’s Open
Access repository and for inter-library loan, and for the title and summary to be made available to
outside organisations.

Signed ............................................ Date ............................................................

STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

I hereby give consent for my thesis, if accepted, to be available online in the University’s Open
Access repository and for inter-library loans after expiry of a bar on access previously approved
by the Academic Standards & Quality Committee.

Signed ............................................ Date ............................................................
Acknowledgements

First and foremost, I wish to thank my supervisor, Prof. Meng Li, for offering me the opportunity to work on this project and guiding me through it, keeping me focused and allowing me to develop my own ideas.

Thank you also to all the present and past members of the Li lab, for creating a great work environment. I could not have hoped for a better group of people to share my days with. Above all, I would like to thank Ines, for teaching me hESCs culture and for being a good friend during both the London and Cardiff days; Xinsheng for always having an answer to my molecular biology problems; Shanthi, for keeping the lab running smoothly; Craig, for the support with iPSCs reprogramming; Dani, for his help with the bioinformatic analysis and Zoe, for her valuable advice on this thesis.

I also need to thank my wonderful family and friends in Italy, for understanding my choices and always encouraging me; in particular, Monica, Laura, Gloria, Stefania e Ilaria, for every time I went back home and felt like I had never left. I miss you every day.

Thank you to my boyfriend and best friend, James, for the continuous love and support, through the highs and lows that come with a career in science. Thank you for being my rock and believing in me more than I do myself. I cannot put on paper what this means to me.

And, finally, thank you to my parents, to whom I owe everything I have and everything I am. I hope I made you proud.
Abstract

Microdeletions and microduplications at the chromosomal locus ch15q11.2 have been linked to an increased risk of developing several psychiatric and neurodevelopmental conditions. Abnormal expression levels of CYFIP1, one of the genes within this region, have been shown by several studies to cause alterations to the morphology and physiology of neuronal cells. However, despite its specific expression in progenitors of the developing cortex, little is known about the role played by CYFIP1 in neural development.

This thesis presents an investigation of the biological mechanisms, through which CYFIP1 affects cortex development in vitro using human embryonic stem cells (hESCs). To this end, lines of hESCs either expressing a CYFIP1 transgene (CYFIP1tg) or harbouring a loss-of-function deletion (CYFIP1ko) were derived to mimic increased and decreased level of CYFIP1 expression, respectively, and differentiated into cortical glutamatergic neurons. Disruptions to the normal levels of CYFIP1 resulted into abnormal formation of neural rosettes and altered kinetics of neuronal differentiation. Importantly, CYFIP1 overexpression seemed to promote self-renewal of the progenitor pool, while loss of this gene had an opposite effect.

Whole-genome transcriptomic analysis revealed the dysregulation of numerous pathways, including WNT signalling, cell adhesion and mitochondria metabolism. In line with this, the expression levels of N-Cadherin and the phosphorylation pattern of β-Catenin suggested that the signalling axis involving these two proteins is altered in CYFIP1tg and CYFIP1ko cortical progenitors. Moreover, pharmacological inhibition of WNT and AKT, which mediate this signalling, was able to rescue the excessive proliferation of NPCs associated with CYFIP1 overexpression. Finally, the presence of altered mitochondrial dynamics in CYFIP1tg and CYFIP1ko NPCs and neurons was validated using a high-content screening system.

These data demonstrate that CYFIP1 plays a role in the regulation of cortex development and, consequently, in the clinical manifestations associated with 15q11.2 CNVs. Furthermore, they strengthen the hypothesis of a developmental origin for psychiatric conditions like schizophrenia and autism spectrum disorders.
Contents
List of figures .......................................................................................................................... viii
List of tables .......................................................................................................................... ix
List of abbreviations .............................................................................................................. x
1 Introduction .......................................................................................................................... 1
  1.1 Relevance of Ch15q11.2 CNVs in the context of psychiatric disorders ...................... 1
  1.1.1 Genetics of psychiatric disorders: an overview ...................................................... 1
  1.1.2 The Ch15q11.2 locus .............................................................................................. 8
  1.2 CYFIP1 biology ............................................................................................................ 11
  1.2.1 Role of Cyfip1 in the WRC .................................................................................. 12
  1.2.2 Role of Cyfip1 in the regulation of mRNA translation ........................................ 13
  1.2.3 Expression pattern of CYFIP1 ............................................................................. 15
  1.2.4 In vivo and in vitro studies on Cyfip1 alterations in neuronal cells ...................... 16
  1.3 The use of human Pluripotent Stem Cells for in vitro modelling of neurodevelopmental psychiatric disorders .......................................................... 17
  1.3.1 Generation of disease relevant neuronal cell types from human PSCs ............... 19
  1.3.2 Current status of iPSCs-based research for psychiatric disorders ....................... 21
  1.3.3 CRISPR-based genome editing as a powerful alternative to model neurodevelopmental disorders .................................................................................. 24
  1.4 Cortex development .................................................................................................... 31
  1.4.1 Early corticogenesis: progenitor subtypes and characteristics ......................... 31
  1.4.2 Formation of the cortical layers .......................................................................... 34
  1.5 Aims of the project ...................................................................................................... 37
2 Methods and materials ...................................................................................................... 38
  2.1 Cell culture .................................................................................................................. 38
  2.1.1 hESC culture ....................................................................................................... 38
  2.1.2 Monolayer differentiation into cortical glutamatergic neurons ............................ 38
  2.1.3 hESC differentiation into brain organoids ........................................................... 40
2.1.4 Other chemicals used in differentiation experiments ........................................40

2.2 Fibroblast reprogramming into iPSCs .................................................................41

2.3 Molecular cloning for CYFIP1-overexpression vector ........................................41

2.4 CRISPR/Cas9 targeting: gRNA design and synthesis .......................................42

2.5 Transfection of hESCs ......................................................................................42

2.5.1 Nucleofection ...............................................................................................42

2.5.2 RNA transfection .........................................................................................43

2.6 Gene expression analysis ..................................................................................43

2.6.1 RNA extraction .............................................................................................43

2.6.2 DNase treatment ..........................................................................................44

2.6.3 Reverse transcription ....................................................................................44

2.6.4 qPCR ............................................................................................................44

2.7 Genotyping .......................................................................................................45

2.7.1 Genomic DNA extraction .............................................................................45

2.7.2 PCR and DNA electrophoresis ....................................................................45

2.7.3 Surveyor assay ..............................................................................................46

2.7.4 PCR cloning and sequencing analysis ..........................................................46

2.8 Western Blotting ...............................................................................................47

2.9 Immunocytochemistry ......................................................................................48

2.9.1 Immunofluorescence staining ......................................................................48

2.9.2 EdU proliferation assay ................................................................................49

2.9.3 Imaging and picture analysis ........................................................................49

2.10 High Content analysis of mitochondria .........................................................49

2.10.1 Mitotracker staining ....................................................................................49

2.10.2 Cell Insight CX7 High Content Screening (HCS) platform ......................49

2.11 Flow cytometry analysis ................................................................................49

2.12 Transcriptome analysis via RNA sequencing (with Daniel Cabezas de la Fuente) 50

3 Derivation of hESCs with altered CYFIP1 dosage and iPSCs carrying 15q11.2 CNVs ....52
6.3 Future directions

7 Bibliography
List of figures

Figure 1.1 Scheme of the 15q11-13 chromosomic locus................................................................. 9
Figure 1.2 CYFIP proteins participation in the regulation of actin polymerisation by the WRC... 13
Figure 1.3 Scheme of the proposed Cap-dependent translation mechanism involving CYFIP1.... 14
Figure 1.4 Cyfip1 expression in the mouse developing brain...................................................... 15
Figure 1.5 Scheme of cortical neurogenesis.................................................................................. 34
Figure 1.6 Schematic representation of cortical layer formation during mouse corticogenesis... 36
Figure 2.1 Monolayer differentiation of hESCs into cortical pyramidal neurons............................ 39
Figure 2.2 Generation of gRNAs.................................................................................................. 42
Figure 3.1 CYFIP1-overexpression vector construction................................................................. 54
Figure 3.2 Characterisation of CYFIP1tg hESCs........................................................................... 56
Figure 3.3 Validation of RNA transfection method....................................................................... 57
Figure 3.4 Screening and characterisation of CYFIP1-mutant hESCs......................................... 59
Figure 3.5 Immunofluorescent staining in iPSCs and iPSCs-derived neuronal cultures.............. 60
Figure 3.6 RT-qPCR for 15q11.2 genes in iPSCs-derived neuronal cells....................................... 62
Figure 3.7 Western blots for 15q11.2 genes in iPSCs-derived neuronal cells............................... 63
Figure 4.1 Expression of cortical progenitor markers in CYFIP1tg and CYFIP1ko cultures........ 69
Figure 4.2 Expression of proliferation markers in CYFIP1tg, CYFIP1ko and control NPCs.......... 70
Figure 4.3 Abnormal neural rosettes and VZ-like areas formed by CYFIP1tg neural progenitors.... 72
Figure 4.4 Disorganised neural rosettes in CYFIP1ko progenitor cultures................................. 73
Figure 4.5 Overproliferation of neural progenitors caused by CYFIP1 overexpression............. 76
Figure 4.6 Delayed appearance of cortical layer markers in CYFIP1t cultures............................. 77
Figure 4.7 Increased neurogenic ability of CYFIP1ko cortical progenitors.................................. 79
Figure 4.8 Reduced proliferation of CYFIP1ko progenitor pool.................................................. 80
Figure 5.1 Biological processes and signalling pathways significantly affected by altered CYFIP1 levels. .................................................................................................................. 86
Figure 5.2 Neurogenesis-related genes affected by abnormal CYFIP1 levels............................ 88
Figure 5.3 Western blot analysis of NCAD and βCAT in CYFIP1tg and CYFIP1ko.......................... 90
Figure 5.4 Reduction of βCAT activity downstream of NCAD rescues the delayed neurogenesis of CYFIP1tg cultures. .......................................................................................... 93
Figure 5.5 Quantification of GABAergic cells in CYFIP1tg and control cultures..................... 94
Figure 5.6 Mitochondria-related genes affected by abnormal CYFIP1 levels............................. 96
Figure 5.7 Alterations of mitochondria dynamics in CYFIP1tg and CYFIP1ko NPCs and neurons. 98
List of tables

Table 1.1 Genes in which GWAS have reported the presence of risk variants linked to ASD or SZ ......................................................... 3
Table 1.2 CNVs increasing risk for ASD and/or SZ .................................................................................................................. 5
Table 1.3 Monogenic syndromes associated with ASD (table based on Caglayan et al 2010 and Sztainberg and Zoghbi 2016) ........................................................................................................................... 7
Table 1.4 List of publications on iPSCs-based models of neurodevelopmental psychiatric disorders published to date ........................................................................................................... 27
Table 2.1 composition of media used for monolayer differentiation .................................................................................. 40
Table 2.2 composition of media used for organoids differentiation ............................................................................... 40
Table 2.3 Primers used for qPCR ..................................................................................................................................................... 45
Table 2.4 Primers used for genotyping of CYFIP1 locus ........................................................................................................ 46
Table 2.5 Antibodies used for western blotting ..................................................................................................................... 47
Table 2.6 Antibodies used for immunofluorescence ................................................................................................................ 48
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>4E-BP</td>
<td>4E-Binding Protein</td>
</tr>
<tr>
<td>ADHD</td>
<td>Attention Deficit Hyperactivity Disorder</td>
</tr>
<tr>
<td>AMPA</td>
<td>α-Amino-3-Hydroxy-5-Methyl-4-Isoxazolepropionic Acid</td>
</tr>
<tr>
<td>Arp2/3</td>
<td>Actin-Related Protein 2/3</td>
</tr>
<tr>
<td>AS</td>
<td>Angelman Syndrome</td>
</tr>
<tr>
<td>ASD</td>
<td>Autism Spectrum Disorder</td>
</tr>
<tr>
<td>βCAT</td>
<td>β-Catenin</td>
</tr>
<tr>
<td>BPs</td>
<td>Break-Points</td>
</tr>
<tr>
<td>CNVs</td>
<td>Copy Number Variants</td>
</tr>
<tr>
<td>CP</td>
<td>Cortical Plate</td>
</tr>
<tr>
<td>CRISPR</td>
<td>Clustered Regularly Interspaced Short Palindromic Repeats</td>
</tr>
<tr>
<td>CTIP2</td>
<td>COUP TF1-interacting protein 2</td>
</tr>
<tr>
<td>CYFIP1/2</td>
<td>Cytoplasmic FMRP interacting protein 1/2</td>
</tr>
<tr>
<td>DEGs</td>
<td>Differentially Expressed Genes</td>
</tr>
<tr>
<td>DISC1</td>
<td>Disrupted in Schizophrenia 1</td>
</tr>
<tr>
<td>E</td>
<td>Embryonic day</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EdU</td>
<td>5-ethyl-2'-deoxyuridine</td>
</tr>
<tr>
<td>eIF4A/E/G</td>
<td>Eukaryotic Initiation Factor 4A/E/G</td>
</tr>
<tr>
<td>FACS</td>
<td>Fluorescence-Activated Cell Sorting</td>
</tr>
<tr>
<td>FGF</td>
<td>Fibroblast Growth Factor</td>
</tr>
<tr>
<td>FMRP</td>
<td>Fragile X Mental Retardation Protein</td>
</tr>
<tr>
<td>FOXG1</td>
<td>Forkhead Box G1</td>
</tr>
<tr>
<td>FXS</td>
<td>Fragile X Syndrome</td>
</tr>
<tr>
<td>GABA</td>
<td>Gamma-Aminobutyric Acid</td>
</tr>
<tr>
<td>gRNA</td>
<td>Guide RNA</td>
</tr>
<tr>
<td>GWAS</td>
<td>Genome-Wide Association Studies</td>
</tr>
<tr>
<td>HCS</td>
<td>High-Content Screening</td>
</tr>
<tr>
<td>hESCs</td>
<td>Human Embryonic Stem Cells</td>
</tr>
<tr>
<td>hPSCs</td>
<td>Human Pluripotent Stem Cells</td>
</tr>
<tr>
<td>ID</td>
<td>Intellectual Disability</td>
</tr>
<tr>
<td>iN</td>
<td>Induced Neurons</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>INs</td>
<td>Interneurons</td>
</tr>
<tr>
<td>iPSCs</td>
<td>Induced Pluripotent Stem Cells</td>
</tr>
<tr>
<td>IRES</td>
<td>Internal Ribosomal Entry Site</td>
</tr>
<tr>
<td>IZ</td>
<td>Intermediate Zone</td>
</tr>
<tr>
<td>KD</td>
<td>Knock-Down</td>
</tr>
<tr>
<td>KO</td>
<td>Knock-Out</td>
</tr>
<tr>
<td>MECP2</td>
<td>Methyl CpG binding Protein 2</td>
</tr>
<tr>
<td>Mg++</td>
<td>Magnesium</td>
</tr>
<tr>
<td>MGE</td>
<td>Medial Ganglionic Eminence</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA</td>
</tr>
<tr>
<td>NCAD</td>
<td>N-Cadherin (Cadherin 2)</td>
</tr>
<tr>
<td>NEUN</td>
<td>Neuronal Nuclei</td>
</tr>
<tr>
<td>NIPA1/2</td>
<td>Non-Imprinted In Prader-Willi/Angelman Syndrome 1/2</td>
</tr>
<tr>
<td>NPCs</td>
<td>Neural Progenitor Cells</td>
</tr>
<tr>
<td>ORF</td>
<td>Open Reading Frame</td>
</tr>
<tr>
<td>PABP</td>
<td>Poly-A Binding Protein</td>
</tr>
<tr>
<td>PAC</td>
<td>Puromycin N-acetyltransferase</td>
</tr>
<tr>
<td>PAX6</td>
<td>Paired Box 6</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>PWS</td>
<td>Prader-Willi Syndrome</td>
</tr>
<tr>
<td>qPCR</td>
<td>Quantitative PCR</td>
</tr>
<tr>
<td>Rac1</td>
<td>Rac Family Small GTPase 1</td>
</tr>
<tr>
<td>RGCs</td>
<td>Radial Glia Cells</td>
</tr>
<tr>
<td>RELN</td>
<td>Reelin</td>
</tr>
<tr>
<td>ROS</td>
<td>Reactive Oxygen Species</td>
</tr>
<tr>
<td>VCA</td>
<td>Verprolin-homology, Central and Acidic regions</td>
</tr>
<tr>
<td>WASP</td>
<td>Wiskott–Aldrich syndrome protein</td>
</tr>
<tr>
<td>WAVE</td>
<td>WASP family verprolin homologous protein</td>
</tr>
<tr>
<td>WES</td>
<td>Whole-Exome Studies</td>
</tr>
<tr>
<td>WNT</td>
<td>Wingless Type</td>
</tr>
<tr>
<td>WRC</td>
<td>Wave Regulatory Complex</td>
</tr>
<tr>
<td>WT</td>
<td>Wild Type</td>
</tr>
<tr>
<td>SATB2</td>
<td>Special AT-rich sequence-binding protein 2</td>
</tr>
<tr>
<td>SD</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>SHANK1/3</td>
<td>SH3 And Multiple Ankyrin Repeat Domains 2</td>
</tr>
<tr>
<td>SHH</td>
<td>Sonic Hedgehog</td>
</tr>
<tr>
<td>shRNA</td>
<td>Silencing RNA</td>
</tr>
<tr>
<td>SVZ</td>
<td>Subventricular Zone</td>
</tr>
<tr>
<td>SZ</td>
<td>Schizophrenia</td>
</tr>
<tr>
<td>TBR1/2</td>
<td>T-Box, Brain 1</td>
</tr>
<tr>
<td>TG</td>
<td>Transgene</td>
</tr>
<tr>
<td>TSC1/2</td>
<td>Tuberous Sclerosis Complex 1/2</td>
</tr>
<tr>
<td>TUBGCP5</td>
<td>Tubulin Gamma Complex Associated Protein 5</td>
</tr>
<tr>
<td>VZ</td>
<td>Ventricular Zone</td>
</tr>
</tbody>
</table>
1 Introduction

Psychiatric disorders such as autism spectrum disorders (ASD) and schizophrenia (SZ) are common conditions that, combined, are estimated to affect more than 1% of the population worldwide (Baxter et al., 2015; Saha et al., 2005 and WHO.int). Many people affected by these disorders require long-term care and represent a high economic burden for society (Buescher et al., 2014; Chong et al., 2016).

SZ usually becomes manifest post-puberty and is characterised by a combination of symptoms that are classified into “positive”, such as hallucinations and delusions, “negative”, such as apathy and lack of speech, and “cognitive”, which include defects of working memory, attention and verbal abilities (Patel et al., 2014). ASD are a group of heterogeneous conditions that usually arise during early childhood and are defined by the presence of some common aspects, such as restrictive and repetitive behaviours and deficits in social interaction and communication (Fakhoury, 2015).

Both SZ and ASD arise from a combination of genetic and environmental factors. The last few years have seen the identification of hundreds of chromosomic loci responsible for an increased risk of developing these conditions. However, the way in which genetic mutations contribute to the disease pathology and, ultimately, the cellular mechanisms underlying these conditions are still largely unknown. To date, this represents perhaps the biggest obstacle to the development of new therapies. In vitro models in the form of human pluripotent stem cells (hPSCs)-derived neurons represent a promising approach to investigate the function of disease-risk genes and can also provide a platform for drug development studies.

1.1 Relevance of Ch15q11.2 CNVs in the context of psychiatric disorders

1.1.1 Genetics of psychiatric disorders: an overview

Twin and adoption studies have been fundamental to understand the contribution of genetic and environmental factors to the aetiology of psychiatric disorders. The environmental factors most commonly associated with SZ are birth or residence in urban areas, season of birth, usually linked to a higher risk of pre-natal infections, starvation, obstetric complications, and advanced paternal age (Gejman et al., 2011). However, all these factors seem to have a relatively small effect compared to the influence of genetics.
Heritability, a parameter used to estimate how much of the variability of a phenotypic trait is due to the genetic variability in a given population, is estimated to be above 80% for SZ (Sullivan, 2005). In fact, monozygotic twins have a higher concordance rate than dizygotic twins and the risk of SZ in the offspring of schizophrenic parents is the same whether they are raised by their biological or adoptive unaffected parents (Gejman et al., 2011). Similarly, several studies estimated ASD heritability between 50% and 90% and the most commonly implicated environmental factors include prenatal viral infection and exposure to toxins, advanced paternal and maternal age and prenatal and perinatal stress (Freitag et al., 2010; Grabrucker, 2013; Kim & Leventhal, 2015).

The first attempts to discover the genetic causes of SZ and ASD were represented by linkage and candidate gene studies. Linkage studies are based on the concept that alleles close together are transmitted together, as a unit, through meiosis (Lander & Kruglyak, 1995). Thus, they were used to map chromosomal regions that follow the same inheritance pattern between affected family members. In the context of SZ, some of the most significant chromosomic regions that emerged from these studies are represented by 1q21-22 (Brzustowicz et al., 2000; Rosa et al., 2002) and 6p24-p22 (Moises et al., 1995; Schwab et al., 2000; Straub, 1995). Examples of loci associated to ASD by at least two independent linkage studies are represented by 3p25 (Lauritsen et al., 1999; McCauley et al., 2005; Ylisaukko-Oja et al., 2006) and 7q35 (Alarcón et al., 2008; Alarcón et al., 2002; Ylisaukko-Oja et al., 2006).

The candidate gene approach is a hypothesis-driven method, which aims to test if the frequency of a specific allele/haplotype of a gene of interest is higher in affected than unaffected subjects (Kwon & Goate, 2000). Well known candidate genes include Neuregulin 1 (NRG1) (Stefansson et al., 2002), Disrupted in schizophrenia 1 (DISC1) (Blackwood et al., 2001) and Dysbindin (DTNBP1) (Straub et al., 2002) for SZ and the oxytocin receptor (OXTR) (Wu et al., 2005), Contactin-associated protein-like 2 (CNTNAP2)(Arking et al., 2008) and Reelin (RELN) (Persico et al., 2001) for ASD.

Several more chromosomic loci and genes were identified by these early genetic studies, but most of them were found difficult to reproduce. Two reasons for this could be the small effect of the genetic variants analysed and the small sample size (Farrell et al., 2015; Freitag et al., 2010). Indeed, linkage studies had been successful in identifying the cause of disorders following a Mendelian pattern of inheritance, such as mutations of presenilin 1 in Alzheimer’s disease, but it is now known that most psychiatric disorders are likely caused by several common genetic variations with different degrees of penetrance (George-hyslop, 2000; Schellenberg et al., 1992; Sherrington et al., 1995).
More significant advances have been achieved in recent years thanks to the rapid evolution of new high-throughput microarray platforms and the formation of large scientific consortia. This has facilitated efficient processing of the genetic information of large cohorts of patients and controls. Genome-wide association studies (GWAS), analysis of chromosomic copy number variations (CNVs) and exome sequencing are the new techniques that significantly contributed to such advances.

(I) **GWAS**

GWAS are based on the analysis of markers across the whole genome to determine if some genetic variations are more abundant in individuals affected by a disease than the unaffected subject population. A genetic marker associated with the disease can represent in itself a causative variant, or it can be in linkage disequilibrium (LD) with the causative variant (Bush & Moore, 2012). This approach is hypothesis-independent and has led to the identification of many common variants, in the form of single nucleotide polymorphisms (SNPs) or chromosomic copy number variations (CNVs) (McCarroll, 2008).

SNPs are single base substitutions, which are relatively common in the genome (~1 every 300 bases) and within populations (>5%) (Schwab & Wildenauer, 2013). Overall it is estimated that each SNP associated with a psychiatric disorder is responsible for a small contribution towards the total genetic risk (Gejman et al., 2011; Yoo, 2015). Table 1.1 contains a list of genes in which SNPs for ASDs and SZ have been confirmed by more than one study.

<table>
<thead>
<tr>
<th>Disorder</th>
<th>Gene symbol</th>
<th>Gene name</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASD</td>
<td>CDH9</td>
<td>Cadherin 9</td>
<td>(Ma et al., 2009; Wang et al., 2009)</td>
</tr>
<tr>
<td>ASD</td>
<td>CDH10</td>
<td>Cadherin 10</td>
<td>(Ma et al., 2009; Wang et al., 2009)</td>
</tr>
<tr>
<td>ASD</td>
<td>CNTNAP2</td>
<td>Contactin-associated protein-like 2</td>
<td>(Anney et al., 2012; Arking et al., 2008)</td>
</tr>
<tr>
<td>ASD</td>
<td>JARID2</td>
<td>Jumonji and AT-Rich Interaction Domain Containing 2</td>
<td>(Liu et al., 2015; Weiss, Arking et al., 2009)</td>
</tr>
<tr>
<td>ASD</td>
<td>MACROD2</td>
<td>MACRO Domain-Containing Protein 2</td>
<td>(Anney et al., 2010, 2012; Vieland et al., 2017)</td>
</tr>
<tr>
<td>ASD</td>
<td>SEMAS5A</td>
<td>Semaphorin 3A</td>
<td>(Vieland, 2017; Weiss et al., 2009)</td>
</tr>
<tr>
<td>SZ</td>
<td>AKT3</td>
<td>AKT Serine/Threonine Kinase 3</td>
<td>(Ripke et al., 2013, 2014)</td>
</tr>
<tr>
<td>SZ</td>
<td>ARL3</td>
<td>ADP Ribosylation Factor Like GTPase 3</td>
<td>(Ripke et al., 2014; H. Yu et al., 2016)</td>
</tr>
<tr>
<td>SZ</td>
<td>AS3MT</td>
<td>Arsenite Methyltransferase</td>
<td>(Ripke et al., 2014; H. Yu et al., 2016)</td>
</tr>
<tr>
<td>SZ</td>
<td>CACNA1C</td>
<td>Calcium Voltage-Gated Channel Subunit Alpha1C</td>
<td>(Purcell et al., 2009; Ripke et al., 2013, 2014)</td>
</tr>
<tr>
<td>SZ</td>
<td>CNNM2</td>
<td>Cyclin And CBS Domain Divalent Metal Cation Transport Mediator 2</td>
<td>(Ripke et al., 2011, 2014)</td>
</tr>
<tr>
<td>SZ</td>
<td>CSMD1</td>
<td>CUB And Sushi Multiple Domains 1</td>
<td>(Ripke et al., 2011, 2014)</td>
</tr>
<tr>
<td>SZ</td>
<td>CSMD2</td>
<td>CUB And Sushi Multiple Domains 2</td>
<td>(Ripke et al., 2014; H. Yu et al., 2016)</td>
</tr>
<tr>
<td>SZ</td>
<td>ITIH3</td>
<td>Inter-Alpha-Trypsin Inhibitor Heavy Chain 3</td>
<td>(Ripke et al., 2013, 2014)</td>
</tr>
<tr>
<td>SZ</td>
<td>MAD1L1</td>
<td>MAD1 Mitotic Arrest Deficient Like 1</td>
<td>(Ripke et al., 2013, 2014)</td>
</tr>
<tr>
<td>SZ</td>
<td>MIR137</td>
<td>MicroRNA 137</td>
<td>(Ripke et al., 2011, 2013)</td>
</tr>
<tr>
<td>SZ</td>
<td>NRGN2</td>
<td>Neurogranin</td>
<td>(Ripke et al., 2014; Stefansson et al., 2009)</td>
</tr>
<tr>
<td>SZ</td>
<td>NT5C2</td>
<td>5'-Nucleotidase, Cytosolic II</td>
<td>(Ripke et al., 2011, 2014)</td>
</tr>
<tr>
<td>SZ</td>
<td>PCGEM1</td>
<td>PCGEM1, Prostate-Specific Transcript (Non-Protein Coding)</td>
<td>(Ripke et al., 2011, 2014)</td>
</tr>
<tr>
<td>SZ</td>
<td>TCF4</td>
<td>Transcription Factor 4</td>
<td>(Ripke et al., 2011, 2014; Stefansson et al., 2009)</td>
</tr>
<tr>
<td>SZ</td>
<td>TSNARE1</td>
<td>T-SNARE Domain Containing 1</td>
<td>(Ripke et al., 2013, 2014)</td>
</tr>
<tr>
<td>SZ</td>
<td>VRK2</td>
<td>Vaccinia Related Kinase 2</td>
<td>(Ripke et al., 2014; H. Yu et al., 2016)</td>
</tr>
<tr>
<td>SZ</td>
<td>ZNF804A</td>
<td>Zinc Finger Protein 804A</td>
<td>(O'Donovan et al., 2008; Ripke et al., 2014; Steinberg et al., 2011; Williams et al., 2011)</td>
</tr>
</tbody>
</table>

ASD, autism spectrum disorder; SZ, schizophrenia.

CNVs are DNA segments of at least 1kb that are found deleted or duplicated in different genomes (Hosak, 2013). They can be inherited or they can be the result of de novo mutations, as they originate through several mutational mechanisms, mainly non-allelic homologous recombination between regions containing low-copy repeats (Gu, Zhang, Lupski, 2008). These chromosomic rearrangements are very frequent in the genome and contribute to human genomic variation.
Interestingly, several studies on both SZ and ASD reported a higher frequency of \textit{de novo} CNVs in the genome of affected individuals than in the control population (International Schizophrenia Consortium, 2008; Marshall et al., 2008a; Sebat et al., 2010; Xu et al., 2008). These findings suggest that genomic instability may be contributing to the pathology of these disorders.

CNVs affect the expression level of the genes located in the deleted or duplicated regions. A popular interpretation of the link between this mechanism and psychiatric disorders is that CNVs alter the tightly regulated expression of dosage-sensitive genes important in neural development (Hosak, 2013; Tam et al., 2009). In agreement with this, it has been demonstrated that pathogenic CNVs are enriched for genes classified as ohnologs, when compared with CNVs identified in the general population. Ohnologs are duplicated genes originated by whole genome duplication events during vertebrate evolution. Many of them are developmental genes and members of protein complexes that are known to be highly dosage-sensitive (McLysaght et al., 2014). The effects of such changes in gene dosage are very variable, as several CNVs are associated with more than one disorder and can be also present, at lower frequency, in healthy control populations (Malhotra & Sebat, 2012). The most characterised CNVs associated to date with SZ and/or ASD are listed in Table 1.2

\textbf{Table 1.2 CNVs increasing risk for ASD and/or SZ}

<table>
<thead>
<tr>
<th>CNV</th>
<th>Disorder(s)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>1q21.1 del/dup</td>
<td>ASD/SZ</td>
<td>(Bucan et al., 2009; Kirov et al., 2014a; Levinson et al., 2011; Pinto et al., 2010a; Rees et al., 2014; Sanders et al., 2011; Stefansson et al., 2008; Vacic et al., 2011a)</td>
</tr>
<tr>
<td>2p16.3 del</td>
<td>ASD/SZ</td>
<td>(Bremer et al., 2011; Glessner et al., 2009a; Kirov et al., 2014b; Pinto et al., 2010b; Rees et al., 2014)</td>
</tr>
<tr>
<td>3q29 del</td>
<td>ASD/SZ</td>
<td>(Kirov et al., 2014b; Levinson et al., 2011; Mulle et al., 2010; Pinto et al., 2010b; Rees et al., 2014; Sanders et al., 2011; Vacic et al., 2011b)</td>
</tr>
<tr>
<td>15q11.2 del/dup</td>
<td>ASD/SZ</td>
<td>(Kirov et al., 2009, 2014b; Rees et al., 2014; Stefansson et al., 2008; Tam et al., 2010)</td>
</tr>
<tr>
<td>15q11.2-13.1 dup</td>
<td>ASD/SZ</td>
<td>(Christian et al., 2008; Glessner et al., 2009b; Marshall et al., 2008a; Pinto et al., 2010b; Sanders et al., 2011; Vacic et al., 2011b)</td>
</tr>
<tr>
<td>15q13.3 del</td>
<td>ASD/SZ</td>
<td>(Bremer et al., 2011; Kirov et al., 2014b; Levinson et al., 2011; Pinto et al., 2010b; Rees et al., 2014; Sanders et al., 2011; Stefansson et al., 2008; The International Schizophrenia Consortium, 2008; Vacic et al., 2011b)</td>
</tr>
<tr>
<td>16p13.11 dup</td>
<td>ASD/SZ</td>
<td>(Ingason et al., 2011; Kirov et al., 2014b; Pinto et al., 2010b; Rees et al., 2014; Sanders et al., 2011; Vacic et al., 2011b)</td>
</tr>
<tr>
<td>16p11.2 del/dup</td>
<td>ASD/SZ</td>
<td>(Bremer et al., 2011; Glessner et al., 2009b; Kirov et al., 2014b; Levinson et al., 2011; Marshall et al., 2008b; McCarthy et al., 2009; Pinto et al., 2010b; Rees et al., 2014; Sanders et al., 2011)</td>
</tr>
<tr>
<td>17p12 del</td>
<td>ASD/SZ</td>
<td>(Kirov et al., 2009, 2014b; Pinto et al., 2010b; Rees et al., 2014; Sanders et al., 2011; Vacic et al., 2011b)</td>
</tr>
<tr>
<td>22q11.2 del/dup</td>
<td>ASD/SZ</td>
<td>(Bremer et al., 2011; Glessner et al., 2009b; Kirov et al., 2014b; Marshall et al., 2008b; Pinto et al., 2010b; Rees et al., 2014; Sanders et al., 2011; Stefansson et al., 2008; The International Schizophrenia Consortium, 2008; Vacic et al., 2011b)</td>
</tr>
</tbody>
</table>

ASD, autism spectrum disorder; SZ, schizophrenia; del, deletion; dup, duplication.

(II) Exome sequencing

GWAS studies are, by definition, limited to the identification of common SNPs and CNVs and may therefore fail to identify highly penetrant mutations that occur at a lower frequency in a population, due to the decreased reproductive fecundity of the affected individuals (Power et al., 2013). Because of the large effect size of such rare variants, they are likely to be easier in their functional interpretation, providing an important contribution to the understanding of the disorders with which they are associated (Gratten, 2016).

The identification of rare genetic mutations requires the application of DNA sequencing technologies, such as whole-exome sequencing (WES). This type of analysis can be performed on case-parents trios, to investigate the effect of de novo mutations, or on cohorts of cases and controls. WES is still an emerging approach in the field of psychiatric disorders, likely because of its relatively high cost compared to older technologies. For the same reason, the studies published so far were conducted on small samples (Kato, 2015; Sener, Canatan, & Ozkul, 2016).

So far in the context of SZ, the application of WES has not strongly implicated any individual genes, with the exception of SET Domain Containing 1A (SETD1A) (Singh et al., 2016). However, several studies have reported that rare (1:10000) and ultra-rare (1:100000) variants disrupting protein-
Coding genes are present at a higher frequency in SZ cases than in controls and that they are particularly enriched in specific gene sets, such as voltage-gated calcium channels and synaptic proteins (Genovese et al., 2016; Purcell et al., 2014). A higher frequency of gene-disrupting mutations in cases versus controls has also been reported by WES studies for ASDs (Iossifov et al., 2012). The sequencing studies on ASDs present in the literature are several, but they are all based on a maximum of a few hundred individuals. Nonetheless, some potential causative genes appear in more than one study. Some of these are Forkhead box protein P1 (FOXP1), a transcription factor involved in gene expression regulation during brain development (Egawa et al., 2015; O’Roak et al., 2011), Glutamate receptor, ionotropic, N-methyl D-aspartate 2B (GRIN2B), and Sodium voltage-gated channel alpha subunit 1 and 2 (SCN1A and 2A). Disruption of these genes is also associated with epilepsy (Codina-Solà et al., 2015; O’Roak et al., 2011; Roak et al., 2012; Sanders et al., 2013).

At present, WES appears to be a promising approach to complement the large number of GWAS data already published, but bigger sample sizes are needed to identify significant rare variants with more confidence.

Finally, it is important to mention that autistic behaviour is also one of the clinical manifestations of several genetic syndromes listed in Table 3. Such disorders have a known high-penetrant genetic cause and their pathology involves several organs. Generally, typical manifestations include dysmorphic features, metabolic dysfunctions, intellectual disability and seizures (Caglayan, 2010; Sztainberg & Zoghbi, 2016). However, these syndromes account for only a small proportion (5-7%) of the total ASDs cases (Schaaf & Zoghbi, 2011).

**Table 1.3 Monogenic syndromes associated with ASD** *(table based on Caglayan et al 2010 and Sztainberg and Zoghbi 2016)*

<table>
<thead>
<tr>
<th>Syndrome</th>
<th>Genetic mutation</th>
<th>Main phenotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fragile X syndrome</td>
<td>Silencing of FMR1 gene by CGG triplet expansion</td>
<td>Large head, protruding ears, macro-orchidism, aversion to social interaction, developmental delay, ID, ASD. FXS accounts for 1-3% of total autism cases</td>
</tr>
<tr>
<td>Rett syndrome</td>
<td>LOF mutation of MECP2</td>
<td>Microcephaly, cognitive and motor impairment, epilepsy, ASD</td>
</tr>
<tr>
<td>Tuberous sclerosis complex</td>
<td>LOF mutation of TSC1 or TSC2</td>
<td>Tumours in multiple organs, ID, ADHD, epilepsy, ASD</td>
</tr>
</tbody>
</table>
In conclusion, SZ and ASDs are characterised by high phenotypic variability and complex genetic architecture, which includes common, rare and de novo risk variants. For this reason, the identification of the genetic causes underlying these diseases remains a challenge. Nevertheless, recent advances in genetic technologies have provided a long list of genetic variants associated with such psychiatric disorders. Research into the genes affected by such mutations and the biological pathways on which they converge can provide important insight into the mechanism underlying psychiatric disorders.

### 1.1.2 The Ch15q11.2 locus

CNVs in the long arm of chromosome 15 are estimated to be present at a frequency above 0.3% in the general population (Grozeva et al., 2012). Several low copy DNA repeat clusters (also called duplicons), containing pseudogenes, are present in this area. Misalignment of these sequences
during meiosis can be responsible for non-allelic homologous recombination and, consequently, the formation of chromosomal abnormalities (Locke et al., 2004). Five breakpoints (BPs) have been characterised in the proximal region of the long arm of chromosome 15. Deletions and duplications in this chromosomal region can involve different combinations of these BPs (Butler, 2017).

Prader-Willi syndrome (PWS) and Angelman syndrome (AS) are two genetic imprinting disorders. They are characterised by neurodevelopmental phenotypes of variable severity and are caused by a deletion of the paternal copy of the 15q11-13 region for PWS, or the maternal one for AS (Butler, 2017; Nicholls & Knepper, 2001). The deletion is classified as type I when it occurs between BP1 and BP3 (6.6 Mb), or type II if it involves BP2 and BP3 (5.3 Mb) (Butler et al., 2008) (See fig. 1.1). Individuals with type I deletion have been found to have more severe intellectual and behavioural deficits, suggesting the presence of some dosage-sensitive genes in this region that play an important role in neurodevelopment (Butler et al., 2004). The region between BP1 and BP2 spans approximately 500 Mb and encompasses four non-imprinted genes, which are \textit{NIPA1} (non-imprinted in Prader-Willi and Angelman syndrome 1), \textit{NIPA2}, \textit{CYFIP1} (cytoplasmic FMRP-interacting protein 1) and \textit{TUBGP5} (tubulin gamma complex associated protein 5) (Chai et al., 2003).

\textbf{Figure 1.1 Scheme of the 15q11-13 chromosomic locus.}

Genes located between BPs 1 and 5 are shown in green when non-imprinted, genes expressed only from paternal allele (PWS critical region) are in blue and genes expressed only from maternal allele (AS critical region) in red. This image is taken from Butler’s review on 15q11.2 microdeletion syndrome (Butler 2017).
**NIPA1** encodes for a magnesium (Mg++) transporter that is widely and highly expressed in the developing and adult brain (Goytain et al., 2007; van der Zwaag et al., 2010). Mutations in this gene are known to cause hereditary spastic paraplegia (Fink, 2003).

**NIPA2** also encodes a highly selective Mg++ transporter that is expressed in several organs including the brain, but is particularly abundant in placenta and kidneys (Chai et al., 2003; Goytain et al., 2008). Several studies have implicated a role for NIPA2 mutations in the aetiology of childhood absence epilepsy (Jiang et al., 2012, 2014; Xie et al., 2014).

**TUBGCP5** is part of the gamma-tubulin complex required for microtubules nucleation at the centrosome (Murphy et al., 2001). Its expression can be detected in many tissues of both mouse and human. In the brain, it is particularly high in subthalamic nuclei, a region important in the neurobiology of ADHD and obsessive-compulsive disorder (OCD) (Brem et al., 2014; Chai et al., 2003; Doornbos et al., 2009).

**CYFIP1** is a cytoplasmic protein which is ubiquitously expressed in many organs and participates in two multi-protein complexes (Chai et al., 2003). It binds to FMRP to regulate mRNA translation and is part of the WAVE regulatory complex (WRC) (Abekhoukh & Bardoni, 2014; De Rubeis & Bagni, 2010; Napoli et al., 2008). Its association with FMRP, the protein mutated in the FXS (see paragraph 1.1), makes CYFIP1 a highly likely candidate to explain the psychiatric phenotypes linked to 15q11.2 chromosomal abnormalities. Its expression and biological function have been the object of numerous studies, which will be reviewed later in this chapter.

Besides modulating the severity of PWS and AS phenotypes, CNVs involving the region between BP1 and BP2 of chromosome 15 are also known to increase risk for neurodevelopmental and psychiatric conditions. The first study reporting the association between neurological disorders and 15q11.2 microdeletion was published in 2007 (Murthy et al., 2007). Since then, many additional cases have been reported. The clinical phenotypes most frequently present in 15q11.2 deletion carriers include developmental delay and psychiatric disorders, including ASD and SZ. Other neurological problems, such as seizures, have also been described (Abdelmoity et al., 2012; Burnside et al., 2011; Cafferkey et al., 2014; Doornbos et al., 2009; Jerkovich & Butler, 2014; Madrigal et al., 2012; Picinelli et al., 2016; Vanlerberghe et al., 2015; von der Lippe et al., 2011).

Clinical reports of 15q11.2 microduplications are less abundant in the literature. Patients usually present with developmental delay and autistic features. In addition, dysmorphic features, motor coordination issues and seizures were also reported in a few cases (Burnside et al., 2011; Picinelli et al., 2016; van der Zwaag et al., 2010).
The association between 15q11.2 CNVs and psychiatric disorders has been further strengthened by the results of several GWAS. In fact, 15q11.2 microdeletion or duplications are reportedly found at a frequency of 0.55-0.82% in SZ cases and of above 0.8% in ASD ones, while only 0.2-0.4% control individuals have been found to carry such mutations (Kirov et al., 2009, 2014a; Rees et al., 2014; Stefansson et al., 2008). Data on inheritability are scarce for 15q11.2 duplications, but, in the context of deletions, it has been reported that most patients inherit the mutation from an unaffected parent (51%), while 5 to 22% of cases represent de novo CNVs (Butler, 2017; Cafferkey et al., 2014).

The frequency of 15q11.2 deletions and duplications observed in control populations provides proof for incomplete penetrance and it has been proposed that other mutations in interacting genes or pathways may play an important role in determining the clinical phenotype (Butler, 2017). Figures on penetrance are variable and incomplete, however two studies report similar values for the risk of SZ and ASD combined, associated with the microdeletion (10 and 13%) (Kirov et al., 2014a; Rosenfeld et al., 2013). Only one study investigates the effect of microdeletion and microduplication on ASD alone, reporting an estimated penetrance of 1.3% and 1.8% respectively (Chaste et al., 2014). Despite the incomplete penetrance for psychiatric disorders, 15q11.2 deletions have been recently shown to also impact the brain structure and IQ of control individuals, who often present with dyslexia and/or dyscalculia (Stefansson et al., 2014).

Considering all the data presented above, it can be concluded that some of the genes located between BP1 and BP2 on chromosome 15 have an important role in brain development and neuronal function and that an incorrect expression level of such genes significantly increases the risk of developing ASD or SZ. The next section reviews the existing literature on CYFIP1, which is considered, from a neurodevelopmental perspective, the most promising candidate gene encoded in the 15q11.2 chromosomal locus.

### 1.2 CYFIP1 biology

CYFIP1 (also known as p140Sra-1) is an evolutionarily conserved gene that was named as “cytoplasmic FMRP-interacting protein 1” by Schenck and colleagues in 2001 (Schenck et al., 2001). In this study, the authors employed a yeast two-hybrid system to screen a mouse embryonic library for proteins binding to the N-terminus of FMRP. This led to the characterisation of two proteins, named CYFIP1 and CYFIP2. The latter is also able to bind the FMRP-related proteins FXR1P/2P, whereas CYFIP1 interacts only with FMRP. The two CYFIP proteins share ~88% homology (Schenck et al., 2001).
However, prior to this study, CYFIP1 had already been identified as a target of Rac1 small GTPase, named p140Sra-1 (Kobayashi et al., 1998), and CYFIP2 was already known as PIRI121 (Saller et al., 1999). Around the time they were established as FMRP binding partners, both CYFIP1/Sra-1 and CYFIP2/PIRI121 were also characterised as being associated with WAVE (WASP (Wiskott–Aldrich syndrome protein)-family verprolin homologous protein) proteins in the formation of the WAVE regulatory complex (WRC), which controls cytoskeleton dynamics (Eden et al., 2002; Steffen et al., 2004). The interest on CYFIP1 as a potential risk gene for psychiatric disorders is therefore explained by its participation to these two fundamental cellular functions, the regulation of mRNA translation with the FMRP and the control of actin polymerisation as part of the WRC.

1.2.1 Role of Cyfip1 in the WRC

The WRC is a hetero-pentameric complex, which controls actin polymerisation by the Arp2/3, a process fundamental for cell adhesion, cell migration, vesicle trafficking and neurite extension (Takenawa & Suetsugu, 2007). Its components include WAVE1/2/3, ABI (Ableson interacting protein), NAP (NCK-associated proteins), CYFIP1/2 and HSPC300 (hematopoietic stem progenitor cell 300) (Chen et al., 2010).

WAVE proteins are characterised by the presence of a VCA (Verprolin-homology, Central and Acidic regions) domain, which binds to Arp2/3 and G-actin monomers to guide actin polymerisation (Mendoza, 2013). The resolution of the WRC structure suggested a possible mechanism for the regulation of this multi-protein complex. It was proposed that the WRC is constitutively inactive, with the VCA occupied by CYFIP1/2, preventing its interaction with Arp2/3. Several processes requiring cytoskeletal rearrangements, such as the extension of lamellipodia, stimulate Rac1 activation, which can induce CYFIP1/2 release from the VCA (Chen et al., 2010). Cooperatively with Rac1, negatively charged phospholipids also participate in WRC activation, recruiting the complex to the cell membrane. Such a step is necessary to initiate the formation of cell protrusions or lamellipodia and can also facilitate WRC oligomerisation, which significantly enhances its activity (Oikawa et al., 2004; Padrick et al., 2008).
1.2.2 Role of Cyfip1 in the regulation of mRNA translation

Together with their role within the WAVE complex, CYFIP1/2 are known to interact with FMRP at the same domain this protein uses for dimerization with other FXR proteins (Schenck et al., 2001). This suggests that CYFIP1/2 have the ability to modify FMRP function, but the consequences of this interaction have not been entirely elucidated.

FMRP is a ubiquitously expressed mRNA-binding protein which is particularly enriched in neuronal dendrites. It is believed that FMRP acts primarily as a translational repressor, but it is also involved in the regulation of localisation and stability of mRNA (Fernández, Rajan, & Bagni, 2013). Several studies conducted in neuronal cells identified over a thousand putative FMRP mRNA targets (Brown et al., 2001; Chen et al., 2003; Darnell et al., 2011). Importantly, a significant proportion of these overlaps with ASD risk genes reported by GWAS studies, while only a smaller fraction is represented by mood disorders- and SZ-related genes (Fernández et al., 2013). This evidence further strengthens the link between FMRP and ASD and provides important support to the hypothesis of an existing shared mechanism across psychiatric disorders.

In the context of mRNA translation, CYFIP1 has been proposed to be part of a complex that binds and controls mRNA translation, interacting with both FMRP and the initiation factor 4E, eIF4E. According to this model, CYFIP1 is recruited by FMRP onto specific mRNAs to act as a 4E-BP (4E-binding protein), blocking the initiation of translation (Napoli et al., 2008; Richter & Sonenberg, 2005). In Cap-dependent translation, the start of this process requires the association between the initiation factors eIF4A-eIF4E-eIF4G (named eIF4F as a complex) and the 5’m7G cap of mRNA.
molecules. 4E-BPs interfere with eIF4E-eIF4G interaction, keeping the translation inhibited (Richter & Sonenberg, 2005).

**Figure 1.3 Scheme of the proposed Cap-dependent translation mechanism involving CYFIP1.**
The translation of mRNAs is inhibited by the FMRP/CYFIP1 complex, which prevents binding of the eIF4E to the PABP. Synaptic stimulation requires mRNA translation at the spine level and, therefore, induces the release of FMRP and CYFIP1 from the translation complex. At the same time, the association between eIF4E, eIF4G and PABP allows the initiation of translation. (from Napoli et al., 2008)

In an attempt to find a single mechanism explaining CYFIP1 participation in actin polymerisation and protein synthesis, it has been suggested that cellular signals, converging on Rac1 stimulation of CYFIP1, trigger a conformational change that induces CYFIP1 release from the complex with FMRP and eIF4E and its recruitment to the WRC (DeRubeis et al., 2013). The authors designed two mutant forms of CYFIP1 to specifically inhibit its interaction with the WRC or eIF4E and transfected each isoform separately in primary neurons, where CYFIP1 had been knocked-down. Transfection of the first isoform was able to restore normal actin levels, while the second only rescued the translation of some known FMRP targets (DeRubeis et al., 2013). However, the interpretation of these results is in conflict with the numerous studies reporting CYFIP1 inhibitory role towards actin polymerisation, within the WRC (see previous paragraph). One explanation is that the mechanism regulating the distribution of the CYFIP1 protein pool between the two complexes could be more complicated than previously believed. Indeed, it has been recently shown that MNKs (Map kinase-interacting kinases) are also involved in the control of translation initiation in neuronal cells and that, by phosphorylating eIF4E, they induce its release from CYFIP1 (Genheden et al., 2015). It is therefore possible that even more proteins are involved in modulating CYFIP1
activity in the context of actin polymerisation and protein synthesis and that future studies will eventually clarify the mechanism.

1.2.3 Expression pattern of CYFIP1

There is little data on CYFIP1 expression pattern, as it has been investigated by very few studies. *In situ* hybridisation analysis by van der Zwaag et al. showed that *Cyfip1* is expressed in the mouse brain as early as E12 (van der Zwaag et al., 2010). At this stage, the expression level appears to be enriched in the ventricular zone (VZ) of the forebrain, an observation in agreement also with data from the Eurexpress database (Eurexpress.org). *Cyfip1* expression is maintained throughout embryonic development up to day E18.5, with strong signal still present in the VZ and cortical plate (CP) of the forebrain. In the adult, *Cyfip1* was found in the olfactory bulb, the hippocampus and the cerebellum.

Support that Cyfip1 protein is present in the developing cortex was recently provided by Yoon and colleague (Yoon et al., 2014). The authors showed specific localisation of Cyfip1 in the ventricular surface of the radial glia cells facing the VZ, co-localising with N-Cadherin and β-Catenin, and lower expression levels in the intermediate zone (IZ). However, in contrast to what had been previously reported by *in situ* studies, Cyfip1 protein did not appear to be expressed in the CP (van der Zwaag et al., 2010; Yoon et al., 2014).

![Figure 1.4 Cyfip1 expression in the mouse developing brain.](image)

*(A) In situ hybridisation for Cyfip1 in E14.5 mouse embryo (from Eurexpress.org). (B) Specific localisation of Cyfip1 in VZ cells of E15.5 mouse forebrain (from Yoon et al 2014).*
1.2.4 In vivo and in vitro studies on Cyfip1 alterations in neuronal cells

The known association between abnormal CYFIP1 levels and the increased risk of acquiring developmental psychiatric disorders has driven research into modelling the consequences of such alterations, using genetically modified mice and rodent or human cultured cells.

Pathania and colleagues demonstrated that Cyfip1 is essential for embryonic development, as Cyfip1−/− mouse embryos are detectable only until day E8.5 (Pathania et al., 2014). They also showed that, in primary cultures of mature mouse hippocampal neurons, Cyfip1 is enriched at excitatory synapses and that haploinsufficiency and overexpression of this gene have opposite effects on dendritic complexity. More precisely, low levels of Cyfip1 decrease dendritic branching, while high levels lead to a more complex dendritic arborisation. Spine maturation is also affected by such alterations (Pathania et al., 2014).

The disruption of normal dendritic outgrowth by CYFIP1 overexpression was also reported in human neuroblastoma cells SY5Y and in cortical cells of transgenic mice by Oguro-Ando and colleagues (Oguro-Ando et al., 2014). According to their results, high CYFIP1 dosage causes a reduction in neurite length, but also a significant increase in branch number and cell size, in both models. With the support of a microarray analysis on E15 cortex of wild type (WT) and transgenic mice, the authors proposed mTOR upregulation as the disrupted pathway at the origin of the phenotype observed. This hypothesis was confirmed treating mouse neuronal progenitors with the mTOR inhibitor rapamycin, which rescued the morphological defects caused by Cyfip1 overexpression (Oguro-Ando et al., 2014).

The disruption of physiological CYFIP1 levels has been investigated also in human PSCs-based in vitro models. A transcriptomic analysis of CYFIP1-knock down (KD) against control iPSCs-derived neural progenitor cells (NPCs) revealed that many of the differentially expressed genes were related to cytoskeletal remodelling and cell cycle (Nebel et al., 2016). In agreement with cytoskeletal dysregulation, CYFIP1KD NPCs had lower levels of WAVE protein, lower levels of F-Actin and bigger nuclei, but not significantly bigger cell size (Nebel et al., 2016).

However, CYFIP1 is only one of the four genes affected by 15q11.2 deletions and duplications (see paragraph 1.1.2). As discussed in the next paragraph, iPSCs represent a useful tool to model CNVs encompassing several genes and they have also been used in the context of 15q11.2 deletions. Yoon and colleagues used iPSCs from three deletion-carriers and five controls to demonstrate that the self-organisation of NPCs into neural rosettes, a stage mimicking the neural tube formation during in vivo embryonic development, was compromised by the deletion (Yoon et al., 2014). The expression pattern of the adhesion proteins marking the rosettes lumen was disrupted and this
defect seemed specifically due to the reduction in CYFIP1 levels. Indeed, it could be reproduced by transfecting short hairpin RNA (shRNA) targeting CYFIP1 in control cells and it could be rescued by increasing CYFIP1 expression in cells carrying the CNV. Significant reduction of WAVE2 protein was also observed in both 15q11.2-deleted and CYFIP1KD NPCs. The authors confirmed the importance of CYFIP1 in the maintenance of the correct polarity for cortical progenitors with in utero electroporation experiments, in which they demonstrated that the correct level of Cyfip1 is necessary for the proper placement of radial glial cells (RGCs) and the correct distribution of neurons across the cortical layers (Yoon et al., 2014).

In another study, Das and colleagues used iPSCs from two deletion carriers and one control to show that the mRNA levels for CYFIP1, NIPA1, NIPA2 and TUBGCP5 were reduced in pluripotent cells and neurons carrying the deletion (Das et al., 2015). In addition, they report that the morphology of mature neurons derived from the two patients suggested the presence of dendritic spine alterations, but no quantitative analysis was reported for this aspect (Das et al., 2015).

To date, only these three studies have used PSCs to investigate the significance of 15q11.2 CNVs and, in particular CYFIP1, in human neuronal cells (Das et al., 2015; Nebel et al., 2016; Yoon et al., 2014). However, all of them represent loss-of-function models. The derivation of iPSCs from duplication carriers, or the development of hPSCs-based models mimicking an increase of CYFIP1 or other 15q11.2 genes, have not been reported.

In the next paragraph, I review the use of human PSCs to model developmental psychiatric disorders. This topic was also the subject of a review published in the Brain Pathology journal, as part of a mini-symposium on “Using iPSCs to understand human neurological disease: potential and limitations” (Tamburini & Li, 2017). The content of this review is reported in the next paragraph with minor modifications.

1.3 The use of human Pluripotent Stem Cells for in vitro modelling of neurodevelopmental psychiatric disorders

Neurodevelopmental disorders are a group of conditions typically manifest early in life and are characterised by developmental deficits that produce impairments of personal, social, academic or occupational functioning. Neurodevelopmental (or neuropsychiatric) conditions include autism spectrum disorders, intellectual disabilities, schizophrenia and bipolar disorder, and have a substantial genetic component (Beneyto & Lewis, 2011; Cristino et al., 2014; O’Shea & McInnis, 2016; Ziats & Rennert, 2016).
For a long time, psychiatric research mainly relied on post-mortem studies and animal models. The first approach has the major limitation of representing the end-stage of the disease, without being informative about its origin and pathological progression. Moreover, changes observed in post-mortem tissues may be secondary effects of a patient’s prolonged use of medications. Animal models, predominantly rodents, are available for several psychiatric disorders, especially autism spectrum disorders and schizophrenia, but they have so far failed to show significant predictive validity for drug discovery (Markou et al., 2009; Pratt et al., 2012). This may be due to the inability of model organisms to represent unique higher human functions and consequently, to recapitulate all the symptoms characterising a particular disorder (Nestler & Hyman, 2010). For example, despite the numerous rodent models available for schizophrenia, these are mostly representative of the psychotic aspects of the disease, but do not reliably reproduce the cognitive and negative symptoms, such as impaired working memory, anhedonia and social withdrawal (Jones, Watson & Fone, 2011). In the case of bipolar disorder, there are no animal models available to represent both the manic and depressive extremes characterising the disorder (O’Shea & McInnis, 2016).

In just a decade from the derivation of the first line (Takahashi et al., 2007; Takahashi & Yamanaka, 2006), induced pluripotent stem cells (iPSCs) have become a fundamental tool for modelling human development and diseases, as well as for drug discovery. Like their mouse counterparts, human iPSCs are believed to have the same self-renewal and pluripotency properties of human embryonic stem cells (hESCs), but are derived (reprogrammed) from somatic cells, such as skin fibroblasts (Takahashi et al., 2007), keratinocytes (Aasen et al., 2008), dental pulp (Tamaoki et al., 2010; Yan et al., 2010) or blood (Loh et al., 2009). The reprogramming is achieved by forcing the expression of key pluripotency genes such as OCT4, SOX2, c-MYC and KLF4 in somatic cells, where the reprogramming factors start a self-regulatory loop that initiates and maintains pluripotency (Masui et al., 2007; Pan et al., 2006). Expression of these reprogramming factors can be induced via viral transduction (Ban et al., 2011; Park, 2008), transfection of polycistronic plasmids (Montserrat et al., 2011), mRNAs (Warren et al., 2010) or direct delivery of recombinant proteins (Hongyan et al., 2009). Moreover, treatment with specific combinations of small molecules has been shown to greatly increase the reprogramming efficiency (Huangfu et al., 2008; Yan Shi et al., 2008).

The use of adult somatic cells as starting material means that iPSCs are free from the ethical concerns that surround the use of hESCs, and that they can be derived from individuals carrying genetic variants that predispose to an increased risk of human diseases. Working on iPSCs, and PSCs in general, has become increasingly important in psychiatric research thanks to the ability to
study the consequences of a large number of disease-associated mutations in the phenotypically relevant cell types. This is playing an important role in advancing our understanding of the cellular mechanisms underlying psychiatric disorders and reinforcing the hypothesis of their developmental origin.

### 1.3.1 Generation of disease relevant neuronal cell types from human PSCs

The brain contains thousands of neuronal types that differ in terms of neurotransmitter identity, electrophysiological properties and afferent/efferent connectivity (Nelson, Sugino & Hempel, 2006). Different neurological diseases often exhibit pathologies specific to certain brain regions and/or cell types. This means that the possibility of observing and, therefore, being able to correct a specific phenotype, is strictly dependent on the presence of the appropriate neuronal type in the system used. Therefore, an important consideration in modelling human neurological diseases is the generation of neural cell types targeted by the disease of interest.

A number of neuronal cell types have been implicated in psychiatric disorders: cortical projection neurons (Donegan & Lodge, 2016), inhibitory interneurons (Donegan & Lodge, 2016; Lewis et al., 2012), hippocampal neurons (Frey et al., 2007), dopaminergic neurons (Donegan & Lodge, 2016) and striatal medium spiny neurons (Fuccillo, 2016; Simpson, Kellendonk & Kandel, 2010). Protocols to generate these neuronal types have been developed and are reviewed in this paragraph. Many lines of evidence demonstrate that in vitro PSC differentiation, to a large extent, mimics vertebrate development. In the context of neuronal conversion, PSCs firstly exit the pluripotent state and acquire a neuroectoderm fate. Therefore, the majority of neuronal differentiation protocols start with the induction of a specific regional neuroepithelial phenotype (region-specific progenitors) from which the target cells arise. This is mostly achieved by artificially recapitulating the signalling environment that the region-specific progenitors normally experience in vivo, by adding an appropriate combination of ‘inductive’ molecules. The aim is to induce a cascade of transcription resembling normal development, leading to the expression of a combinatorial set of transcription factors characteristic to the desired neural progenitor phenotype.

Currently, a popular method for generating neural cells from hPSCs is via monolayer differentiation protocol by dual SMAD inhibition (Chambers et al., 2009). Normal central nervous system development follows an anterior first - posterior later temporal fashion. As such, the first neuroepithelial cells generated from PSCs exhibit features of forebrain regional identity. These anterior progenitors readily mature into neurons with predominantly cortical glutamatergic identity (Boissart et al., 2013; Espuny-Camacho et al., 2013; Shi et al. , 2012). Cell types of all
cortical layers appear in a sequence reminiscent of in vivo corticogenesis, from deep layers to superficial ones. The efficiency of this approach was shown to be very high, with glutamatergic neurons accounting for 80% to nearly 100% of the cells in culture.

The specification of the other neuronal fates requires the use of additional morphogens to mimic the in vivo environment of the corresponding brain region. For example, the derivation of GABAergic interneurons requires the activation of sonic hedgehog (SHH) signalling (Y. Liu, Liu, et al., 2013; Y. Liu, Weick, et al., 2013), in some cases combined with WNT inhibition (Maroof et al., 2013; Nicholas et al., 2013), to induce the medial ganglionic eminence (MGE) fate demonstrated by the expression of transcription factor NXX2.1. However, despite the induction of a high percentage of NXX2.1+ MGE-like progenitors, efficient generation of mature interneurons, including the two major subtypes, somatostatin (SST) and parvalbumin (PV), have so far proved challenging. On the other hand, caudal ganglionic eminence (CGE)-derived interneurons expressing calretinin have been derived with a higher efficiency (>70%), by exposing late neural progenitors to activin A, a member of the TGF-β superfamily (Cambray et al., 2012).

When applied to forebrain neural progenitors at an earlier time window, activin A induces the specification of a lateral ganglionic eminence (LGE) fate, leading to the production of 40-50% of medium spiny neurons (Arber et al., 2015). Other strategies for deriving this type of neurons rely on the use of SHH, alone or in combination with WNT inhibition (X.-J. Li et al., 2009; L. Ma et al., 2012; Nicoleau et al., 2013), in a similar way to MGE protocols. Dopaminergic neurons, which are born in the ventral midbrain, can be differentiated from hPSCs by exposing the neural precursors to different combinations of WNT agonists with SHH (Denham et al., 2012; Kirkeby et al., 2012; Kriks et al., 2011) or FGF8 (Xi et al., 2012). In addition, FGF signalling blockade using a ERK/MEK inhibitor on exit of pluripotency, followed by addition of SHH and FGF8 has also been reported to induce authentic midbrain dopamine neurons (Jaeger et al., 2011). The yield of tyrosine hydroxylase positive cells is elevated in all these cases, reaching above 80%. Finally, only one protocol has been reported so far for the production of hippocampal neurons (Yu et al., 2014). For these cells, the application of WNT3a and BDNF to dorsal forebrain progenitors seems to be essential to induce the expression of PROX1, a marker for dentate gyrus hippocampal neurons.

Monolayer differentiation has proved itself a highly efficient and reliable paradigm for generating a number of neuronal types. Indeed, monolayer based forebrain glutamatergic neuron differentiation has been the platform of choice for the majority of disease modelling papers published so far (Brennand et al., 2015; Brennand et al., 2011; Chen et al., 2014; Griesi-Oliveira et al., 2014; Kim, Hysolli & Park, 2011; Marchetto et al., 2010; Robicsek et al., 2013; Shcheglovitov et al., 2013a; Srikanth et al., 2015; Wen et al., 2014b; Yoon et al., 2014). However, the monolayer
culture system has limitations in visualising potential alterations in the cytoarchitecture of the derived ‘brain tissue’. This shortcoming can be overcome, to a certain extent, by differentiating PSCs in 3 dimensional (3D) structures called organoids (Lancaster & Knoblich, 2014; Mariani et al., 2012; Qian et al., 2016), although at the expense of higher variability, both within and between organoids, and in different preparations. To date, only Mariani and colleagues have used telencephalic organoids for analysing the developmental abnormalities with iPSCs derived from patients with idiopathic autism (Mariani et al., 2015).

The generation of ‘induced neurons’ (iNs) by direct reprogramming represents an alternative to conventional PSC neural differentiation. iNs are produced by forcing the expression of a defined set of transcription factors crucial for the acquisition of neuronal fate in somatic cells or PSCs (Ambasudhan et al., 2011; W. Hu et al., 2015; Ladewig et al., 2012; Pang et al., 2011; A. S. Yoo et al., 2011). iN protocols for the generation of specific neuronal types have also been published (Caiazzo et al., 2011; Colasante et al., 2015; Pfisterer et al., 2011; Sun et al., 2016; Victor et al., 2014; Z. Xu et al., 2016). However, the iN approach may not be suitable for modelling diseases where pathogenesis occurs at neural progenitor stage, because it bypasses the process of neural progenitor specification, proliferation and differentiation choice towards distinct neuronal and glial fates. Moreover, in contrast to hPSC neural differentiation, the number of iNs that can be generated from the donor somatic cells is limited due to the restricted proliferative capacity of somatic cells prior to senescence. These shortcomings combined may explain why there are no studies published to date using iN technology to model psychiatric disorders.

1.3.2 Current status of iPSCs-based research for psychiatric disorders

Brain imaging studies have demonstrated changes in the anatomy and neuronal activity in patients suffering from psychiatric disorders, while post-mortem studies have revealed aberrant cellular morphology (Black et al., 2004; Dinstein et al., 2011; Hutsler & Zhang, 2010; Whitfield-Gabrieli et al., 2009). The iPSC technology has provided an invaluable tool to investigate the cellular basis of such alterations and elucidate the molecular pathways that may be targeted for drug discovery.

The first proof-of-principle study using iPSCs was published by Brennand and colleagues in 2011 (Brennand et al., 2011). The authors reported that neurons derived from schizophrenic patients’ iPSCs differ from those of the controls in neuronal connectivity, morphology and gene expression. This work was followed by several other schizophrenic iPSC studies reporting the emergence of earlier developmental abnormalities.
Robicsek and colleagues differentiated schizophrenia and control iPSCs into glutamatergic and dopaminergic neurons and found defects in the maturation of both cell types, with the dopamine lineage more severely affected (Robicsek et al., 2013). The authors also reported differences in mitochondrial distribution and function. This aspect of the phenotype was more pronounced in dopaminergic progenitors than the glutamatergic cells. This could suggest the presence of a higher vulnerability for dopamine cells to oxidative stress, but it could also reflect, at least in part, their extremely impaired maturation. Together with alterations of mitochondrial membrane potential and neuronal morphology, other defects often reported in neuronal cells from schizophrenic patients iPSCs are relative to WNT signalling and migration (K. Brennand et al., 2015; Topol et al., 2015) (see also Table 1).

All the studies listed above were based on heterogeneous cohorts of schizophrenic patients, selected only on the basis of their diagnosis, without knowledge of their genetic risk variants. This probably played a significant part in the high experimental variability evident in some of the results.

It is known that schizophrenia has a strong genetic component, with rare Copy Number Variations (CNVs) significantly increasing the risk of developing this disorder (reviewed by Kirov (Kirov, 2015)). Stratifying patients based on the presence of specific genetic mutations could help to reduce the degree of variability associated with iPSC work and discover new mechanisms that otherwise may be masked by the heterogeneity of the patients’ samples. Wen and colleagues took the genetics orientated approach by analysing the effect of Disrupted In SChizophrenia 1 (DISC1) mutations, which are known to co-segregate with major psychiatric illnesses (Millar et al., 2000; Wen et al., 2014b). iPSCs lines were derived from two patients carrying the same frameshift mutation in the DISC1 gene and three unaffected individuals. The authors reported altered morphology and electrophysiological properties in DISC1 neurons, as well as the expression of genes related to synaptic transmission, neural development and major mental disorders. They also established the causality between DISC1 mutations and the changes observed by repeating some of the analysis in several isogenic cell lines. These were derived by correcting the DISC1 mutation in one of the mutant lines and by introducing the same frameshift deletion present in patients in two control lines. Both synaptic and vesicular release properties were restored to normal levels in the cell line in which the DISC1 gene sequence had been corrected, while the control lines carrying DISC1 deletion recapitulated the original mutant phenotype. However, the majority of CNVs associated to psychiatric disorders contain multiple genes. This genetic feature makes iPSC modelling particularly advantageous over the generation of animal models. Microdeletion of the 15q11.2 locus has been reported by several studies as an important risk
factor for schizophrenia (Stefansson et al., 2008; The International Schizophrenia Consortium, 2008). Yoon and colleagues demonstrated that iPSCs derived from 15q11.2 deletion carriers had significant defects in neural rosette formation (Yoon et al., 2014). The 15q11.2 region contains four genes, CYFIP1 was proposed to be the likely responsible gene for the observed phenotypes. Indeed, increasing CYFIP1 expression by lentiviral transduction in differentiating 15q11.2del iPSCs rescued the abnormal expression of apical polarity markers, while reduction of CYFIP1 expression by shRNA in a control line mimicked the phenotype observed in the deleted progenitors. The authors extended their quest into CYFIP1 function by shRNA knockdown of Cyfip1 in mouse embryos via in utero electroporation, which resulted in incorrect localisation of radial glia cells and nascent neurons. However, the consequences of CYFIP1 disruption in post-mitotic neurons were not explored.

As for autism spectrum disorders, in vitro modelling was first applied to syndromic autism, such as Rett syndrome (RTT) and Phelan-McDermid syndrome (PMDS). RTT syndrome is one of the most common causes of mental retardation mainly affecting girls and is caused by mutations of the methyl CpG binding protein 2 (MECP2) (Amir et al., 1999). iPSCs-based investigations into RTT reported defects in neuronal maturation in patients’ cells (K. Y. Kim et al., 2011; M. C. N. Marchetto et al., 2010) (See also Table 1). In particular, Marchetto and colleagues demonstrated the presence of morphological alterations and reduced number of glutamatergic synapses in RTT neurons (M. C. N. Marchetto et al., 2010). This defect could be rescued by IGF1, a neurotrophic factor capable of promoting synaptogenesis. The same group subsequently reported the involvement of MeCP2 in regulating the expression of TRPC6, one of the genes disrupted by a translocation recently found in an autistic patient (Griesi-Oliveira et al., 2014).

Similar cellular phenotypes were reported in in vitro models of PMDS. Shcheglovitov and colleagues investigated the cellular phenotypes of iPSCs-derived neurons from two patients carrying heterozygous deletion of chromosome 22q13.3, the mutation responsible for PMDS (Shcheglovitov et al., 2013b). Cortical neurons derived from these iPSCs displayed impaired excitatory synaptic transmission, while the properties of their inhibitory synapses were not affected. These deficits could be either rescued by increasing the expression of SHANK3, a gene included in the deleted locus, or by IGF1 treatment, which did not affect SHANK3 levels.

Taking a different approach, Mariani and colleagues derived iPSCs from four probands with idiopathic autism, carrying no known genetic mutation previously associated with autism spectrum disorders, and unaffected family members (Mariani et al., 2015). Transcriptomic analysis of telencephalic organoids derived from patients and control cells revealed many differentially expressed genes, mainly relative to cell fate, proliferation, axonal guidance, synaptic
function and ion channels. Consistently with these results, probands’ cells showed dysregulated cell cycle and overproduction of inhibitory neurons, a phenotype that could be rescued by attenuating FOXG1 levels.

In the context of bipolar disorder, after performing a microarray analysis of iPSC-derived neurons from three patients and three controls, Chen and colleagues found a similar alteration in the expression of transcription factors regulating dorso-ventral telencephalic patterning (Chen et al., 2014). These results, however, do not significantly overlap with those from Mertens et al., the only other publication employing iPSCs for bipolar disorder modelling to date (Mertens et al., 2015). Their work, based on hippocampal neurons derived from bipolar patients, showed that the hyperexcitability and abnormal mitochondria size of these cells could be rescued by lithium treatment, while alteration of mitochondrial membrane potential could not be improved.

In the interest of space, table 1.4 summarises the above works together with additional studies not individually discussed in this mini review. It is evident that differences exist between studies within the context of the same disorder. These discrepancies may attribute to different differentiation protocols, culture systems and analysis methods. Nevertheless, some common themes emerge from these studies. In general, genes involved in nervous system development are reported to be affected by many studies and, in line with this, several phenotypes are already present at the neural progenitor stage. In particular, altered WNT signalling and mitochondria dysfunction seem to be frequently reported in patient-derived cells (K. Brennand et al., 2015; K. J. Brennand et al., 2011; Mertens et al., 2015; Robicsek et al., 2013; Srikanth et al., 2015; Topol et al., 2015; P. Wang, Lin, et al., 2015). They may represent shared mechanisms in the aetiology of neurodevelopmental psychiatric disorders.

1.3.3 CRISPR-based genome editing as a powerful alternative to model neurodevelopmental disorders

iPSCs represent a virtually unlimited and bankable source of patient-specific cells that can be differentiated into many disease-relevant neuronal types. Therefore, they allow the investigation of cellular phenotypes in cohorts of patients sharing a specific genetic mutation or the same clinical manifestation with unknown genetic background. This is particularly advantageous in the context of those disorders for which a genetic cause has not been identified, such as idiopathic autism cases (Ardhanareeswaran, Coppola & Vaccarino, 2015). However, iPSC-based studies can suffer from high variability due to differences in the genetic background of different patients, reprogramming methods and culture conditions (Sandoe & Eggan, 2013). Working with genetically modified hESCs or a well characterised reference line of hiPSCs and their isogenic
controls can represent an alternative to the use of patient iPSCs. This strategy will avoid the variability linked to the different genetic backgrounds of distinct patient iPSC lines and reduce the necessary work load of studying multiple patient iPSC lines required for identifying true phenotype and establish causality.

The CRISPR (Clusters of Regularly Interspaced Short Palindromic Repeats)/Cas9 technology allows genome editing more easily and efficiently than traditional gene targeting via Homologous Recombination (Cong et al., 2013; Ding et al., 2013; Mali et al., 2013; Ran, Hsu, Wright, & Agarwala, 2013). CRISPR/Cas9 is a type II CRISPR system (reviewed by Makarova et al., 2011), which is naturally present in many bacteria as an immunity mechanism to protect them against foreign DNA (Gasiunas et al., 2012; Jinek et al., 2012). It consists of the Cas9 nuclease and a guide RNA (gRNA), a chimeric RNA molecule combining a CRISPR RNA (crRNA) and a transactivating RNA (tracrRNA), which together direct the Cas9 to cleave the target DNA sequence. This mechanism also requires the presence of a protospacer adjacent motif (PAM) upstream of the binding region (Jinek et al., 2012). As a genome editing tool, CRISPR/Cas9 can be used to target virtually any genomic sequence next to a PAM site, by simply designing appropriate gRNAs. The generation of a double strand break (DSB) by the Cas9 induces cellular DNA repair mechanisms, like non-homologous ends joining (NHEJ), which is likely to introduce indels disrupting the targeted DNA sequence, or homology-directed repair (HDR), if a donor construct is present. This system has been optimised for its application to human cells, including hPSCs, by transfecting a single or multiple vectors to co-express the Cas9 nuclease and the gRNAs (Cho et al., 2013; Cong et al., 2013; Jinek et al., 2013; Mali et al., 2013). As an alternative to the generation of a DSB by the wild-type Cas9 enzyme, a mutated nickase version has also been developed, to facilitate HDR and reduce off-target mutations (Cong et al., 2013). Delivery of the CRISPR/Cas9 components via lentiviruses (Shalem et al., 2014; X. Wang et al., 2015), or adenoviruses (Maggio et al., 2014), has also shown a high efficiency in targeting the human genome. To further improve the flexibility and rapidity of genome editing in hPSCs, Danwei Huangfu’s group developed the iCRISPR platform, consisting of hESCs lines with doxycycline-inducible Cas9 expression (indicated as iCas9) (González et al., 2014). Transfection of iCas9 cells with gRNAs, derived via in vitro transcription, lead to over 40% efficiency for single gene targeting or around 10% for triple gene targeting.

Genome editing technologies are evolving fast and, as previously discussed, they represent a valid alternative to the use of iPSCs, because of the reduced variability associated with the use of isogenic cell lines. Alternatively, CRISPR/Cas9 could also be used to correct a specific genetic mutation in patient-derived cells. In this case, the rescue of the phenotype would provide the definitive proof of the connection between disease and genotype. When investigating the effects
of CNVs including several genes, the creation of isogenic models would be excessively time consuming, but it should at least be considered to confirm the causality link between the gene(s) of interest within a CNV and the phenotype observed in differentiated iPSCs. In conclusion, the careful planning of in vitro modelling experiments and parallel use of engineered hESCs and iPSCs, allows the generation of very elegant systems for the investigation of the cellular pathology underlying complex neurodevelopmental disorders.
<table>
<thead>
<tr>
<th>Disease</th>
<th>Genotype</th>
<th>Cellular model</th>
<th>Findings</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schizophrenia</td>
<td>Mixed</td>
<td>iPSCs from 5 patients and 6 controls Differentiation to predominantly glutamatergic neurons</td>
<td>Differences in neuronal connectivity, morphology and gene expression (glutamate, cAMP and WNT signalling, synaptic connection and long-term potentiation).</td>
<td>(Brennand et al., 2011)</td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>Mixed</td>
<td>iPSCs from 3 patients and 2 control individuals Differentiation to dopaminergic (DA) and cortical glutamatergic fate</td>
<td>Defective differentiation to DA fate (lower TH expression, reduced number and length of neurites) Impaired maturation towards cortical glutamatergic fate (lower expression of synaptic markers) Mitochondrial abnormalities</td>
<td>(Robicsek et al., 2013)</td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>Mixed</td>
<td>iPSCs from 4 affected patients and 6 controls Differentiation to neural progenitor cells (NPCs)</td>
<td>Many differences in gene and protein expression (mainly relative to synapses, cellular adhesion and oxidative stress) Aberrant migration and increased oxidative stress in SZ NPCs</td>
<td>(Brennand et al., 2015)</td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>Mixed</td>
<td>iPSCs from 4 affected patients and 6 controls Differentiation to NPCs</td>
<td>Increased canonical WNT signalling in SZ cells No differences in migration</td>
<td>(Topol et al., 2015)</td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>Frameshift deletion in DISC1 gene</td>
<td>iPSCs from 2 deletion carriers and 3 controls Differentiation to forebrain neurons</td>
<td>Morphological, electrophysiological and synaptic defects in DISC1-mutant neurons Phenotype rescued after correction of DISC1 mutation in patients’ cells Phenotype mimicked in control cells after introduction of DISC1 mutation</td>
<td>(Wen et al., 2014a)</td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>15q11.2 deletion</td>
<td>iPSCs from 3 deletion carriers and 3 control individuals Differentiation to NPCs</td>
<td>Disrupted expression of apical adhesion proteins and formation of neural rosettes Phenotype rescued by lentiviral transduction to increase CYFIP1 expression</td>
<td>(Yoon et al., 2014)</td>
</tr>
<tr>
<td>Disorder</td>
<td>Chromosome Deletion</td>
<td>iPSCs Details</td>
<td>Summary Details</td>
<td>Reference</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>---------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>22q11.2 deletion</td>
<td>iPSCs from 6 deletion carriers and 6 control individuals</td>
<td>Reduction of CYFIP1 level by shRNA recapitulated the phenotype</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Embryonic bodies (EBs)-based differentiation to a mixed population of excitatory and inhibitory neurons</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Altered expression of miRNA regulating genes associated to psychiatric disorders</td>
<td></td>
<td>(Zhao et al., 2015)</td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>22q11.2 deletion</td>
<td>IPSCs from 8 deletion carriers and 7 controls</td>
<td>Differentially expressed genes linked to cell cycle, cell survival and apoptosis and MAPK pathway</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Differentiation to a mixed population of excitatory and inhibitory neurons</td>
<td></td>
<td>(Lin et al., 2016)</td>
</tr>
<tr>
<td>Syndromic Autism (Rett syndrome)</td>
<td>MeCP2 mutations</td>
<td>iPSCs from 4 RTT patients and 6 healthy controls</td>
<td>Reduced number of glutamatergic synapses, fewer spines and smaller soma in RTT neurons Synaptic defects rescued by IGF1 treatment</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Embryoid bodies (EBs)-based differentiation to a mixed population of excitatory and inhibitory neurons</td>
<td></td>
<td>(Marchetto et al., 2010)</td>
</tr>
<tr>
<td>Syndromic Autism (Rett syndrome)</td>
<td>MeCP2 mutations</td>
<td>iPSCs from 5 RTT patients and 4 WT PSCs lines (3 iPSCs and HUES1)</td>
<td>Impaired neuronal maturation of RTT iPSCs (lower TuJ and sodium channels expression)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Kim et al., 2011)</td>
</tr>
<tr>
<td>Syndromic Autism (Phelan-McDermid syndrome)</td>
<td>22q11.3 deletion</td>
<td>iPSCs from 2 patients and 3 control PSCs lines (2 iPSCs and HUES9)</td>
<td>Altered excitatory post-synaptic transmission (reduced amplitude and frequency of excitatory post-synaptic currents, EPSCs) Inhibitory synaptic transmission unaffected Synaptic deficits restored to control levels by lentiviral transduction increasing SHANK3 expression, or IGF1 treatment</td>
<td>(Shcheglovitov et al., 2013a)</td>
</tr>
<tr>
<td>Syndrome</td>
<td>Description</td>
<td>iPSCs Details</td>
<td>Findings</td>
<td>References</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>Syndromic Autism 15q11-13 duplication 15q11-13 deletion (Angelman syndrome, AS)</td>
<td>4 duplication carriers, 1 AS patient and 1 control Embryonic bodies (EBs)-based differentiation to a mixed population of excitatory and inhibitory neurons</td>
<td>Altered level of transcripts related to neuronal differentiation, cell cycle and protein catabolism Several autism and epilepsy candidate genes amongst the differentially expressed genes</td>
<td>(Germain et al., 2014)</td>
<td></td>
</tr>
<tr>
<td>Syndromic Autism (Fragile X Syndrome)</td>
<td>FMR1 gene silencing caused by CGG expansion</td>
<td>iPSCs from 3 patients and 2 hESCs control lines (H9 and HUES13) Cortical glutamatergic neurons differentiation</td>
<td>Morphological abnormalities Reduced ability to fire trains of action potentials Abnormal synaptogenesis</td>
<td>(Telias et al., 2015)</td>
</tr>
<tr>
<td>Autism Spectrum Disorder</td>
<td>Translocation in 3p21 and 11q22 chromosomes (Disruption of VPRBP and TRPC6 genes respectively)</td>
<td>iPSCs from one patient and 2 control lines (one iPSCs and 1 HUES6) Differentiation to forebrain NPCs and neurons</td>
<td>Differentially expressed genes relative to nervous system development and function (several CREB-target genes) Alterations of Ca++ influx, morphology and synapses Normal phenotype restored by lentiviral-mediated WT TRPC6 expression or hyperforin (TRPC6 agonist) treatment Involvement of MeCP2 in TRPC6 regulation</td>
<td>(Griesi-Oliveira et al., 2014)</td>
</tr>
<tr>
<td>Autism Spectrum Disorder</td>
<td>Mixed</td>
<td>iPSCs from 4 patients and 8 unaffected first-degree family members Differentiation to cerebral organoids</td>
<td>Altered expression of genes involved in cell fate, proliferation, axonal guidance, synaptic function and ion channels Faster proliferation of ASD iPSCs and NPCs Overproduction of GABAergic neurons Phenotype rescued by reduction of FOXG1 level via shRNA</td>
<td>(Mariani et al., 2015)</td>
</tr>
<tr>
<td>Autism Spectrum Disorder</td>
<td>Mixed</td>
<td>iPSCs from 8 patients with early brain overgrowth and 5 controls Differentiation to forebrain excitatory and inhibitory neurons</td>
<td>Faster proliferation (altered β-catenin/BRN2 signaling) Reduced synaptogenesis and network activity</td>
<td>(Marchetto et al., 2016)</td>
</tr>
</tbody>
</table>
| Bipolar Disorder | Mixed | iPSCs from 3 patients and 3 controls  
Differentiation to forebrain neurons | Dysregulated expression of several transcripts regulating dorso-ventral telencephalic patterning, calcium signalling and miRNAs processing | (Chen et al., 2014) |
|------------------|-------|---------------------------------|-------------------------------------------------|-------------------|
| Bipolar Disorder | Mixed | iPSCs from 6 patients and 4 controls  
Differentiation to hippocampal dentate gyrus neurons | Upregulation of mitochondrial genes, higher mitochondrial membrane potential and smaller mitochondria size in BD neurons  
Hyperexcitability of BD neurons (higher Na+ activation, increased action potential firing and amplitude)  
Mitochondrial and electrophysiological changes rescued by lithium (Li) treatment in neurons derived from Li-responsive patients | (Mertens et al., 2015) |
1.4 Cortex development

The cerebral cortex is considered to be the most complex structure of the human brain, which regulates its highest functions, such as consciousness and perception of reality, making it particularly relevant in the context of psychiatric disorders (Frith & Dolan, 1996). For this reason and because differentiation protocols for PSCs towards this fate are well characterised and reproducible, this approach is the most used in in vitro models of psychiatric disorders (see paragraph 1.3). The correct analysis and interpretation of the cellular phenotypes observed during the application of such protocols depends on a deep understanding of the corresponding in vivo process.

Most of our knowledge about in embryonic cortical development derives from mouse studies, thanks to the availability of tissue and the ease of deriving mutant animals to study the function of specific genes. The following paragraphs describe embryonic cortex development primarily in the mouse model. Human data are included where important differences between the two models need to be highlighted.

1.4.1 Early corticogenesis: progenitor subtypes and characteristics

The most abundant class of neurons found in the cerebral cortex is represented by excitatory pyramidal neurons (~85%), which originate from the embryonic dorsal telencephalon, the rostral-most region of the neural tube. GABAergic interneurons (INs), which migrate into the cortex from the MGE, in the ventral telencephalon, constitute the remaining 15-20% (Marín & Müller, 2014).

The term corticogenesis is mainly used to indicate the generation of pyramidal neurons from progenitors located in the VZ and subventricular zone (SVZ), which together represent the proliferative area adjacent to the ventricles in the dorsal telencephalon (Fig. 5) (Tiberi, Vanderhaeghen & van den Ameele, 2012). The tissue forming the wall of the neural tube, where neurogenesis occurs, is characterised by an intrinsic polarity, with the side facing the ventricle defined as apical and the one in contact with the basal lamina (or pial surface) defined as the basal side (Fig. 5) (Taverna, Götz & Huttner, 2014).

The first type of progenitors to arise early in development are neuroepithelial cells (NEs). These are bipolar cells that contact both apical and basal surfaces. The apical side is marked by the presence of tight junctions and a primary cilium, which detect signals present in the cerebrospinal fluid (CSF) filling the ventricle (Laguesse, Peyre & Nguyen, 2014). Signalling molecules in the CSF include
Fibroblast Growth Factors (FGF), Insulin-like Growth Factor (IGF), Sonic Hedgehog (SHH), Bone Morphogenetic Proteins (BMPs), Retinoic Acid (RA) and WNT (Taverna et al., 2014).

NEs undergo a first phase of symmetric divisions to amplify the initial progenitor pool and later, with the onset of neurogenesis (~E12), switch to asymmetric divisions. The proliferation of this progenitor pool is characterised by a mechanism defined as Interkinetic Nuclear Migration (INM), which indicates the migration of the cell soma between the apical and basal side of the VZ at different phases of the cell cycle. Briefly, cell division during the M phase happens in contact with the apical surface, while other phases happen during migration towards the basal side (G1), while the nucleus is localised at the basal side of the VZ (S phase) or during migration from the basal to the apical side (G2). This mechanism, which allows all NEs to remain in contact with the apical surface, and to simultaneously migrate away from it at different times, leads to formation of what is defined as a pseudostratified epithelium (Fig 51a) (Taverna & Huttner, 2010).

With the switch to asymmetric division, NEs undergo several changes, resulting in transformation into radial glia cells (RGCs). RGCs maintain some similarities to NEs, such as the extension of a primary cilium into the ventricle and the mechanism of INM (Laguesse et al., 2014). However, important differences include the loss of tight junctions, which are substituted by adherent junctions, and the expression of a different set of markers, including Paired box 6 (Pax6), Brain Lipid-Binding Protein (BLBP), Vimentin (VIM) and Nestin (Feng et al., 1994; Götz et al., 1998; Laguessed et al., 2014).

The adherent junctions, found at the apical surface, are marked by the presence of N-Cadherin (NCAD or CDH2) and Zona-Occludens 1 (ZO-1). Cadherins are transmembrane proteins which carry out, with their extracellular domain, calcium-dependent homophilic interactions, while, at the intracellular level, they recruit complexes of α/β-catenin. Catenin proteins mediate the interaction between Cadherins and the actin cytoskeleton, which is represented by a “belt” of F-Actin lining the apical surface of these cells (Suzuki & Takeichi, 2008; Thumkeo et al., 2011).

The apical domain of RGCs is a fundamental signalling centre in the regulation of the balance between proliferation and differentiation of the cells occupying the VZ. The key signalling pathways active in this area include Wnt/β-Catenin, Notch, Fgf and Shh, all supporting the proliferation of the progenitor pool (Tiberi et al., 2012). Perturbation of the elements located at the apical surface of RGCs has deleterious consequences on the balance between self-renewal and differentiation of this class of progenitors. For instance, the interaction between Ncad and β-catenin contributes to regulating the pool of free β-catenin that can migrate into the nucleus and activate transcription (Taverna et al., 2014). It has been demonstrated that in vivo knock-down of Ncad in mouse cortical progenitors at E13.5, causes premature differentiation and migration away from the VZ (Zhang et
In addition to this, overexpression of a constitutively active form of β-catenin in mouse neural precursors causes abnormally enlarged brains due to the over-proliferation of neuronal progenitors (Chenn & Walsh, 2002). However, a similar hyperactivation of the WNT pathway in basal progenitors has been reported to promote neuronal differentiation by activating the proneuronal gene Neurogenin 2 (NGN2), suggesting that most of these mechanisms are stage-specific (Hirabayashi et al., 2004).

Other proteins recruited to the apical domain and involved in signalling mechanisms include the atypical Protein Kinase C (aPKC) and Par complexes (Costa et al., 2007; Insolera, Chen & Shi, 2011). In particular, the level of Par3 in daughter cells regulates asymmetric division through a Notch-mediated mechanism (Bultje et al., 2009). Because of the connection between apical adhesion molecules and F-actin fibres, cytoskeleton dynamics are also involved in regulating RGCs behaviour. Indeed, conditional KO mouse models for RhoA and cdc42, two GTPase proteins participating in the control of actin polymerisation, showed alterations in RGC self-renewal (Cappello et al., 2006, 2012).

Asymmetric divisions of RGCs give rise to intermediate progenitors (IPs), another class of cortical progenitor cells, which are localised in the SVZ, and are characterised by a multipolar morphology and expression of the transcription factor Tbr2 (Englund, 2005). In mouse, these cells divide one or two times before undergoing the last terminal neurogenic division (Noctor et al., 2004). In humans and other gyrencephalic species, IPs can undergo more symmetric divisions, contributing to the expansion of the progenitor pool (Betizeau et al., 2013). The primate SVZ is significantly thicker and can be subdivided into “inner” SVZ (ISVZ), containing TBR2+ cells, and “outer” SVZ, containing another type of progenitors defined as outer radial glia cells (oRGCs). Such cells are derived from apical RGCs by asymmetric division and are highly proliferative, generating more oRGCs, IPs or neurons directly (Betizeau et al., 2013). ORGCs possess a basal, but not an apical, fibre and express some pan-radial glia genes, such as VIM and HES1, and some unique markers, such as ITGB5, HOPX and TNC, which were recently identified. Moreover, the proliferation of this class of progenitors seems to be regulated by LIF/STAT signalling (Pollen et al., 2015). The expansion of the oSVZ is considered the most important event in the evolution of the neocortex (Dehay, Kennedy & Kosik, 2015).
**Figure 1.5 Scheme of cortical neurogenesis.**
(a) Interkinetic Nuclear Migration (INM) in ventricular zone (VZ) cells. (b) Neurogenic steps and progenitors subtypes in the mouse developing cortex. (c) Diversity of cortical progenitors in the human foetal cortex at approximately 15 weeks post-conception (PCW) (From Jiang and Nardelli 2016).

### 1.4.2 Formation of the cortical layers

Before the generation of neurons that will occupy the cortical plate (CP, layers I to V), Reelin+ Cajal-Retzius cells appear close to the pial surface, in the so-called marginal zone (MZ). These early-born neurons are primarily generated in the cortical hem and septum and, in humans, can also originate from NE cells. A clear population of MZ cells is already present at gestational week (GW) 9 in the human developing cortex (Bystron, Blakemore & Rakic, 2008). Cajal-Retzius cells secrete Reelin, a large glycoprotein, which provides fundamental signalling for the correct formation of the cortical layers (Frotscher, 1998).

Cortical neurons of layers I to V are generated either directly from RGCs in the VZ or indirectly from intermediate progenitors in the SVZ. This happens via an asymmetric division, giving rise to a progenitor still capable of self-renewal and a cell destined to terminal differentiation, or via a symmetric neurogenic division (Noctor et al., 2004). Newly-born neurons migrate to the CP by somal translocation or locomotion. The first process happens when the cell extends a basal process and the cell body can move basally within it. In the second case, the neuron develops a short leading
process in the direction of the pial surface, but not attached to it, and the whole cell moves in that direction (Nadarajah et al., 2001).

The formation of cortical layers happens in an inside-out manner, with early-born neurons populating the deep layers and late-born ones migrating past them to settle in more superficial positions. As soon as they leave the cell-cycle, the expression of specific transcription factors acts on post-mitotic neurons to specify their identity (Greig, Woodworth et al., 2013). This process is summarised in figure 1.6.

Briefly, layer VI cortico-thalamic projection neurons are characterised by the expression of T-box brain protein 1 (Tbr1), a transcription factor that was demonstrated to be fundamental in the development of this class of neurons and that acts mainly by repressing FEZ Family Zinc Finger 2 (Fezf2), which is in turn essential for specification of layer V sub-cerebral projection neurons (Chen et al., 2005; McKenna et al., 2011). A typical marker of layer V neurons, named COUP-TF-interaction protein 2 (Ctip2) acts downstream of Fezf2 to regulate the extension of axonal projections to subcortical targets and determines in this way the acquisition of the correct fate (Chen et al., 2008).

Layers II and III are primarily formed by neurons that project to other cortical structures, including callosal projection neurons, the axons of which extend across the corpus callosum into the contralateral hemisphere. The expression of the special AT-rich sequence binding protein 2, Satb2, which represses Ctip2, is essential for specifying the identity of this class of neuronal cells and for the development of their axonal projections towards the appropriate targets (Alcamo et al., 2008; Britanova et al., 2008). Satb2 is also weakly expressed by layer IV granular neurons, which project mostly locally (Alcamo et al., 2008; Britanova et al., 2008). They have been shown to derive from a population of SVZ progenitors labelled by Svet1 (Subventricular tag 1), and to be characterised also by the expression of the two homologues of the Drosophila homeobox Cut gene (Cux1/2) (Nieto et al., 2004; Tarabykin et al., 2001). Cux1/2 are specifically found in all upper-layer neurons, including Satb2+ layer II/III (Nieto et al., 2004).

The specification of the various projection neuron subtypes happens in parallel to other processes, such as the specialisation of the different cortical areas and the generation and migration of inhibitory interneurons. Although very important in the context of cortex development, these events are not essential for the understanding and interpretation of the results presented in this thesis and, therefore, they are not reviewed in this chapter.
Figure 1.6 Schematic representation of cortical layer formation during mouse corticogenesis.

Neurons of all six cortical layers are generated sequentially, starting from E11.5. Cajal-Retzius (CR) cells are the first ones to appear and populate the top layer of the developing cortex. Subplate neurons (SPN) represent the first group of neurons to be generated from radial glia (RG) cells. They are followed by cortico-thalamic projection neurons (CThPN), subcortical projection neurons (SCPN), granular neurons (GN) and callosal projection neurons (CPN), in this order. At the end of corticogenesis, the remaining progenitors switch to the generation of astrocytes and oligodendrocytes (from Greig, Woodworth et al 2013).
1.5 Aims of the project

Deletions and duplications of the 15q11.2 locus are known risk factors of psychiatric disorders and CYFIP1 is one of the four genes encoded in this region. In neuronal cells, CYFIP1 contributes to the regulation of actin polymerisation and mRNA translation in association with FMRP, the protein which mutation causes FXS. Because of this, CYFIP1 is considered the strongest candidate to be responsible for the development of psychiatric phenotypes in 15q11.2 CNVs carriers.

The overall aim of this thesis is to investigate how altered levels of CYFIP1 affect cortical development of hPSCs, an aspect that may underpin its contribution to the phenotypes associated with 15q11.2 CNVs. To achieve this, I aim to generate hESCs lines with increased CYFIP1 expression and derive CYFIP1-knock out stem cells using CRISPR/Cas9. The parallel characterisation of these cell lines during cortical differentiation protocols will allow to uncover the mechanisms in which CYFIP1 is involved and, with the help of RNA-sequencing, those biological pathways that are specifically altered by changes in the expression of this gene. Manipulation of these pathways will be used to try and reverse the phenotypes associated to CYFIP1 alterations.

Together with the derivation of hESCs with modified CYFIP1 levels, I will also derive iPSCs from some 15q11.2 CNVs carriers recruited during the time of my project. Such iPSCs lines will be used to characterise the expression of the genes encoded in this locus during human cortical differentiation. This would help to understand if, together with CYFIP1, any of the other 15q11.2 genes may be involved in the regulation of brain development.
2 Methods and materials

2.1 Cell culture

2.1.1 hESC culture

The cell lines used in this study were H7 (wicell.org), iCas9 (González et al., 2014), their engineered derivatives, CYFIP1tg and CYFIP1ko, as well as iPSC lines derived from 15q11.2 deletion carriers. Human ESCs were maintained in mTesR1 or TesR-E8 media (STEMCELL technologies) under standard culture conditions (37°C, 5% CO2) in 6 well-plates coated with Matrigel® (Corning, VWR). mTesR1 was used during critical steps, such as sub-cloning of genetically modified cells, or recovery after thawing or transfection.

The cells were passaged every 3-4 days, when 70-90% confluent. Briefly, stem cells were washed once with DPBS and incubated in 0.02% EDTA (Sigma) for about 3 minutes at 37°C. After removing EDTA, cells were dissociated into small clumps in fresh medium and were then seeded into a new plate. Usually, H7 and H9-iCas9 were split at ratios between 1:3 and 1:5.

For freezing, approximately 1-2x10⁶ cells were dissociated with EDTA as described above, centrifuged at 200g for 5 minutes and resuspended in 1ml of cold hESC medium with 10% DMSO (Sigma). Cryo-vials containing the cell suspension were transferred in an appropriate freezing container, placed at -80°C, to achieve 1°C/min cooling rate. Once frozen, cryo-vials were transferred to nitrogen tanks.

For thawing, cryo-vials were placed in a water bath at 37°C, to make this process as quick as possible. The thawed cell suspension was then transferred to 10 ml of pre-warmed hESC medium and centrifuged at 200g for 5 minutes. Finally, the cell pellet was resuspended in hESC medium and plated.

2.1.2 Monolayer differentiation into cortical glutamatergic neurons

The cortical pyramidal neurons differentiation protocol used for the experiments presented in this thesis is an adapted version of previously published ones (Espuny-Camacho et al., 2013; Kirwan et al., 2015; Yichen Shi et al., 2012). An outline of the procedure is shown in Figure 2.1.
Figure 2.1 Monolayer differentiation of hESCs into cortical pyramidal neurons. The timing for both parental lines is shown in grey, while the substrate and the type of media used at each stage are shown in blue and purple, respectively.

Initially, stem cells were plated on Growth Factor Reduced Matrigel (Corning, VWR) and kept in hESCs medium until 80-90% confluent. The day when hESCs medium was switched to neural induction medium was set as day 0. The composition of all media used in differentiation experiments are listed in Table 2.1. The timing of neuronal induction and maturation for H7 and iCas9 cells is remarkably different. Therefore, based on previous work of our group, the differentiation protocol was adapted accordingly. The cells were kept in neural induction medium until day 8 for the iCas9 line and day 12 for the H7, when LDN and SB were removed.

Two days before the end of neural induction, the cells were split at a 2:3 ratio on fibronectin-coated plates. Coating with fibronectin (Millipore) was performed incubating a solution of 15μg/ml at 37°C for at least 1 hour. For the splitting, the cells were first pre-treated with 100μM ROCK inhibitor (Y-27632, STEMCELL Technologies), for 1 hour to prevent cell death. Then they were incubated in EDTA for 3 minutes and manually dissociated in N2B27 medium, using a 2ml serological pipette, keeping large cell clumps. These were carefully resuspended in the appropriate amount of neural induction medium and seeded in a new plate.

After 8-10 days, the neural progenitors that form the differentiating cultures at this stage, were split again, at a ratio of 1:4, on Poly-D-Lysin/Laminin-coated plates. This coating was performed incubating first a solution of 10 μg/ml of Poly-D-Lysin (Sigma) for 30 minutes at room temperature. This was followed by 3 washes with DPBS and an overnight incubation at 37°C with a 10μg/ml laminin solution (Sigma). A few days after this second passage, when cells displayed a clear neuronal morphology, the media is switched to N2B27, in which the B27 supplement contains vitamin A, to promote maturation. Throughout the whole differentiation protocol, cells were fed every other day.
Table 2.1 composition of media used for monolayer differentiation

<table>
<thead>
<tr>
<th>Media</th>
<th>Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>N2B27 (without vitamin A)</td>
<td>2:1 DMEM-F12 and Neurobasal, 1x N2, 1x B27 without vitamin A, 2Mm Pen/Strep-Glutamine, 0.1 mM Beta-Mercapto-Ethanol (All from Thermo Fisher)</td>
</tr>
<tr>
<td>Neural induction medium</td>
<td>N2B27 without vitamin A, 10μM SB-431542 (TOCRIS), 100nM LDN-193189 (TOCRIS)</td>
</tr>
<tr>
<td>N2B27 (with vitamin A)</td>
<td>2:1 DMEM-F12 and Neurobasal, 1x N2, 1x B27 without vitamin A, 2Mm Pen/Strep-Glutamine, 0.1 mM Beta-Mercapto-Ethanol (All from Thermo Fisher)</td>
</tr>
</tbody>
</table>

2.1.3 hESC differentiation into brain organoids

Human ESCs were differentiated into brain organoids according to the protocol published by Lancaster and Knoblich (Lancaster & Knoblich, 2014). Briefly, 9000 cells per well were seeded in U-bottom ultra-low attachment 96w-plates to form embryoid bodies (EBs). EBs were fed every other day until 500-600 μm in diameter, when they were transferred to bacterial dishes in neural induction medium. After 4-5 days, the EBs showing a radially organised neuroepithelium were embedded in Matrigel droplets and cultured in differentiation medium. After a few days, when the embedded organoids started showing neuroepithelial buds, they were transferred to a spinning bioreactor in differentiation medium containing B27 with vitamin A and cultured in constant agitation until the time chosen for analysis. The composition of all media is listed in Table 2.2.

Table 2.2 composition of media used for organoids differentiation

<table>
<thead>
<tr>
<th>Media</th>
<th>Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>EBs medium</td>
<td>DMEM-F12, KOSR (1:5), FBS (1:30, Biosera), 1x Pen/Strep-Glutamine, 1x NEAA, bFGF (4ng/ml, PeproTech) and ROCK inhibitor (50mM, STEMCELL Technologies)</td>
</tr>
<tr>
<td>Neural induction medium</td>
<td>DMEM-F12, N2 supplement (1x), Pen-Strep-Glutamine (1x), NEAA (1x) and heparin (1μg/ml).</td>
</tr>
<tr>
<td>Differentiation medium</td>
<td>1:1 DMEM-F12 and Neurobasal, 0.5x N2, 0.5x B27 (with or without vitamin A), 1x Pen-Strep-Glutamine and 1x NEAA (All reagents from Thermo Fisher, unless otherwise stated)</td>
</tr>
</tbody>
</table>

2.1.4 Other chemicals used in differentiation experiments

For the experiment aimed at rescuing the neurogenesis phenotype in CYFIP1tg cells (Chapter 5.2.2), neural progenitors were treated for three days with XAV939 (1μM, TOCRIS), API2 (1μM, TOCRIS) or
an antibody blocking N-Cadherin function (ACAM GC-4, Sigma). An equivalent amount of DMSO (Sigma) was also added to the cultures as vehicle.

2.2 Fibroblast and keratinocytes reprogramming into iPSCs

The acquisition of primary tissue was performed in collaboration with the National Centre for Mental Health, with ethical approval from the Cardiff University Research Tissue Bank (11/WA/0255). Human dermal fibroblasts and keratinocytes were collected, cultured and reprogrammed by Dr. Craig Joyce. The reprogramming was carried out using the CytoTune®-iPS 2.0 Sendai Reprogramming Kit (Thermo Fisher). Briefly, somatic cells from an early passage and at about 80% of confluency, were transduced with the appropriate amount of virus, according to manufacturer instructions. A week after transduction, the cells were re-plated on Matrigel-coated plates and, 24 hours later, their culture media was replaced with stem cells media. During the following 2 weeks, the media was replaced daily and the cultures were monitored for the appearance of iPSCs colonies. When they reached the appropriate size, these were manually picked, expanded and banked as described above. This part of the work was done in conjunction with Craig Joyce (NMHRI technician).

2.3 Molecular cloning for CYFIP1-overexpression vector

For the generation of the CYFIP1-overexpression vector, the full length open reading frame (ORF) of the human CYFIP1 gene was cut out of a commercial expression vector (Origene, SC100426) and cloned downstream of the CAG promoter in a pCAG-IRES-GFP plasmid available in the lab. To this end, the Origene plasmid was digested with the restriction enzyme NotI (New England Biolab, NEB) and run on a 2% agarose gel, to isolate the DNA band corresponding to CYFIP1 (4400 pb) from the backbone. The band corresponding to the CYFIP1 ORF was then cut and purified using the ISOLATE II PCR and Gel Kit (Bioline). At the same time, NotI was used to cut also the pCAG-IRES-PAC vector at the level of the insertion site, just downstream of the CAG promoter. This cut vector was then combined with a 3-fold molar excess of insert, the purified CYFIP1 ORF, for the ligation reaction. This was performed at room temperature, following the Quick ligation Kit (NEB) protocol. An aliquot of 5 µl from the ligation reaction was used to transform chemically competent cells 5-alpha F'Iq Competent E.coli (NEB) according to manufacturer’s instructions. Transformed bacteria were plated on LB agar plates containing ampicillin and incubated overnight at 37°C. On the following day, 10-20 miniprep reactions were set up from single ampicillin-resistant colonies. These were left to grow overnight in LB media, in constant agitation, at 37°C. The next day, the plasmid DNA was
extracted using QIAprep Spin Miniprep Kit (Qiagen) and screened with BglII and PvuI (both NEB) to reveal the presence and orientation of the insert. A representative image of the digested CYFIP1-overexpression (pCAG-hCYFIP1-IRES-PAC) vector can be found in Chapter 3 (Fig. 3.1). All digestions were performed in the recommended buffers, at 37°C, for at least 1 hour.

2.4 CRISPR/Cas9 targeting: gRNA design and synthesis

Guide-RNAs (gRNAs) targeting the first exon of CYFIP1 were designed using DNA2.0, now ATUM (www.atum.bio) and the online CRISPR Design tool from the Massachusetts Institute of Technology (www.crispr.mit.edu). Three sequences were chosen based on the off-targets scores given by both online tools. The gRNAs were then generated via *in vitro* transcription. To this purpose, a single-strand DNA oligo containing a T7 promoter followed by a 20 nucleotides-long guide sequence and the tracr-RNA, was synthetized by Sigma (Fig. 2.1 A). This single-strand construct was used as template for PCR using a primer complementary to the T7 region and one complementary to the last 20 nucleotides of the tracr-RNA sequence (Fig. 2.1 A, red arrows). The resulting double stranded DNA amplicon was converted into RNA using the MEGAshortscript T7 kit for *in vitro* transcription (Thermo Fisher), to produce an RNA molecule composed by guide sequence and tracrRNA in a 5’ to 3’ order. The yield was determined by running 2 µl of the *in vitro* transcription reaction on a 2% TBE gel (Fig. 2.1 B), while the remaining product was purified using the MEGAClear Kit (Thermo Fisher), to obtain the gRNAs ready for transfection.

![Figure 2.2](image)

Figure 2.2 Generation of gRNAs.

(A) Nucleotide sequence of one of the DNA oligo used as template for *in vitro* transcription. The T7 promoter sequence is shown in green, the guide RNA in black and the tracrRNA sequence in blue. Red arrows represent the primers used for PCR amplification of the ssDNA oligo. (B) RNA gel showing the products of *in vitro* transcription with three different guide sequences.

2.5 Transfection of hESCs

2.5.1 Nucleofection

Nucleofection was used to deliver the pCAG-hCYFIP1-IRES-PAC plasmid (see paragraph 2.2) into H7 hESCs, to derive CYFIP1-Overexpressing (CYFIP1tg) cells (see also Chapter 3). For the transfection,
the Lonza 4D-Nucleofector® and the P3 Primary Cell 4D-Nucleofector® X Kit were used, according to manufacturer’s instructions. For one nucleofection, about $2 \times 10^6$ H7 hESCs were dissociated with Gentle Cell Dissociation Reagent (STEMCELL Technologies) for 10 minutes and spun down in stem cells medium. The cell pellet was resuspended in 100 µl of P3 buffer together with 2 µg of plasmid DNA and placed into an electroporation cuvette. The CB-150 program, pre-installed in the 4D-Nucleofector, was applied to deliver the plasmid. Immediately after this step, the cells were replated in hESCs medium, supplemented with 100μM Y-27632, and left to recover for 2 days. After this time, puromycin selection ($1 \mu g/ml$) was applied for ten days, until a few colonies of cells with stable expression of the transgene remained and grew to a suitable size to be manually picked.

### 2.5.2 RNA transfection

RNA transfection was carried out to deliver gRNAs into iCas9 hESCs for the derivation of the CYFIP1ko line (see also Chapter 3). The Lipofectamine® RNAiMAX transfection reagent (Thermo Fisher) was used for this purpose. Following the product protocol, 1µl of Lipofectamine and the appropriate amount of RNA were diluted into two separate aliquots of stem cell medium (250µl each, for one well of a 6 well-plate). These were then combined and incubated at room temperature for 20 minutes, to allow the formation of RNA-Lipofectamine complexes. After this time, such complexes were added dropwise to two wells of ~50% confluent iCas9 cells. The final concentration of RNA in the well was of 10 nM for the fluorescent Alexa555-Block iT RNA control (Thermo Fisher) and 20 nM for the guide RNAs used for CYFIP1-targeting.

### 2.6 Gene expression analysis

#### 2.6.1 RNA extraction

RNA was isolated from cultured cells using a phenol/chloroform extraction protocol. Firstly, cultured cells from one well of a 12-well plate were washed once with PBS and lysed in 1 ml of TRIzol® (Thermo Fisher). Cell lysates from the same experiment were stored at -80°C until all the samples, from multiple cell lines and/or multiple time points, were collected. Biological duplicates for each sample of each experiment, were taken.

The extraction was performed following the manufacturer’s protocol. Briefly, the samples were thawed and 200 µl of chloroform were added to each of them. The samples were then shaken and centrifuged at 12,000g for 15 minutes to allow the separation of the mixture into a lower red phenol-chloroform phase, a white interphase, and a colourless upper aqueous phase. The RNA-containing aqueous phase was transferred to a new tube, mixed with 500 µl isopropanol and left
to stand for 5 minutes at room temperature. The RNA was precipitated by centrifuging the samples at 12,000g for 10 minutes. The supernatant was discarded and the pellet washed in 75% ethanol. The RNA was pelleted again by centrifugation at 7,500g for 5 minutes. After removing ethanol RNA samples were air dried and resuspended in DEPC-treated bi-distilled water (ddH2O). All centrifugation steps were performed at 4°C. The RNA concentration was measured using a BioSpectometer (Eppendorf).

2.6.2 DNase treatment

RNA was treated with the TURBO DNA-free kit (Thermo Fisher) to remove any presence of contaminant DNA that could compromise the qPCR analysis. For each sample, 10 µg of RNA were diluted with nuclease-free water to a final volume of 17 µl, to which 2 µl TURBO DNase 10x buffer and 1 µl TURBO DNase were added. The samples were incubated at 37°C for 30 minutes. After this time, 2 µl of DNase inactivation reagent were added to stop the reaction and incubated for 5 minutes at room temperature. Finally, the samples were centrifuged for 1.5 minutes at 10,000g to pellet the DNase enzyme and inactivation reagent and the supernatant, containing DNA-free RNA, was transferred to a new clean tube. At the end of the DNase treatment the RNA concentration was re-measured.

2.6.3 Reverse transcription

The qScript cDNA Supermix (Quantabio) was used for reverse transcription. Following the manufacturer protocol, 1 µg of RNA for each sample was diluted in DEPC-treated ddH2O up to a volume of 16 µl, which was combined with 4 µl of 5X Supermix. Samples were placed in a T100 Thermal Cycler (Biorad) and the recommended program was run (25°C for 5 minutes, 42°C for 60 minutes and 85°C for 5 minutes). The resulting cDNA was diluted 1:10 with ddH2O and stored at -20°C.

2.6.4 qPCR

Quantitative PCR (qPCR) was used to quantify transcripts of interest using fluorescent dsDNA-binding dye and measuring its signal as the specific sequence of interest is amplified during the reaction. For this process, each sample was run as a triplicate. Firstly, a mastermix containing 10 µl of 2X PerfeCTa SYBR Green SuperMix (Quantabio), 2µl of primers (from a 10 µM stock) for the gene of interest and 6 µl of ddH2O was prepared for each well of 96-well plate needed for the analysis. A volume of 18 µl of this mastermix was pipetted into each well together with 2µl of the cDNA solution of the appropriate sample. The qPCR reaction was run on a Bio-Rad CFX Connect Real-Time System. The program used included an initial incubation at 95°C for 4 mins, followed by 40 cycles
of 94°C for 30 secs, 60°C for 15 secs and 72°C for 30 secs. A melting curve was also generated to check for the specificity of the product at the end of the run. The data were analysed on Microsoft Excel using the ΔΔCT method for relative quantification (Livak & Schmittgen, 2001; Pfaffl, 2001).

**Table 2.3 Primers used for qPCR**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Fw sequence (5'-3')</th>
<th>Rev sequence (5'-3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>IRES</td>
<td>CCACCATAATTGCGGTCTTT</td>
<td>GAGGAACTGCTTCTTTCA</td>
</tr>
<tr>
<td>CYFIP1</td>
<td>GAAAACCGTGGAGGTCTGGA</td>
<td>GCTCAATGCGATTTCTTGGAAG</td>
</tr>
<tr>
<td>NIPA1</td>
<td>TGGGCACCTTGGAGTA</td>
<td>TGCCCAAGATGTTAGCTTTTT</td>
</tr>
<tr>
<td>NIPA2</td>
<td>GGCACACTTGCGGTGTTAT</td>
<td>GGCACTTACTAGCCGCTGAGA</td>
</tr>
<tr>
<td>TBGCP5</td>
<td>TGGTTACTTTCCTGTAGGTTAG</td>
<td>TGGCAACTGCTTCTTGTTCATGT</td>
</tr>
<tr>
<td>GAPDH</td>
<td>ATGACATCAAGAAGGTGTTG</td>
<td>CACATCGGAAATGAGCTTTG</td>
</tr>
<tr>
<td>18S rRNA</td>
<td>GAAACCCGTGAACCCATT</td>
<td>CCAACCCGTGAACCCATT</td>
</tr>
</tbody>
</table>

(NIPA1, NIPA2 and TUBGCP5 primers were published by Picinelli et al., 2016)

2.7 Genotyping

2.7.1 Genomic DNA extraction

Cultured cells were washed once with PBS and incubated at 37°C overnight in lysis buffer (10 mM Tris-pH8.0, 50 mM EDTA, 100 mM NaCl, 0.5% SDS) supplemented with 0.5mg/ml of Proteinase K (all components from Sigma). The following day the lysis buffer was mixed with an equal volume of isopropanol and the DNA was precipitated at 15,000g for 20 minutes. The resulting pellet was washed with 70% ethanol, air-dried and resuspended in an appropriate amount of ddH2O. DNA concentration was measured with a Biospectrometer (Eppendorf).

2.7.2 PCR and DNA electrophoresis

PCR was used for the screening of CYFIP1-targeted clones. Each PCR reaction included about 100ng of template genomic DNA, 10µl of 5X MyTaq Buffer, 1 µl of MyTaq Polymerase (Bioline), 2 µl of primers from a 20µM stock and ddH2O up to 50µl. Usually, 30 amplification cycles were performed in a T100 Thermal Cycler (BioRad). Each cycle included 15 seconds of denaturation at 95°C, 15 seconds of annealing at 55-57°C and 10-20 seconds of extension at 72°C. The annealing temperature and length of the extension interval were optimised for each pair of primers. Finally, 5 µl of the final PCR product were run on agarose gel (1 to 4% agarose, according to the size of the PCR amplicon).
### Table 2.4 Primers used for genotyping of CYFIP1 locus

<table>
<thead>
<tr>
<th>Gene</th>
<th>Fw sequence (5’-3’)</th>
<th>Rev sequence (5’-3’)</th>
<th>Product length</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYFIP1</td>
<td>ATGTGTTGTTCCAGCCCAGG</td>
<td>CATCATGTGGGTCGGAGC</td>
<td>190bp</td>
</tr>
<tr>
<td>CYFIP1</td>
<td>CCCTGAGAGAGACACGCAAC</td>
<td>CCTCACTGCATAGTCTATTGGGA</td>
<td>557bp</td>
</tr>
</tbody>
</table>

#### 2.7.3 Surveyor assay

The Surveyor assay allows the quantification of the targeting efficiency in a pool of transfected cells using the activity of the Surveyor Nuclease, which recognizes and cleaves mismatches caused by small indels in a DNA fragment. The Surveyor® Mutation Detection Kit (Integrated DNA Technologies) was used for this purpose.

Following the manufacturer protocol, a PCR product derived from a pool of targeted cells was placed in the T100 Thermal Cycler (Biorad) to undergo a cycle of denaturation and re-annealing (95°C for 10 minutes, ramp down to 85°C at -2°C /s; ramp down to 25°C at -0.1°C /s; hold at 4°C). This allows the formation of heteroduplexes between the different DNA molecules within the PCR mixture, which was then mixed with MgCl2 (1:10) and SURVEYOR Enhancer S (1µl). The reaction was split in two equal parts, of which only one was treated with SURVEYOR Nuclease S (1µl), while the other was used as “non-treated” control. Both aliquots were incubated at 42°C for 60 minutes and then mixed with Stop Solution (1:10). Finally, reactions were run on a 4% agarose gel and imaged. The intensity of the bands, measured by ImageJ, was used to calculate the targeting efficiency according to the formula below (Ran et al., 2013).

\[ \%\text{indels} = 100 \times (1 - (1- \text{fraction cleaved})^{1/2}) \]

#### 2.7.4 PCR cloning and sequencing analysis

Candidate CYFIP1 mutant hESC lines identified by PCR were selected for sequencing to verify the presence of indels and out-of-frame mutations. To this end, the targeted locus was amplified with a pair of primers producing a 600 bp product followed by cloning into the pGEM-T vector (Promega) following the manufacturer’s protocol. A ligation containing pGEM-T Easy vector (1µl), 2X reaction buffer (5µl), T4 DNA ligase (1µl) and purified PCR product (3µl) was set up and incubated for 1 hour at room temperature. A volume of 2 µl from this reaction was used to transform competent cells (NEB® 5-alpha F‘lq Competent E. coli) with a heat-shock mediated transformation. The competent cells were then plated on LB plates and incubated overnight at 37°C. For each transformation, the plasmid DNA was extracted from 10 minipreps as described above (2.3) and EcoRI digestion was used to verify the presence of the insert. The plasmids that has successfully incorporated the PCR amplicon were sent to GATC Biotec (gatc-biotec.com) for Sanger sequencing. The results were analysed with the BioEdit software (www.mbio.ncsu.edu/BioEdit). Each sequence derived from a
targeted clone was aligned with the WT sequence and translated into the predicted protein, to characterise the type of mutation generated during the targeting.

2.8 Western Blotting

Cultured cells were lysed on ice using RIPA buffer (Abcam) supplemented with a protease and phosphatase inhibitors cocktail (Sigma). Cell lysates were centrifuged for 15 minutes at 12000g and the resulting supernatant was combined with 1X Bolt®LDS Sample Buffer (Thermo Fisher) and 1X Bolt®Sample Reducing Agent (Thermo Fisher) and boiled at 97°C for 5 minutes. Equal amounts of proteins for each sample were separated on 4-12% Bolt® Bis-Tris Plus gels (ThermoFisher) and then transferred to a PVDF membrane (0.45 µm pore size, Amersham Hybond, GE Healthcare) via electro-blotting. The membrane was blocked in 5% BSA (Sigma) in Tris Buffered Saline containing 0.1% Tween (TBS-T) and incubated with primary antibodies overnight at 4°C. The membrane was then washed three times in TBS-T, incubated 1 hour at room temperature with secondary antibodies and washed again before imaging. The secondary antibodies used were conjugated to horse-radish peroxidase (HRP) (Abcam) or to infra-red dyes (IRDye 800CW or 680RD) (LI-COR). HRP secondary antibodies were detected following 5 minutes incubation with the reagent Luminata Crescendo HRP substrate (Millipore). Chemiluminescent detection was carried out with a BioRad ChemiDoc™ XRS+ system. Alternatively, when IR secondary antibodies were used, the blots were imaged with the LI-COR Odyssey CLx system. In both cases, the images were exported to ImageJ for quantification.

Table 2.5 Antibodies used for western blotting

<table>
<thead>
<tr>
<th>Target</th>
<th>Species</th>
<th>Dilution</th>
<th>Cat. Number</th>
<th>Supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>βCAT</td>
<td>mouse</td>
<td>1:1000</td>
<td>Sc-7963</td>
<td>Santa Cruz</td>
</tr>
<tr>
<td>βCAT (p-ser33/37)</td>
<td>rabbit</td>
<td>1:500</td>
<td>2009</td>
<td>Cell Signalling Technologies</td>
</tr>
<tr>
<td>βCAT (p-Ser552)</td>
<td>rabbit</td>
<td>1:500</td>
<td>9566</td>
<td>Cell Signalling Technologies</td>
</tr>
<tr>
<td>CYFIP1</td>
<td>rabbit</td>
<td>1:3000</td>
<td>AB6046</td>
<td>Millipore</td>
</tr>
<tr>
<td>GAPDH</td>
<td>mouse</td>
<td>1:10000</td>
<td>AB8245</td>
<td>Abcam</td>
</tr>
<tr>
<td>GAPDH</td>
<td>rabbit</td>
<td>1:10000</td>
<td>AB9485</td>
<td>Abcam</td>
</tr>
<tr>
<td>NCAD</td>
<td>mouse</td>
<td>1:1000</td>
<td>18-0224</td>
<td>Thermo Fisher</td>
</tr>
<tr>
<td>NIPA1</td>
<td>rabbit</td>
<td>1:500</td>
<td>AB128640</td>
<td>Abcam</td>
</tr>
<tr>
<td>NIPA2</td>
<td>rabbit</td>
<td>1:500</td>
<td>AB84343</td>
<td>Abcam</td>
</tr>
<tr>
<td>TUBGCP5</td>
<td>mouse</td>
<td>1:1000</td>
<td>AB168325</td>
<td>Abcam</td>
</tr>
</tbody>
</table>
2.9 Immunocytochemistry

2.9.1 Immunofluorescence staining

Cultured cells were washed with DPBS and fixed with cold 3.7% PFA for 15 minutes. PFA incubation was followed by 3 PBS (Sigma) washes before proceeding with staining. In the case of cerebral organoids, fixing was performed in a well of 24-well plate. Here, the desired number of organoids was washed with DPBS and incubated in cold 4% PFA for 15 minutes. They were then incubated overnight at 4°C in a 30% sucrose solution in PBS. The organoids were then embedded in a pre-warmed solution containing 7.5% gelatin (Sigma) and 10% sucrose (Sigma) in PBS. Gelatin blocks were cut with a cryostat (Leica) into 13 µm-thick sections and stored at -80°C.

For staining, cells or organoids sections were first permeabilised in PBS-T (0.3% Triton-X-100 in PBS) for 10 minutes, and then blocked in PBS-T with 2% BSA and 3% donkey serum (Gentaur) for 30 minutes at room temperature. Cells or sections were incubated with primary antibodies in blocking solution overnight at 4°C. They were then washed 3 times for 10 minutes in PBS-T and AlexaFluor secondary antibodies (Thermo Fisher), diluted in PBS-T, were added and incubated for 2 hours at room temperature, in the dark. Nuclei were stained with DAPI (Sigma), diluted 1:3000 in PBS. After 3 more washes with PBS, stained cells or sections were mounted with DAKO fluorescent mounting medium and stored at 4°C.

Table 2.6 Antibodies used for immunofluorescence

<table>
<thead>
<tr>
<th>Target</th>
<th>Species</th>
<th>Dilution</th>
<th>Cat. Number</th>
<th>Supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTIP2</td>
<td>Rat</td>
<td>1:500</td>
<td>AB18465</td>
<td>Abcam</td>
</tr>
<tr>
<td>FOXG1</td>
<td>Rabbit</td>
<td>1:1000</td>
<td>AB18259</td>
<td>Abcam (ChIP grade)</td>
</tr>
<tr>
<td>GABA</td>
<td>Rabbit</td>
<td>1:1000</td>
<td>A2052</td>
<td>Sigma</td>
</tr>
<tr>
<td>KI67</td>
<td>Mouse</td>
<td>1:1000</td>
<td>550609</td>
<td>BD Pharmigen</td>
</tr>
<tr>
<td>MAP2</td>
<td>Mouse</td>
<td>1:1000</td>
<td>M1406</td>
<td>Sigma</td>
</tr>
<tr>
<td>NCAD</td>
<td>Mouse</td>
<td>1:1000</td>
<td>18-0224</td>
<td>Thermo Fisher</td>
</tr>
<tr>
<td>NESTIN</td>
<td>Rabbit</td>
<td>1:500</td>
<td>ABD69</td>
<td>Millipore</td>
</tr>
<tr>
<td>NEUN</td>
<td>Mouse</td>
<td>1:500</td>
<td>MAB377</td>
<td>Millipore</td>
</tr>
<tr>
<td>NEUN</td>
<td>Rabbit</td>
<td>1:500</td>
<td>ABN78</td>
<td>Millipore</td>
</tr>
<tr>
<td>OCT4</td>
<td>Goat</td>
<td>1:500</td>
<td>SC-8629</td>
<td>Santa Cruz</td>
</tr>
<tr>
<td>P27kip1</td>
<td>Rabbit</td>
<td>1:500</td>
<td>06-445</td>
<td>Upstate (Millipore)</td>
</tr>
<tr>
<td>PAX6</td>
<td>Mouse</td>
<td>1:1000</td>
<td>PAX6</td>
<td>DSHB</td>
</tr>
<tr>
<td>PHH3</td>
<td>Rat</td>
<td>1:1000</td>
<td>641002</td>
<td>Biolegend</td>
</tr>
<tr>
<td>SATB2</td>
<td>Mouse</td>
<td>1:1000</td>
<td>AB51502</td>
<td>Abcam</td>
</tr>
<tr>
<td>SSEA3</td>
<td>Rat</td>
<td>1:500</td>
<td>MAB4303</td>
<td>Millipore</td>
</tr>
<tr>
<td>TBR1</td>
<td>Rabbit</td>
<td>1:500</td>
<td>AB31940</td>
<td>Abcam</td>
</tr>
<tr>
<td>TBR2</td>
<td>rabbit</td>
<td>1:500</td>
<td>AB23345</td>
<td>Abcam</td>
</tr>
</tbody>
</table>
2.9.2 EdU proliferation assay

The Click-iT® EdU Alexa Fluor Imaging Kit (Thermo Fisher) was used to measure proliferation of neural progenitors at several stages of differentiation. Cultured cells were incubated with EdU (1:1000, final concentration 10 µM) in culture media at 37°C, usually for 2 hours, then fixed in PFA, as described above. EdU detection was performed by incubating the cells for 20 minutes at room temperature in the Click-iT reaction cocktail, prepared as per manufacturer’s protocol. If co-staining with other nuclear markers, this step was performed after the incubation with secondary antibodies. Finally, EdU detection was followed by nuclei stain with DAPI and mounting, as described above.

2.9.3 Imaging and picture analysis

Stained cells were imaged using a Leica DM6000B inverted microscope, if fixed on plates, or a Leica DMI6000B upright microscope, if fixed on coverslips and mounted on slides. An average of 10 random fields was acquired for each well or coverslip at a 20x magnification, if intended for quantification. Cell counting was performed automatically, using Cell Profiler (cellprofiler.org), for DAPI and most nuclear markers. Other measurements, such as number and morphology of neural rosettes, were performed manually, using the ImageJ software. Data analysis and representation was carried out using Microsoft Excel and R (www.r-project.org). Unless otherwise stated, all the immunohistochemistry quantifications were collected from at least two independent experiments, with at least two biological replicates for each marker counted.

2.10 High Content analysis of mitochondria

2.10.1 Mitotracker staining

Neuronal cells were seeded, 2 or 3 days before the analysis, in 96 well- plates at a density of 2000 cells per well. On the day of the analysis, they were incubated with 100 nM Mitotracker (ThermoFisher) for 1 hour at 37°C. Cultures were then fixed in PFA and the nuclei were stained with DAPI, as described above. The stained cells were preserved in DAKO for a few days at 4 °C.

2.10.2 Cell Insight CX7 High Content Screening (HCS) platform

The Cell Insight CX7 HCS platform (ThermoFisher) combines fluorescent microscopy (wide-field or confocal), automated image acquisition and quantitative analysis. The Spot Detector application allows quantitative analysis of fluorescence punctate stainings and was therefore used for the
analysis of Mitotracker-labelled mitochondria. Before running the analysis, the “miniscan” function was used to set up and test all the parameters necessary for the assay. These include the background removal, thresholding and segmentation for the DAPI staining. The correct identification of all the nuclei in each field, also depends on the elimination of DAPI+ “objects” of abnormal shape or size, like small debris or clusters of nuclei not correctly segmented. For the Mitotracker staining, only background removal and thresholding were applied. Images were acquired using the confocal function. At least 10 fields from about 10 wells for each cell line and time point were imaged without changing any of the acquisition and processing settings. Quantitative data relative to spot count, intensity and nuclear morphology were exported and analysed in Microsoft Excel and R.

2.11 Flow cytometry analysis

Flow cytometry was used to evaluate the efficiency of RNA transfection during the generation of CYFIP1-tageted cell lines (described in Chapter 3) and to analyse the cell cycle profile after DAPI staining (results in Chapter 4). Before the analysis, both transfected (or stained) and non-transfected (or unstained) cells were dissociated in Accutase (Thermo Fisher) for 10 minutes at 37°C. When this reagent was removed, the cells were resuspended in cold DPBS and analysed on a BD LSR Fortessa cytometer (BD Biosciences). Lasers of the appropriate wavelength were used for exciting the samples and gates were set using the unstained samples as negative controls. The instrument was set up with the help of Mark Bishop, lab manager for the European Cancer Stem Cells Research Institute of Cardiff University. Data were analysed with FlowJo_V10 software.

2.12 Transcriptome analysis via RNA sequencing (with Daniel Cabezas de la Fuente)

RNA was collected for sequencing at three time points of differentiation, for each cell line, in triplicate. The extraction was performed with the PureLink® RNA minikit (Thermofisher), following manufacturer’s instructions. RNA quality was checked on an Agilent Bioanalyzer 2100, by the Central Biotechnology Services (CBS) facility of Cardiff University. After confirmation that all the samples had a RNA Integrity Number (RIN) above 8, proving the absence of degradation, they were sent to the Oxford Genomics Centre for library construction and RNA sequencing (RNA-seq) (http://www.well.ox.ac.uk/ogc). Here, 1 μg of RNA was used for library construction with a TruSeq mRNA Library Preparation kit (Illumina). Paired-end sequencing was performed on an Illumina HiSeq 4000 with a depth above 28 million reads per sample.
Processing of the raw data, quality control and differential gene expression were performed by Daniel Cabezas de la Fuente, with the help of Dr Robert Andrews (Cardiff University Data Clinic) and Dr Andrew Pocklington (MRC Centre for Neuropsychiatric Genetics and Genomics). Briefly, the FASTQ files generated from the sequencing were trimmed from any residual Illumina adapter sequence and from low quality bases, then mapped to the Ensembl human genome GRCh38.84 (hg38) using STAR (Bolger, Lohse & Usadel, 2014; Dobin et al., 2013). The FastQC software was used to perform a quality control of the data (www.bioinformatics.babraham.ac.uk/ projects/fastqc). Finally, the expression values for each gene, intended as the number of reads aligning to its sequence, were generated using the FeatureCounts software (Liao, Smyth & Shi, 2014). These values were used for differential gene expression analysis. This was performed in R, using the Bioconductor package DESeq2 (Anders & Huber, 2010). The Benjamini-Hochberg (BH) correction method was applied to control for false discovery rate (Benjamini & Hochberg, 1995). The differentially expressed genes (DEGs) resulting from this process were analysed with the Bioconductor package ClusterProfiler to identify dysregulated pathways via KEGG gene set enrichment analysis (Kanehisa et al., 2012; Yu et al., 2012a). The selection of genes related to a specific function was carried out using the R package dplyr (CRAN.R-project.org/package=dplyr). With this, the data-frame of DEGs between two cell lines at one time point was cross-compared with the desired list of EnrezIDs. Neurogenesis-related transcripts were obtained from the Gene Ontology website (geneontology.org) and mitochondria-related transcripts were from the MitoCarta2.0 website (broadinstitute.org) (Calvo, Clauser & Mootha, 2016). Graphic representation of RNA-seq data was carried out with the R package ggplot2 (CRAN.R-project.org/package=ggplot2).
3 Derivation of hESCs with altered CYFIP1 dosage and iPSCs carrying 15q11.2 CNVs

3.1 Introduction

Psychiatric disorders are caused by a complex combination of genetic and environmental factors (Robinson, Neale & Hyman, 2015; Sullivan, 2005). Many common genetic variants have been demonstrated to increase the risk of developing diseases such as ASD and SZ, with variable penetrance. Some of these variants are represented by CNVs, which often include more than one gene (Gratten et al., 2014).

In this context, the use of patient-derived iPSCs represents an advantageous alternative to the generation of animal or cellular models carrying mutations in several genes, the generation of which would be technically challenging. However, because of the genetic variability between individuals, iPSC-based experiments need to include several clonal lines from different patients, in order to produce reliable results. A solution to this problem is represented by working with genetically-modified hESCs and their parental isogenic line, which are not subject to the same degree of variability. This approach also allows to explore the phenotype specifically caused by one gene of interest, within a CNV, independently from the other genes in the same region and from the genetic background of a patient. Therefore, modified hESCs and patient-derived iPSCs represent two complementary aspects, which should both be considered for in vitro disease modelling.

Since the start of iPSC research, an efficient and popular method for the reprogramming of patients’ somatic cells has been the ectopic expression of pluripotency-inducing factors using viral vectors, in particular retrovirus (Park, 2008; Takahashi et al., 2007). However, retroviral vectors have been demonstrated to be easily silenced, eventually resulting in incomplete reprogramming, and require integration into the host genome, leading to the disruption of random genes, with unknown consequences (Hu, 2014).

Sendai virus-based (SeV) vectors have been developed as a safer alternative to express transgenes in a variety of host cells and have become an increasingly popular method for iPSC reprogramming. This is because SeVs are not pathogenic for humans and do not require integration into the host DNA. Moreover, the development of SeV vectors with depletion of fusion proteins and/or introduction of temperature-sensitive mutations into key viral proteins has opened the possibility of deriving virus-free stably reprogrammed iPSCs from a variety of somatic cells (Ban et al., 2011; Fusaki et al., 2009; Nakanishi & Otsu, 2012). These advantages have led our choice of
reprogramming system towards a commercially available kit containing SeV-based vectors for the delivery of Oct3/4, Sox2, c-Myc and Klf4, which have been successfully used by several published studies (Fusaki et al., 2009; MacArthur et al., 2012; Tucker et al., 2013).

The generation of iPSCs carrying 15q11.2 CNVs is only one of the tools necessary to investigate the role played by the genes in this locus, particularly CYFIP1, during brain development. Together with patients’ cells, the derivation of isogenic hESC lines carrying CYFIP1 mutations alone is also necessary and, because 15q11.2 CNVs can occur as deletions or duplications, hESCs with increased and decreased expression of this gene are equally important.

The overexpression of a gene of interest in hESCs can be challenging, as transgenes tend to be silenced over subsequent passaging of undifferentiated cells and also during the course of in vitro differentiation (Liew et al., 2007). However, a comparison of several vectors where the expression of eGFP (enhanced Green Fluorescent Protein) was driven by different promoters, showed that the pCAG promoter allows the generation of hESCs lines with stable transgene expression, maintained also in the differentiated progeny (Liew et al., 2007). This promoter is formed by the sequences of the cytomegalovirus early enhancer element, the promoter and first exon and intron of the chicken β-Actin gene and the splice acceptor site of the rabbit β-Globin gene and appears to be efficient also in bicistronic vectors. For instance, transfection of pCAG-GFP-IRES-PURO, where the GFP and puromycin-resistance gene (PURO) are linked by an internal ribosomal entry site (IRES), has been applied to several hESC lines to derive stably transfected clones, without affecting their pluripotency (Braam et al., 2008). Therefore, this approach appears to be promising also for the generation of CYFIP1-overexpressing (CYFIP1tg) hESCs.

Gene targeting in hPSCs has recently become significantly easier by the discovery of the CRISPR/Cas9 system and its optimisation for use in mammalian cells. This technique requires the construction of plasmids or viral vectors allowing co-expression of the Cas9 and gRNAs into the target cells (Cho et al., 2013; Cong et al., 2013; Jinek et al., 2013; Mali et al., 2013). The need for several cloning steps can be bypassed using the iCRISPR platform, represented by hESCs containing a doxycycline-inducible Cas9 stably integrated in their genome. Once the Cas9 expression is induced, the only step required is the transfection of gRNAs, which can be generated by in vitro transcription or purchased. This method has been reported to have an efficiency as high as 40% (González et al., 2014). For this reason, the iCRISPR platform was chosen for the generation of CYFIP1-knockout (CYFIP1ko) hESCs.

In this chapter, I describe the derivation and validation of CYFIP1tg and CYFIP1ko hESCs and the reprogramming of 15q11.2-deleted fibroblasts and keratinocytes into iPSCs. The characterisation of these patient-derived lines includes the analysis of protein levels for the four 15q11.2 genes in
iPSCs-derived NPCs and neurons, an aspect that has not been reported to date. The generation of iPSCs carrying 15q11.2 duplication has not been possible, because no patients with such mutation were recruited during my PhD.

3.2 Results

3.2.1 Generation of CYFIP1tg hESCs

(I) Construction of Cyfip1 overexpression vector

For the generation of CYFIP1tg hESCs, a bicistronic vector containing the CYFIP1 and puromycin acetyltransferase (PAC) genes under the control of the pCAG promoter was created (pCAG-hCYFIP1-IRES-PAC). Such a construct allows high constitutive CYFIP1 expression and provides puromycin resistance to all stably transfected cells.

The full-length human CYFIP1 open reading frame (ORF) was cut out of a commercially available expression vector and cloned downstream of the CAG promoter in a pCAG-IRES-PAC plasmid (Fig. 3.1 A). A digestion with the restriction enzymes PvuI and BglII was performed to confirm the presence and orientation of the right insert (Fig. 3.1 B). The plasmid, the digestion of which is shown in lane 1, was used for transfection of H7 hESCs.

**Figure 3.1 CYFIP1-overexpression vector construction.**

(A) Map of CYFIP1-overexpression plasmid and (B) DNA electrophoresis gel showing BglII/PvuI digestion pattern of the same plasmid with CYFIP1 insert in the correct orientation (lane 1), in the wrong orientation (lane 2) or without insert (lane 3).
Isolation of Puromycin-resistant hESCs clones and validation of increased Cyfip1 during neuronal differentiation

The pCAG-Cyfip1-IRES-PAC plasmid was delivered into H7 hESCs via nucleofection, as described in Chapter 2. Puromycin selection was applied 48h after transfection and maintained for 10 days. During this time, 11 colonies of puromycin-resistant cells emerged and were picked manually and expanded as independent clones. RT-qPCR was performed to measure CYFIP1 and IRES mRNA levels at the stem cell stage. The results revealed that CYFIP1tg clones #3, #4 and #5 expressed the highest level of CYFIP1 transgene (Fig. 3.2 A and B). Despite having the correct ESCs morphology (Fig. 3.2 C), cells from clone #4 showed a poor efficiency of differentiation towards the neuronal lineage and were discarded from subsequent differentiation experiments. Conversely, CYFIP1tg #3 and #5 turned efficiently into cortical neurons and, when compared to the parental line H7, maintained a higher CYFIP1 expression throughout the whole process, as demonstrated by the qPCR and western blot data (Fig. 3.2 E, F).
Figure 3.2 Characterisation of CYFIP1tg hESCs.

RT-qPCR results for IRES (A) and CYFIP1 (B) transcripts in all the hESC clones isolated after puromycin selection. Quantification data are normalised to GAPDH and 18S rRNA. Expression values are relative to the H7-CTRL, taken as 1. Error bars represent SD between 3 technical replicates. (C) Bright field images of three CYFIP1tg clones, showing normal hESC-like morphology. Scale bar = 100 μm. (D) RT-qPCR results for CYFIP1 expression at different stages of cortical differentiation of CYFIP1tg, clones #3 and #5 (clone numbers defined in A). Data are normalised to GAPDH and 18S rRNA. Expression values are relative to the control line at day 5, taken as 1. Quantification (E) and western blots (F) for CYFIP1 protein during cortical differentiation of CYFIP1tg (#3 and #5) and CTRL cells. RT-qPCR and western blot data (D and E) represent the mean value ± SD from 4 independent experiments (N=4, 2 experiments for each CYFIP1tg clone). Data were compared by two-way ANOVA, followed by Tukey HSD test (*p<0.05, **p<0.01).
3.2.2 CYFIP1 targeting in iCas9 hESCs

(I) Validation of RNA transfection procedure

CYFIP1ko hESCs were derived from iCas9 cells, a line of genetically modified pluripotent cells with doxycycline-inducible expression of Cas9 nuclease (González et al., 2014). Gene targeting with this system requires the transfection of gRNA molecules into the cells, after the induction of Cas9 expression.

For this reason, I firstly verified the efficiency of RNA transfection using an Alexa555-conjugated RNA molecule delivered into the cells via a lipid-based transfection reagent optimised for RNA delivery. Flow-cytometry was used to analyse the population of transfected cells 24 hours later. The results showed that over 50% of the cells were fluorescent when compared to the non-transfected (negative) control sample (Fig.3.3).

![Figure 3.3 Validation of RNA transfection method.](image)

*Figure 3.3 Validation of RNA transfection method.*

Histograms of FACS analysis showing a control (A) and transfected (B) sample of iCas9 cells. The analysis was carried out 24 hours post-transfection.

(II) Transfection and screening of targeted clones

After confirming the efficiency of the transfection method, iCas9 hESCs were transfected with gRNAs targeting the first exon of CYFIP1 (Fig. 3.4 A). The synthesis and transfection of the gRNAs is described in Chapter 2.2 and 2.4. Twenty-four hours after transfection, one of the transfected wells was used for extraction of the genomic DNA, together with a control (non-transfected) one. A PCR reaction was carried out for both samples to amplify the targeted locus and the product was then digested with Surveyor nuclease, to evaluate the targeting efficiency. Indel occurrence was estimated as explained in Ran et al. and resulted to be about 15% (Ran et al., 2013). A DNA gel showing the result of the Surveyor assay is shown in figure 3.4 B.
The remaining cells were dissociated, plated at a clonal density (250-500 cells per 6cm dish) and left to grow until small colonies emerged. About 100 of these colonies were manually isolated and expanded as single clones. Genomic DNA was extracted from some cells of each clone and used for PCR screening of the targeted CYFIP1 locus. This process led to the identification of several clones carrying insertions and/or deletions (indels) in the region of interest. An example of an agarose gel used to analyse the PCR results is shown in figure 3.4 B.

Sanger sequencing was then used to verify the results of the PCR screening and to determine the exact nucleotide sequence of the targeted locus. As shown in figure 3.4 D and E, clones #31, #41 and #70 had mutations in both alleles, but in the case of #41 and #70, one of the alleles had a deletion of 3 and 6 nucleotides respectively, leaving the gene sequence in frame. Other clonal lines containing mutations were #53 and #97, with a 1bp insertion in one allele, and #71 and #97, with only an in-frame deletion in one allele. Therefore, only clone #31 resulted to be a functional homozygous knock-out, with all the others being genetically or functionally heterozygous. Clone #31 was used for most of the experiments presented in this thesis and will be referred to as CYFIP1ko hereafter.

(III) Validation of targeted clones

The CYFIP1ko and two of the heterozygous lines (#41 and #70) were differentiated into cortical projection neurons in parallel to the control line to confirm that the mutations introduced did not affect the differentiation process in a way that would prevent the use of these mutant lines. During this process, proteins were collected from neural progenitors, a stage when CYFIP1 is known to be expressed. The results shown in figure 3.4 F and G confirm a significant reduction of CYFIP1 protein in all the targeted clones, particularly CYFIP1ko, when compared to the parental line.

These data confirm that the cell lines generated by CRISPR/Cas9 targeting of CYFIP1 are suitable for modelling the effects of a reduction of this gene in human neurons, using in vitro differentiation from hESCs.
Figure 3.4 Screening and characterisation of CYFIP1-mutant hESCs.

(A) Scheme of the CYFIP1 gene locus, highlighting the first translated exon (red) and the positions targeted by the three guide sequences used. gRNAs are reported with 5’ to 3’ orientation followed by the PAM (blue). (B) Surveyor nuclease assay of genomic DNA extracted from an untransfected pool of iCas9 hESCs (lane 1) and a well of transfected cells (lane 2). (C) Example of a DNA electrophoresis gel showing PCR products of the targeted locus from iCas9 hESCs (lane 1) and three targeted clonal lines where indels are present (lanes 2, 3, 4). (D) Summary of the genotyping results for all the clonal lines containing indels in one or both alleles. (E) Partial DNA sequence of CYFIP1 exon 1, in the parental iCas9 line and all the targeted clones where indels were found. (F) Western blot for CYFIP1 in CTRL (iCas9), CYFIP1ko (clone#31) and clonal lines #41 and #70 (heterozygous mutants).

3.2.3 Derivation of iPSCs from 15q11.2-deleted fibroblasts and keratinocytes

iPSCs were derived from fibroblasts and keratinocytes of two different patients carrying 15q11.2 deletions using a commercial Sendai virus-based kit. Briefly, a population of somatic cells was transduced with viral vectors to induce the expression of several pluripotency genes (see also Chapter 2). Over time, colonies showing ESC-like morphology started to appear. More than 10
colony per patient were isolated and expanded. These were named after the same identification number (ID) given to the patient followed by another number indicating the cellular clone.

Clones EA008.4, EA008.21, EA062.13, and EA062.19 were stained for pluripotency markers and differentiated into cortical glutamatergic neurons to verify their ability to generate neurons. During this experiment, RNA and proteins were also extracted to quantify the expression levels of the deleted genes. In parallel to the patient-derived iPSCs, two control lines from a male and female healthy volunteer were also differentiated (CTRL#900 and CTRL#202, respectively).

The results of the staining for the pluripotency markers SSEA3 and OCT4 for one control line and one iPSC clone per patient are shown in Figure 3.5 A. The same figure also demonstrates that both patient- and control-derived iPSCs were equally able to generate neuronal cells, as shown by the expression of typical neural progenitor (fig. 3.5 B) and post-mitotic neuron markers (fig. 3.5 C).

Figure 3.5 Immunofluorescent staining in iPSCs and iPSCs-derived neuronal cultures.
(A) Undifferentiated iPSCs stained for the pluripotency markers SSEA3 (green) and OCT4 (red). (B) iPSC-derived neural progenitors stained for a neuroepithelial marker nestin (green). (C) iPSC-derived neurons stained for the neuron-specific cytoskeletal protein MAP2 (green). All nuclei are counterstained with Dapi (blue). Scale bars=50 μm.
The analysis of the expression levels of the four 15q11.2 genes showed that CYFIP1 mRNA and protein were significantly reduced in patient-derived iPSCs compared to controls (fig. 3.6 A and fig. 3.7 A).

NIPA1 mRNA levels in patient iPSC-derived neuronal cells appear to be similar to those from the controls (fig. 3.6 B). However, the threshold cycles detected by RT-QPCR were all considerably high. This could mean the gene is expressed at a very low level, which may compromise the reliability of the quantification. In support to this hypothesis, western blot for NIPA1 showed only a very weak band at the neuronal stage and seemed almost absent in progenitor cells. Moreover, the intensity of NIPA1 bands did not appear to be different between patients and control cells (fig. 3.7 B).

Interestingly, NIPA2 mRNA was particularly abundant at the NPCs stage and the protein was detected only at this time point. In neural progenitors, both the transcript and protein quantification showed significant reduction in cells carrying 15q11.2 deletion when compared to any of the controls used (fig. 3.6 C and fig. 3.7 C). In neurons, only the two clonal lines from EA062 had a significant reduction in NIPA2 mRNA (fig. 3.6 C).

The levels of TUBGCP5 transcript and protein were very variable between lines. Despite an apparent reduction in TUBGCP5 mRNA in EA008 progenitors and EA062 neurons, these differences were not statistically significant. This was reflected in the protein levels, which were similar in all lines and clones, at both differentiation stages (fig. 3.6 D and fig. 3.7 D).

In conclusion, these data describe the characterisation of iPSCs carrying a 15q11.2 deletion. The expression of CYFIP1 and NIPA2 generally correlate with the number of copies of the gene present in the genome, both at the mRNA and protein level. On the other hand, the regulation of TUBGCP5 and NIPA1 expression does not appear to be strictly related to the gene copies carried by an individual. However, NIPA1 seems to be present only at a very low level in in vitro differentiated post-mitotic neurons.
Figure 3.6 RT-qPCR for 15q11.2 genes in iPSCs-derived neuronal cells.

Data were normalised to GAPDH and ribosomal RNA 18S and expressed relatively to the CTRL#900, taken as 1, for each time point. Data represent mean ± SD (N=2 independent experiments) and were compared with a Two-way ANOVA, followed by Tukey HSD test (* p<0.05, **p<0.01).
Figure 3.7 Western blots for 15q11.2 genes in iPSCs-derived neuronal cells.

The band intensity for each gene was normalised to that of GAPDH from the same membrane. Data represent mean ± SD (N=2 independent experiments) and were compared with a Two-way ANOVA, followed by Tukey HSD test (* p<0.05, **p<0.01).
3.3 Discussion

This chapter describes the derivation of several hPSCs lines that lay the basis for the study of CYFIP1 function in human brain development and its contribution to the increased risk for psychiatric disorders in 15q11.2 CNVs carriers.

CYFIP1tg hESCs were derived from H7 hESCs via transfection of a vector where the CYFIP1 gene is driven by the CAG promoter. The use of a strong promoter is necessary to drive high level of constitutive transgene expression in hESCs, because they tend to be silenced during prolonged cultures and differentiation protocols. In agreement with published studies, the CAG promoter proved to be efficient in maintaining a high level of CYFIP1 both in stem cells and in differentiated neurons (Braam et al., 2008; Liew et al., 2007). Indeed, two of the CYFIP1tg hESCs lines tested, clone #3 and #5, were able to successfully differentiate into cortical projection neurons, maintaining a higher level of CYFIP1 expression than the parental H7 line, at any of the time points tested.

However, the stable integration of the CYFIP1-overexpression plasmid in the host’s genome confers risk of interrupting important genes, which could affect the pluripotency and/or differentiation of the transgenic cells. Despite this, none of the CYFIP1tg hESCs clones showed abnormal morphology or behaviour, while the differentiation potential was compromised in only one of the transgenic lines (clone #4). Therefore, these two CYFIP1tg lines (#3 and #5) can be considered as a suitable gain-of-function model to investigate the consequences of an abnormally high level of CYFIP1 during human neural development. Moreover, both clones were differentiated in parallel in most of the experiments, in particular during the initial stage of the project, to allow a confident identification of the cellular phenotypes specifically due to CYFIP1 overexpression.

Human ESCs modelling the loss of CYFIP1 function were derived from the H9-iCas9 line (González et al., 2014), via transfection of gRNAs targeting the first translated exon of CYFIP1. When tested with a fluorescent RNA molecule, the transfection method used to deliver the gRNAs showed an efficiency comparable to that stated in the original publication on the iCRISPR system (González et al., 2014). However, the targeting efficiency reported in the same publication was, on average, around 40%, while that observed during CYFIP1 targeting was around 15%. This is not surprising and is likely due to the efficiency of the guide sequences chosen and to the genetic region targeted. Indeed, the data presented by Gonzalez et al also showed high variability in the percentage of mutant clones generated by targeting of different genes, ranging from as low as 20% to as high as 70%. In addition to this, they also reported the preferential generation of the same non-random deletions in several clones, possibly due to microhomology-directed repair caused by the presence of small repeated sequences adjacent to the region where the DSB occurs (González et al., 2014).
Non-random deletions were detected also during the generation of CYFIP1-targeted lines. These were a 17 nucleotides-deletion present in clone #31 and #41 and a 6 nucleotides deletion present in one allele of clones #70, #71 and #107. Despite the efficiency observed being lower than expected, it was possible to generate a hESC line carrying a full knock-out of the CYFIP1 gene. Inactivation of this gene did not affect the behaviour of the targeted cells and resulted into a strong reduction of CYFIP1 protein in the line named CYFIP1ko. Therefore, hESCs from this clonal line represent a suitable model to investigate the effects of CYFIP1 loss during cortex development.

Finally, iPSCs carrying a 15q11.2 deletion were also generated from two patients. Most of the interest in this CNV has been so far focused on CYFIP1, because of its functions within neuronal cells. The other genes encoded in this region have received little attention and their expression pattern has been investigated by just few studies and only at the mRNA level (Das et al., 2015; Picinelli et al., 2016; van der Zwaag et al., 2010). To carry out an initial characterisation of the patient-derived iPSCs, the expression of all the genes encoded in the 15q11.2 locus at both the mRNA and protein levels in two clonal iPSC lines per patient were compared to two control lines available in our group. RNA and protein samples were collected at the neural progenitor and neuronal stages from cultures differentiating into cortical pyramidal neurons, to carry out the analysis in a cell type relevant to neurodevelopmental psychiatric disorders.

NIPA1 transcripts and protein levels were detected but at a very low level. Neuronal cells derived from patients carrying 15q11.2 deletion did not show a significant reduction of NIPA1 mRNA, in contrast with data from other studies (Das et al., 2015). NIPA1 protein was difficult to detect by western blot and, when present, the bands observed were very weak. Several different combinations of antibodies and protein concentrations were tried to obtain this result. An increase in protein or antibody concentration resulted in the appearance of several unspecific bands rather than a more intense band of the right molecular weight. On the basis of these preliminary data, a small amount of NIPA1 seems to be present only in in vitro-derived cortical neurons and its level at this stage does not appear to be specifically linked to the number of NIPA1 copies in the genome.

TUBGCP5 mRNA and protein were expressed in iPSC-derived neuronal cells at both time points tested. However, transcripts levels were very variable between cell lines and especially between patients. Also, the lack of one gene copy in 15q11.2-deleted cells did not result in significantly reduced protein levels. These results are different from what was expected on the basis of the qPCR data published by Das and colleagues (Das et al., 2015). However, the small sample used in that study and the high degree of variability observed in the results presented here may explain this discrepancy.
CYFIP1 mRNA and protein were detected at both the progenitor and neuronal stages and, in agreement with published data, loss of a copy of the gene results in significantly reduced expression levels for both mRNA and protein (Das et al., 2015; Picinelli et al., 2016).

Finally, NIPA2 transcript was found at higher levels in progenitors than in neurons in all the lines tested. Interestingly, and in line with qPCR data, NIPA2 protein was detected only in neural progenitors and seemed to be absent in neurons. Moreover, the presence of 15q11.2 deletion significantly affects the expression of this gene, resulting in a clear reduction of its mRNA and protein. This is in agreement with published studies reporting that NIPA2 expression is positively correlated with the number of copies of the gene (Das et al., 2015; Picinelli et al., 2016; van der Zwaag et al., 2010). However, only one of these studies is based on in vitro-derived neurons and does not report any indication of the expression trend of NIPA2 mRNA during differentiation, as data are normalised to the control line for each time point analysed (Das et al., 2015). This makes it impossible to compare the interesting expression pattern observed in this study with any published data. NIPA2 specific expression in neural progenitors suggests a role for this gene in the developing cortex and opens a question about a possible contribution of NIPA2 to the phenotypes observed in 15q11.2 CNVs carriers.

NIPA2 is a Mg$^{++}$ transporter, mutations of which have been linked by some studies to an increased risk of epilepsy in the Chinese population (Jiang et al., 2012). Few data are available on its function in neuronal cells. Xie and colleagues used rat primary foetal neurons to demonstrate that mutations of NIPA2 interfere with its correct cellular localisation, lowering intracellular Mg$^{++}$ and disrupting NMDAR currents (Xie et al., 2014). The presence of NIPA2 in rat foetal cortical neurons seems to be in contrast with the impossibility to detect the same protein in in vitro-derived human neurons during the experiments presented here. The samples taken for this study are collected from young hESCs-derived neurons, which should be comparable to the E16-18 cortical cells used by Xie et al. Nevertheless, differences in the model and the maturity of cells could be at the origin of this inconsistency.

Literature on the function of Mg$^{++}$ transporters in NPCs is scarce. However, Mg$^{++}$ is a very abundant ion in mammalian cells and a few studies linked its concentration to the proliferation and differentiation potential of neuronal precursors during adult neurogenesis (Jia et al., 2016; W. Liao et al., 2017). In light of these recent data, the investigation of a possible role of NIPA2 in NPCs represents an interesting venue for future studies.

For most of the time spent working on the experiments presented in this thesis, iPSCs from only one 15q11.2 deletion carrier were available. Because of this, most of my work was focused on CYFIP1tg and CYFIP1ko cells, while work on patients-derived cells is still at a preliminary stage.
4 Phenotypic Analysis of CYFIP1tg and CYFIP1ko cells

4.1 Introduction

The expression of *Cyfip1* in the neuroepithelial cells facing the ventricle in the developing mouse cortex (Yoon et al., 2014 and Eurexpress.org) suggests a role for this gene in cortical development. Indeed, Yoon and colleagues demonstrated that knock-down of CYFIP1 in iPSC-derived neural progenitors disrupts the formation of neural rosettes, a defect observed also in NPCs carrying 15q11.2 deletion (Yoon et al., 2014). The same authors also reported that *Cyfip1* knock-down during mouse embryonic development results in aberrant localisation of radial glial cells and the neurons to which they give rise (Yoon et al., 2014). The apical side of neural rosettes is an important source of signals regulating the survival and neurogenic potential of the neural progenitors forming these structures (Banda et al., 2015; Elkabetz et al., 2008). Therefore, the disruption of the normal rosette morphology associated with reduced level of CYFIP1 is likely to impact on the generation of neurons at a later stage. However, the consequences of neural rosettes alterations in human neuronal cells and the mechanism behind the defects observed were not thoroughly investigated.

The effects of CYFIP1 reduction in NPCs was analysed also by Nebel et al (Nebel et al., 2016). Transfection of shRNAs targeting CYFIP1, followed by RNA sequencing, revealed alterations in genes regulating cell cycle, cytoskeletal remodelling and cell adhesion. The differentially expressed genes were also enriched for FMRP targets and post-synaptic density (PSD) genes. This study also reported differences in the amount of F-Actin and the nuclear and cellular size of CYFIP1-deficient NPCs, but the authors did not examine all the cellular functions that appeared to be affected in the transcriptomic analysis (e.g. cell cycle) nor did they explore which pathway to manipulate to rescue the abnormalities caused by CYFIP1 knock-down (Nebel et al., 2016).

In addition to 15q11.2 deletions, duplications of the same locus are also associated with psychiatric and developmental disorders (Burnside et al., 2011; Picinelli et al., 2016; van der Zwaag et al., 2010). Despite this, the investigation of increased CYFIP1 level in human cortical progenitors and neurons is still missing.

In this Chapter, I report the phenotypic analysis of the CYFIP1tg and CYFIP1ko hESCs in cortical development *in vitro*. To this end, CYFIP1tg and CYFIP1ko neuronal cells were compared to their respective parental lines at different stages during differentiation and were analysed using several approaches. These include assays to evaluate the efficiency of neuronal induction, neural rosette formation and generation of neurons of all cortical layers.
4.2 Results

4.2.1 Changes in CYFIP1 level do not affect neuronal induction efficiency

CYFIP1tg and CYFIP1ko hESCs and their respective parental lines were differentiated to cortical pyramidal neurons using a protocol adapted from published studies (Espuny-Camacho et al., 2013; Shi et al., 2012). Under our conditions, PSCs differentiate into NPCs in about 18 days for H7 and CYFIP1tg cells and in about 12 days in the case of iCAS9 and CYFIP1ko lines. At this time, cells were stained for markers of cortical radial glia cells to analyse the efficiency of NPCs production and the positional identity of the differentiated cells. As shown in figure 4.1 A and B, virtually all cells in CYFIP1tg, CYFIP1ko and control cultures were positive for NESTIN, an intermediate filament widely expressed by proliferating cells of the developing central nervous system (Dahlstrand, Lardelli & Lendahl, 1995). A high proportion of cells was also positive for PAX6 and FOXG1, two transcription factors expressed in dorsal forebrain progenitors and in all forebrain cells, respectively, (Greig, Woodworth et al., 2013), without significant differences between the CYFIP1 engineered and control cell lines (Fig.4.1 A-D).

Moreover, at this stage, the number of cells expressing proliferation markers in CYFIP1tg and CYFIP1ko NPCs was also similar to the respective controls. KI67, a protein present during all active phases of the cell cycle (Scholzen & Gerdes, 2000), was present in more than 60% of cells for all the lines analysed (Fig. 4.2 A-D). Another important parameter in the context of proliferation is the incorporation of the thymidine analogue EdU (Salic & Mitchison, 2008). The ratio between EdU and KI67 is inversely proportional to the cell cycle length and it is often used as an indirect measure of the proliferation rate in cultured cells (Mariani et al., 2015; Qu et al., 2013). Both the number of Edu⁺ cells and the value of the EdU/KI67 ratio were similar for CYFIP1tg, CYFIP1ko and their respective isogenic control lines (Fig 4.2 C, D).

Together these data demonstrate that an altered dosage of CYFIP1 does not affect the efficiency of neural induction, nor the positional identity or proliferation of cortical progenitor cells at an early stage of PSC differentiation.
Figure 4.1 Expression of cortical progenitor markers in CYFIP1tg and CYFIP1ko cultures.

Immunofluorescence for PAX6 (red), FOXG1 (green) and Nestin (red) in CYFIP1tg and H7 cultures at day 18 (A) and in CYFIP1ko and iCas9 cultures at day 14 (B). All nuclei were counterstained with Dapi (blue). Scale bars=50 μm. (C) Quantification of cells expressing PAX6 and FOXG1 in CYFIP1tg and H7 NPCs at day 18 of differentiation. CYFIP1tg data represent an average of two different clonal lines. (D) Percentage of CYFIP1ko and iCas9 NPCs expressing PAX6 and FOXG1 at day 14 of differentiation. Data represent mean ± SD from 4 independent experiments (2 experiments for both CYFIP1tg clones #3 and #5). All quantification data were compared by t-test and no significant differences were found.
Figure 4.2 Expression of proliferation markers in CYFIP1tg, CYFIP1ko and control NPCs.

Immunofluorescence staining for Ki67 (green) and Alexa-555 labelling of EdU (red) in day 18 CYFIP1tg and H7 neural progenitors (A) and in day 14 CYFIP1ko and iCas9 neural progenitors (B). Nuclei are stained with Dapi (blue). Scale bars = 50 μm. (C, D) Quantification of Ki67+ and EdU+ cells and Ki67/EdU ratio in day 18 CYFIP1tg and H7 cultures and in day 14 CYFIP1ko and iCas9 cultures, respectively. Data represent mean ± SD from 3 independent experiments. The CYFIP1tg line used was #5. Data were compared by t-test and no significant differences were found.
### 4.2.2 Abnormal levels of CYFIP1 interfere with the self-organising ability of cortical progenitors

Cultures of CYFIP1tg and control progenitors were stained for NCAD (N-CADHERIN or CADHERIN 2), a protein that \textit{in vivo} is localised at the apical surface of cortical progenitors of the VZ and that \textit{in vitro} marks the centre of neural rosettes. These are polarised structures originated, during monolayer differentiation, by the self-organisation of neural progenitors around a centre devoid of cells (Elkabetz et al., 2008; Miyamoto, Sakane & Hashimoto, 2015).

At a time when control cultures reached their peak of neural rosette formation, CYFIP1-overexpressing progenitors formed significantly fewer rosettes (Fig. 4.3 C). Moreover, while NCAD staining appeared to exhibit the typical circular expression pattern in H7 control cultures, the CYFIP1tg cultures showed irregular NCAD staining, highlighting the aberrant and incomplete morphology of a significant proportion of rosettes (Fig 4.3 A and D). The disorganised nature of CYFIP1tg neural rosettes was more apparent at a higher magnification (Fig. 4.3 B). Nestin-stained CYFIP1tg progenitor cells did not show an apico-basal radial pattern as those in the control rosette (Fig. 4.3 B, white arrowhead). Mitotic phospho-histone H 3 (pHH3)-expressing cells, which \textit{in vivo} are only found adjacent to the ventricle, were correctly positioned around the lumen of neural rosettes in control cultures, but not in CYFIP1tg rosettes (Fig. 4.3 B, green arrowhead).

In order to explore how these cells performed in a more complex system, CYFIP1tg and H7 hESCs were differentiated into brain organoids. Following a published protocol, PSCs were cultured in suspension to form embryoid bodies-like structures, which were then embedded in Matrigel droplets to allow expansion of the neuroepithelial tissue (Lancaster & Knoblich, 2014) (Fig. 4.3 E).

Around 15 days post-Matrigel embedding, several VZ-like areas were evident in the control line. At this time, the thickness and area of each VZ region were measured from the most central section showing the widest ventricular area. The thickness was measured by averaging the length of four radial segments for each VZ. The area was selected with the free hand tool of ImageJ and calculated using the “Measure” function of the same program. The results of this analysis showed that CYFIP1tg hESCs-derived organoids had significantly smaller and thinner VZ-like areas than control ones, at the same time point (Fig. 4.3 F, G, H).
Figure 4.3 Abnormal neural rosettes and VZ-like areas formed by CYFIP1tg neural progenitors.

(A) Immunostaining for NCAD (red) in control and CYFIP1tg cultures at day 18 of differentiation. Scale bars=50μm; (B) Immunostaining for Nestin (white) and pHH3 (green). White arrowhead indicates an area of disorganised Nestin staining. Green arrowhead indicates a pHH3+ cells not adjacent to the rosette lumen. Nuclei were counterstained with Dapi (blue). Scale bars=25μm. (C, D) Quantification of NCAD+ rosettes and percentage of incomplete rosettes present in control and CYFIP1tg cultures at day 18 of differentiation. Quantification was carried out on two different CYFIP1tg clones each time (N=4, two independent experiments for both CYFIP1tg clone #3 and clone #5). (E) Bright field images of brain organoids at day 2 post-seeding and day 5 after Matrigel embedding. Scale bars=100μm; (F) Immunostaining for NCAD and Nestin on a section of a control and a CYFIP1tg organoid at day 15 post-embedding. Nuclei were counterstained with Dapi (blue). Scale bars=50μm; (G and H) Quantification of thickness and surface area of VZ-like structures in organoids derived from H7 and CYFIP1tg cells on day 15 post-embedding (N=3. Four organoids measured for each experiment). All data are represented as mean ± SD and were compared with t-test (*p<0.05, **p<0.01, ***p<0.001).
The formation of neural rosettes was analysed also in CYFIP1ko cultures (Fig. 4.4). Following, immunostaining for NCAD, no differences were found in the number of rosettes between CYFIP1ko and the isogenic control cultures (Fig. 4.4 C). However, the expression pattern of NCAD was found to be abnormal in a high proportion of CYFIP1ko rosettes, as it lacked the typical circular configuration and, instead, appeared in a disperse pattern. In contrast, neural progenitors derived from the control line showed correct radial organisation and circular pattern of NCAD staining (Fig. 4.4 A, B). Quantification of the proportion of incomplete rosettes over two independent experiments confirmed the presence of a significant difference (Fig. 4.4 C).

Overall, these data indicate that abnormal levels of CYFIP1 impact negatively on the self-organisation of cortical neural progenitors into correctly polarised structures.

Figure 4.4 Disorganised neural rosettes in CYFIP1ko progenitor cultures.

(A) Immunostaining for NCAD (green) and Nestin (red) in day 14 cultures of NPCs derived from control and CYFIP1ko cells. Nuclei are marked by Dapi staining (blue). Scale bars=50μm. (B) High magnification field showing one NCAD+ (green) neural rosette for both lines. Nuclei were stained with Dapi (blue). Scale bars=25μm. (C, D) Quantification of the number of neural rosettes and percentage of incomplete ones present in control (iCas9) and CYFIP1ko cultures at day 14 of in vitro differentiation (N=3). Data from the two cell lines are represented as mean ± SD and were compared by t-test (*p<0.05, **p<0.01, ***p<0.001).
4.2.3 CYFIP1 overexpression results in prolonged proliferation of the progenitor pool

To investigate whether the disruption to normal rosettes morphology impacts the generation of neurons at later stages, the number of PAX6⁺ progenitors and post-mitotic NEUN⁺ neurons was analysed in H7 and CYFIP1tg cultures during a window of active neurogenesis, between day 30 and 40 of in vitro differentiation. At both time points the percentage of PAX6⁺ cells was significantly higher in CYFP1tg cultures, while the number of neurons was considerably higher in the control population (Fig. 4.5 A, B, D, E).

A 2-hour EdU incorporation experiment was then performed to quantify the fraction of proliferating cells in both CYFIP1tg and control cultures. This experiment confirmed the presence of a larger number of actively dividing cells within the CYFIP1-overexpressing population both at day 30 and 40, with the difference at day 40 being highly significant (Fig. 4.5 F). EdU incorporation can also be used to determine the neurogenic rate of a population of neural progenitors, as EdU is retained by the neuronal progeny upon exit from the cell cycle. In this case, an EdU pulse is given to the progenitors on a certain day and the percentage of cells co-labelled with EdU and a post-mitotic neuronal marker is analysed a few days later (Otani et al., 2016). This approach was used to analyse the production of neuronal cells in CYFIP1tg and control cultures. Co-staining for NEUN and EdU 5 days after the EdU pulse was given revealed a significantly lower proportion of EdU⁺ NEUN⁺ cells in CYFIP1tg cultures than in the control ones (Fig. 4.5 C and G), suggesting that CYFIP1tg progenitors have a higher tendency to proliferate than to undergo terminal differentiation into neurons.

This was further supported by flow-cytometry analysis of the control and CYFIP1tg cells stained with the DNA dye, Dapi. The amount of DNA present is indicative of whether the cell is in the G0/G1, S or M phase of the cell cycle (Fig. 4.5 H and I). Such analysis showed that at day 35 the majority of control cells had exited the cell cycle, while a significant proportion of CYFIP1tg cells was still in the S and M phases (Fig. 4.5 K). However, no significant differences were found when the same analysis was performed 1 week earlier (Fig. 4.5 J), confirming that proliferating progenitors behave in similar way at early stages of differentiation but persist for longer in CYFIP1tg cultures.

This over-proliferation of CYFIP1tg PAX6⁺ progenitors was also reflected by the delayed appearance of TBR2⁺ intermediate progenitors and of the typical cortical markers for deep and superficial layers (Fig. 4.6). The transcription factors TBR2, TBR1, CTIP2 and SATB2 appear in this order during in vivo corticogenesis and this pattern is maintained in in vitro differentiation (Espuny-Camacho et al., 2013; Yichen Shi et al., 2012).
The percentage of TBR2 progenitors was significantly lower in CYFIP1tg cultures than in controls at day 30, but no difference was present at days 40 or 50. Both control and CYFIP1tg cultures were able to produce all types of pyramidal neurons, but significant differences were present in the numbers of TBR1+, CTIP2+ and SATB2+ cells at day 30. There were no differences at later stages, which is consistent with a delayed exit from the cell cycle of CYFIP1tg progenitors (Fig. 4.6 C).

Together, these data provide evidence that CYFIP1 overexpression promotes expansion of the progenitor pool, at the expense of delayed neurogenesis, during *in vitro* cortical differentiation.
Figure 4.5 Overproliferation of neural progenitors caused by CYFIP1 overexpression.

(A, D) Immunofluorescent staining and quantification of PAX6+ cortical progenitors (red) as a percentage of the total number of cells, marked by Dapi staining (blue). (B, E) Immunofluorescent staining and quantification of NEUN+ post-mitotic neurons (green). Nuclei were counterstained with Dapi (blue). (C) Double staining for EdU (red) and NEUN (green), 5 days after EdU incubation (2hours). All scale bars = 50 μm. (F) Quantification of EdU+ cells at day 30 and 40 of differentiation, right after EdU incubation (2hours). (G) Quantification of EdU+/NEUN+ cells at day 40 of differentiation, 5 days after EdU incubation (2hours). (H, I) Flow-cytometry histogram of Dapi+ cells at day 35 of in vitro differentiation. The intensity of Dapi fluorescence allows to divide the cell population into three sub-populations, corresponding to the different cell cycle phases, G0/G1, S and M. (J, K) Percentage of cells in the different phases of the cell cycle at day 28 and 35 respectively. Quantitative data relative to control and CYFIP1tg lines were compared by t-test (*p<0.05, **p<0.01, ***p<0.001) (N=3, all experiments were performed on the CYFIP1tg clone#5). All data are represented as mean ± SD.
Figure 4.6 Delayed appearance of cortical layer markers in CYFIP1g cultures.

Immunostaining for TBR2, TBR1, CTIP2 and SATB2 in CYFIP1g and control (H7) cultures at day 30 (A) and 50 (B) of in vitro-cortical differentiation. Nuclei are labelled by Dapi (blue). Scale bars = 50 μm. (C) Quantification of TBR2, TBR1, CTIP2 and SATB2 in CYFIP1g and control (H7) cultures at day 30, 40 and 50 of differentiation. For CYFIP1g, data represent an average of two different cellular clones (CYFIP1g clone #3 and clone#5, N=2 experiments for each clone). The percentage of positive cells in overexpressing and control lines were compared using Two-way ANOVA (*p<0.05). Data represent mean ± SD.
4.2.4 CYFIP1ko cortical progenitors undergo premature neuronal differentiation

The analysis of the proliferative behaviour of CYFIP1ko neural progenitors revealed an opposite trend to that of the CYFIP1tg cells. Around day 20, in control iCas9 cultures, a substantial population of NEUN+ neurons began to emerge. However, at the same time, CYFIP1ko cultures already contained a significantly higher proportion of neurons and, accordingly, a much smaller pool of PAX6+ progenitors (Figure 4.7A-D).

A pulse-chase EdU incorporation experiment was performed to analyse the neurogenic rate of the progenitors present at this stage. EdU was given for 2 hours at day 20 and cells co-labelled for EdU and NEUN were quantified 5 days later. More EdU+ neurons were found in the CYFIP1ko cultures than the controls, suggesting that more CYFIP1ko progenitors had given rise to neurons during the 5 day period (Fig. 4.7 E, F). Consistently with this, the proportion of cells expressing TBR1 and CTIP2 at day 20 was higher for CYFIP1ko cells than for the parental line. At day 30, more CTIP2+ cells were found in CYFIP1ko cultures and SATB2 was also significantly increased. However, by day 40, the quantification of these cortical neuronal subtypes did not show any difference (Fig. 4.7 G-I).

Several cell cycle and proliferation markers were also analysed, to gain a more complete understanding of the phenotype caused by CYFIP1 loss. At day 20, when more neurons are known to be present in CYFIP1ko cultures, the number of KI67+ cells, as well as the fraction of proliferating cells detected by flow-cytometry, were significantly smaller (Fig. 4.8 A, C). At the same time, a higher proportion of neuronal cells were found to be positive for p27KIP1, a cyclin-dependent kinase inhibitor, the upregulation of which promotes neuronal differentiation (Nguyen et al., 2006) (Fig. 4.8 B).

Overall, these data demonstrate that loss of CYFIP1 is associated with a reduced proliferative ability of cortical progenitors, which results in early loss of the self-renewing pool and, ultimately, premature neuronal differentiation.
Figure 4.7 Increased neurogenic ability of CYFIP1ko cortical progenitors.

Immunostaining for PAX6 (red) (A) and NEUN (green) (C) in CYFIP1ko and control (iCas9) cultures at day 20 of in vitro cortical differentiation. (B, D) Quantification of PAX6+ and NEUN+ cells in CYFIP1ko and control (iCas9) cultures at day 20 of differentiation. (E, F) Staining for NEUN (green) and EdU (red) and quantification of double-labelled cells at day 25, 5 days after EdU pulse. (G-I) Immunofluorescent staining for TBR1 (green in G), Ctip2 (green in H) and SATB2 (red in H) at day 20 and quantification of these markers (I) at day 20, 30 and 40 of in vitro differentiation. All nuclei are counterstained with Dapi (blue). All Scale bars = 50 μm. Quantification data in B, D and F were analysed by t test. Quantification data in I were compared by two-way ANOVA (*p<0.05, **p<0.01). All data are represented as mean ± SD and were collected from 3 independent experiments.
Figure 4.8 Reduced proliferation of CYFIP1ko progenitor pool.
Immunostaining and quantification for Ki67 (green in A) and P27kip1 (red in B) in CYFIP1ko and control (iCas9) cultures at day 20 of in vitro cortical differentiation. All nuclei are counterstained with Dapi (blue). Scale bars = 50 μm. (C) Flow-cytometry histogram of Dapi^+ cells at day 20 of in vitro differentiation and quantification of the percentage of cells in G0/G1, S and M phases of the cell cycle, on the basis of Dapi content. Data represent mean values from 3 independent experiments ± SD. Significant differences were analysed by t-test (*p<0.05, **p<0.01).

4.3 Discussion

Human ESCs with an excess (CYFIP1tg) and a loss of CYFIP1 (CYFIP1ko) were differentiated into cortical pyramidal neurons to investigate the consequences of abnormal CYFIP1 levels in the context of human cortical development. Neither the overexpression nor the loss of CYFIP1 appeared to compromise the efficiency of neuronal induction or the specification of the correct positional identity of the differentiated cells, as the percentage of neuronal progenitors expressing
dorsal telencephalic markers were comparable between both CYFIP1-mutant lines and the respective parental controls.

One hallmark of early stage in vitro cortical differentiation is the appearance of neural rosettes, radial arrangements of NPCs with very similar characteristics to the RGCs present in the VZ of the embryonic cortex, such as the presence of intracellular nuclear migration (Banda et al., 2015; Elkabetz & Studer, 2009; Ziv et al., 2015). Neural rosettes formed by both CYFIP1tg and CYFIP1ko NPCs were morphologically different from those formed by controls, suggesting that CYFIP1 is fundamental for the self-organisation of NPCs in appropriately polarised structures. The differentiation of CYFIP1tg hESCs into a 3D model of cerebral organoids provided further evidence towards this conclusion.

Aberrant neural rosette formation was previously reported by Yoon and colleagues in 15q11.2-deleted iPSCs and CYFIP1KD cells (Yoon et al., 2014). Notably, the “scattered” expression pattern of rosette apical markers described in their study was very similar to that shown here for CYFIP1tg and CYFIP1ko NPCs. The same authors also reported that CYFIP1 reduction was associated with a lower level of WAVE2 protein, a fundamental component of the WRC, which regulates actin polymerisation (Takenawa & Suetsugu, 2007). For this reason, they argued that the destabilisation of this complex, caused by low CYFIP1 levels, was the cause of the defective neural rosettes (Yoon et al., 2014).

However, actin is only one of the several proteins present at the apical side of these structures, which form closely interconnected system. Of these proteins, NCAD was shown to also be a key FMRP target, in a mouse model of FXS, during embryonic cortical development (La Fata et al., 2014). Therefore, malfunction of the CYFIP1-FMRP complex could alter the normal levels of NCAD translation and result in the formation of faulty apical junctions. Hence, each of these mechanisms, or a combination of the two, could explain CYFIP1 contribution to the neural rosette phenotype observed.

Both the signalling molecules present within the rosettes and their cytoarchitectural organisation play an important role in the regulation of the neurogenic capabilities of the NPCs forming these structures (Banda et al., 2015; Ziv et al., 2015). Consistent with this, CYFIP1tg and CYFIP1ko cells showed differences in their proliferative behaviour and neurogenic rate. EdU incorporation experiments and flow cytometry analysis of the cell cycle profile revealed that overexpression of CYFIP1 was associated with increased progenitor proliferation accompanied by a delayed generation of post-mitotic neurons, while CYFIP1 loss had the opposite effect. Nonetheless, neurons of deep and upper cortical layers were formed by all the cell lines analysed, confirming
that the differences observed concerned only the kinetic of differentiation but not the fate of the cells generated.

Importantly, neurogenesis mechanisms have been shown to be affected also in several FXS models. Loss of FMRP leads to increased proliferation and neuronal production in Drosophila *larvae* and increased generation of TuJ+ cells was observed in neurospheres derived from Fmr1-/- mouse embryos and FXS foetuses (Callan et al., 2010; Castrén et al., 2005). Moreover, loss of Fmr1 was shown to impact adult neurogenesis, by increasing proliferation and affecting neuronal differentiation and survival (Luo et al., 2010). In this model it was also demonstrated that Cyclin D1, CDK4, and GSK3β transcripts are FMRP targets and that Fmr1-/- NPCs have reduced WNT signalling levels (Luo et al., 2010). Recently, the correlation between FMRP and GSK3β was confirmed in NPCs derived from human FXS iPSCs, but it could not be verified in neurons, leading the authors to postulate that this association may be stage and species-specific (Telias et al., 2015). These data support the hypothesis that the FMRP-CYFIP1 complex may be controlling the translation of some key neurogenesis regulators and be ultimately responsible for the differentiation phenotype observed in CYFP1tg and CYFIP1ko cultures.

Deficits of proliferation and differentiations have been described by several other *in vitro* models of SZ and ASD. For instance, two separate studies on iPSCs derived from individuals with idiopathic autism reported enhanced proliferation of neural progenitors, during both monolayer and organoid differentiation (Marchetto et al., 2016; Mariani et al., 2015). More specifically, the defects observed were relative to a decrease in cell cycle length, which caused a fast expansion of the progenitor population. Interestingly, in one of these studies, the increased proliferation was also shown to correlate with the degree of macrocephaly present in the donor patients (Marchetto et al., 2016). This type of alterations could be relevant in the context of the structural abnormalities observed in the brains of patients affected by psychiatric disorders, especially ASD. Indeed, the most commonly reported alterations in ASD subjects are larger brain volume and excess of cortical neurons, as well as abnormalities in cortical minicolumns (Casanova et al., 2006; Courchesne et al., 2011; McKavanagh et al., 2015; Sacco et al., 2015). Moreover, mild brain structural abnormalities have been specifically reported for 15q11.2 CNV-carriers (Stefansson et al., 2014; Ulfarsson et al., 2017).

In conclusion, the phenotypes detected in CYFIP1tg and CYFIP1ko neural cells confirm the hypothesis of an important role played by CYFIP1 during embryonic cortex development and fit in the broader scenario of proliferation and differentiation deficits as a possible common cellular pathology of neurodevelopmental psychiatric disorders.
5 Molecular basis of CYFIP1tg and CYFIP1ko phenotypes

5.1 Introduction

Data presented in the previous chapter showed that the overexpression and loss of CYFIP1 affect the formation of neural rosettes and the kinetics of neurogenesis during in vitro cortical differentiation. More precisely, increased and decreased levels of CYFIP1 had opposite effects on the proliferative and neurogenic ability of the cortical progenitor pool, suggesting that this phenotype is specifically caused by CYFIP1 changes. Therefore, investigation of the mechanism underlying this defect should provide important insight into the role of CYFIP1 in the context of cortex development.

In vitro, neural rosette NPCs experience very similar signals to those present in the VZ during mouse corticogenesis (Banda et al., 2015). A large number of pathways have been implicated in the regulation of neuronal differentiation during mouse cortex development. Of these, Notch and WNT pathways have been reported to be linked to CYFIP1 or its partner FMRP in various models.

The Notch pathway, activated by Delta-like1 ligand (Dll1), is well known for promoting the maintenance of radial glia cells (RGCs) in an undifferentiated state. Upon activation, the intracellular domain of the Notch receptor is released and translocated to the nucleus to activate its target genes, HES1 and HESS (Androutsellis-Theotokis et al., 2006; Imayoshi et al., 2010; Schroeter, Kisslinger & Kopan, 1998). These transcription factors repress the expression of pro-neural genes and prevent the initiation of neuronal differentiation (Bertrand, Castro & Guillemot, 2002; Ohtsuka et al., 1999). Pro-neural genes and Dll1 are expressed within the same cells, making them able to undergo neuronal differentiation whilst inhibiting the same fate in neighbouring cells, in a process known as “lateral inhibition” (Kageyama et al., 2008). Aberrant NOTCH expression was detected during in vitro neuronal differentiation of FXS iPSCs (Telias, Segal & Ben-Yosef, 2013). In addition, CYFIP1 was shown to be a NOTCH target in human keratinocytes (Dziunycz et al., 2017).

WNT signalling also plays a fundamental role in this context, as it regulates both patterning and neurogenesis of the developing telencephalon. Here, WNT molecules are expressed in the cortical hem and are necessary to maintain the dorsal identity of pallial cells (Backman et al., 2005). In addition, WNT has been shown to promote self-renewing symmetric divisions of cortical precursors of the VZ and its down-regulation is needed to initiate differentiation (Chenn & Walsh, 2002; Pöschl et al., 2013; Woodhead et al., 2006). Activation of canonical WNT signalling is mediated by stabilisation of β-catenin, which would be otherwise degraded by a destruction complex, formed by glycogen synthase kinase 3β (GSK-3β), adenomatous polyposis coli (APC), axis inhibition protein.
(Axin) and serine/threonine kinase casein kinase 1α (CK1α) (Moon et al., 2004). The accumulation of active β-catenin in the cytoplasm results in its translocation into the nucleus and the transcripational activation of WNT target genes (Moon et al., 2004). Beta-Catenin is also a binding partner of N-Cadherin (NCAD) and together, they form complexes present at the apical side of the VZ, which are also connected to the actin cytoskeleton by α-catenins (Nelson, 2004; Stocker & Chenn, 2015). Disruption of this type of complexes has been shown to severely affect the differentiation and migration of neural precursors (Stocker & Chenn, 2009; Zhang et al., 2010, 2013). In addition to these, more signalling pathways contribute to the regulation of cortical neurogenesis, making the dissection of the mechanisms affected by CYFIP1 challenging.

RNA-seq (RNA-sequencing) offers an absolute genome-wide quantification of all transcripts present in a sample of cells. This technique overcomes a number of limitations associated with other methods of transcriptomic analysis, such as microarrays, which are limited to the detection of known sequences and have a lower dynamic range (Wang, Gerstein & Snyder, 2009). RNA-seq was used to gain insight into the mechanisms driving the changes in neurogenesis observed in CYFIP1tg and CYFIP1ko cells and to allow a broader and unbiased characterisation of the pathways dysregulated in these cells. Importantly, this technique is increasingly used to characterise animal and cell-based models of ASD and SZ, as well as to analyse patients’ post-mortem brains, in the hope of uncovering the complexity underlying these disorders. Therefore, the comparison of the RNA-seq results obtained from CYFIP1tg and CYFIP1ko neural cells to existing datasets from ASD and SZ models will also allow to explore the presence of shared or convergent mechanisms, providing useful insights into the biology of these diseases.

5.2 Results

5.2.1 Molecular pathways affected by overexpression and loss of CYFIP1

RNA-seq was performed on control and modified cells at three stages of differentiation. These were a neuroepithelial stage (day 10 for H7 and CYFIP1tg, day 5 for iCas9 and CYFIP1ko lines), a NPC stage (day 18 for H7 and CYFIP1tg cells, day 12 for iCas9 and CYFIP1ko) and a neuronal stage (day 45 for H7 and CYFIP1tg, day 35 for iCas9 and CYFIP1ko cells). After the necessary quality control steps on the raw data (performed by Daniel Cabezas de la Fuente), differential gene expression analysis was performed to compare each CYFIP1-modified line to their respective controls, for each time point (Anders & Huber, 2010). Differentially expressed genes (DEGs) with an absolute fold change value of at least 1.5 were used for KEGG (Kyoto encyclopaedia of genes and genomes) gene set enrichment analysis (Kanehisa et al., 2012; Yu et al., 2012). The most significant pathways affected in both the overexpression and knock-out datasets are shown in figure 5.1.
Firstly, the KEGG analysis revealed that most of the normalised enriched scores (NES) follow opposite directions for the overexpression and knock-out data sets. Because of this, these pathways can be considered to be specifically affected by changes in CYFIP1 levels, independently of the genetic background of the parental lines, the integration site of the overexpression plasmid or the effects of off-target mutations.

Moreover, several of the pathways identified by this approach, such as WNT, PI3K, cell cycle and focal adhesion, play an important role in the regulation of neurogenesis. Therefore, they represent interesting candidates to investigate in the context of the phenotype presented in Chapter 4. Together with these, oxidative phosphorylation, which has been previously reported to be altered by several in vitro studies on psychiatric disorders, also appeared to be significantly affected in CYFIP1tg and CYFIP1ko cells at all time points (Ebrahimi-Fakhari et al., 2016; Mertens et al., 2015; Robicsek et al., 2013). For these reasons, neurogenesis and mitochondria-related genes were chosen for follow-up experiments on CYFIP1tg and CYFIP1ko lines.

The same set of RNA-seq results is also being analysed in depth by Daniel Cabezas de la Fuente, another PhD student in our group. His work includes the characterisation of several aspects linked to the pathways illustrated in figure 5.1 (e.g. Calcium signalling), as well as the comparison between the transcripts altered by CYFIP1 mutations with existing databases of psychiatric disease-risk genes.
Figure 5.1 Biological processes and signalling pathways significantly affected by altered CYFIP1 levels.

The bar graph represents the results of KEGG enrichment gene set analysis carried out on the DEGs of each modified line compared to the respective parental line, for each time point. Only DEGs with an absolute fold change value of 1.5 were used for this analysis. A p value cut-off of 0.05 and a p-adjusted cut-off of 0.1 were also set. (NES, normalised enrichment score). This image was provided by Daniel Cabezas de la Fuente.

5.2.2 Mechanisms underlying the neurogenesis phenotype in CYFIP1tg and CYFIP1ko neural cells

(i) Neurogenesis-related genes affected by CYFIP1 expression levels

To gain further insight into the genes responsible for the abnormal neurogenesis in CYFIP1tg and CYFIP1ko cells, DEGs associated with the GO (Gene Ontology) term “neurogenesis” (GO:0022008) were selected from the NPC and neuronal data sets. For both these time points, the neurogenesis-related genes affected by both an increase and a reduction of CYFIP1 were chosen for subsequent analysis (Fig. 5.2 A and B).

This analysis revealed that, in the context of neurogenesis, the most substantial differences (highest fold changes and p-adjusted values) were observed at the NPC stage for the CYFIP1tg-control data set and at the neuronal stage for the CYFIP1ko-control dataset. For both time points and datasets,
CYFIP1 and CYFIP2 transcripts were found to be significantly changed. They were upregulated in CYFIP1tg samples and downregulated in CYFIP1ko samples. In addition, WNT signalling genes, such as frizzled (FZD) receptors and WNT ligands, were highly represented in this data set, especially at the NPC stage. At the same stage, several Cadherins transcripts were also found to be altered (CDH1/2/4/11), as expected on the basis of the KEGG results.

CYFIP1tg and CYFIP1ko NPCs showed opposite changes in many, but not all, genes expressed in the developing cortex. These included dopachrome tautomerase (DCT), expressed by VZ progenitors between E10.5 and E17.5 (Jiao et al., 2006), doublecortin (DCX), marking immature neuronal cells (Gleeson et al., 1999), LHX2, which is necessary for βCAT-driven cortical progenitor proliferation (Hsu et al., 2015) and forkhead box O3 (FOXO3), a downstream effector of PI3K-AKT that regulates proliferation and cell cycle exit of embryonic and adult neural stem cells (Paik et al., 2009; Renault et al., 2009; Vezzali et al., 2016). Strong upregulation of FOXG1 was also detected in CYFIP1tg NPCs (Fig. 5.2 A).

At the neuronal time point, transcripts that were found to be significantly altered included many transcription factors involved in the fate specification of cortical neurons, Empty Spiracles Homeobox 1 (EMX1/2), Neurogenic differentiation 1/4 (NEUROD1/4), RELN, TBR1 and FEZF2 (Martynoga et al., 2012; Mattar et al., 2008) (Fig. 5.2 B). Moreover, at this stage, DEGs related to synaptic maturation (Erb-B2 Receptor Tyrosine Kinase 4 or ERBB4, Discs Large MAGUK Scaffold Protein 4 or DLG4) and axonal growth and migration (Semaphorin 5A/B or SEMAS5A/B) appeared to be also affected (El-Husseini et al. 2000; Li et al., 2007; Purohit et al., 2014; Yoshida, 2012). Finally, increased and decreased CYFIP1 levels also caused alterations in the expression of several psychiatric disorder risk genes, such as CNTNAP2, NXRN3 , SHANK1/3 and TSC at the progenitor stage and NRXN1 and SHANK3 in neurons (Alarcón et al., 2008; Costales & Kolevzon, 2015; Curatolo et al., 2015; Gong & Wang, 2015; Reichelt et al., 2012).

Importantly, a number of DEGs present in CYFIP1tg and CYFIP1ko datasets, such as WNT5A, FOXG1 and CDH2, did not show the same direction of fold change. This is in part expected, as CYFIP1, contrarily to transcription factors, does not affect gene transcription directly. For instance, some of the observed changes could be due to compensatory mechanisms activated by the cells, rather than being a direct consequence of CYFIP1 alterations. Nevertheless, variations in these genes prove the presence of a significant dysregulation of the pathways to which they belong.
Figure 5.2 Neurogenesis-related genes affected by abnormal CYFIP1 levels.

Bar graphs illustrating the DEGs associated with the GO term “neurogenesis” (GO:0022008) for CYFIP1tg and CYFIP1ko NPCs (A) and neurons (B). The height of the bars is proportional to the gene expression fold change (FC) in respect to the appropriate control parental line. The colour represents the significance of the FC, expressed as -LOG_{10} of the p adjusted value (padj).
Dysregulated Cadherin-Catenin signalling in CYFIP1tg and CYFIP1ko cells

The analysis of the DEGs present in CYFIP1tg and CYFIP1ko data sets revealed a high representation of genes coding for members of the Cadherin family and genes related to WNT and PI3K-AKT pathways (Fig. 5.1 and 5.2). A point of convergence of these biological processes is represented by β-catenin (βCAT). The stabilisation of this protein in VZ progenitors requires the activation of WNT and AKT pathways and includes a strong contribution from NCAD, the levels of which correlate positively with βCAT activity (Zhang et al., 2010). If affected, this mechanism could provide an explanation to the altered neurogenesis caused by CYFIP1 changes. To verify this, the amount of NCAD and βCAT, both active and primed for degradation, were analysed by western blot at three time points between the neural progenitor and neuronal stages.

The levels of NCAD were increased in CYFIP1tg progenitors (day 18) and early neurons (day 35) compared to those derived from the parental line. Conversely, a lower amount of NCAD was detected in CYFIP1ko neural cells than in those derived from the parental line, at day 12 and 20. The differences in NCAD expression between CYFIP1tg or CYFIP1ko and their respective controls were visibly reduced at the last time point analysed (Fig. 5.3 A, B, E, F). Phosphorylation of βCAT by GSK3-β of the residues serine 33 and serine 37 (p-ser33/37) primes this protein for degradation, while phosphorylation of serine 552 (p-ser552) enhances its transcriptional activity (Fang et al., 2007; Liu et al., 2002; Zhang et al., 2013). The amount of inactive βCAT, represented by the ratio between p-S33/37 βCAT over the total was reduced in CYFIP1tg early and late neurons compared to levels observed in the parental line (Fig. 5.3 A, C, E, G). Conversely, the quantity of active βCAT, expressed by the ratio between p-S552-βCAT/total βCAT was found to be generally higher in CYFIP1tg neural cells than in control ones. CYFIP1ko displayed overall an opposite pattern of βCAT phosphorylation, with a generally higher amount of p-S33/37-βCAT/ total βCAT and a lower ratio of active βCAT (p-ser552) in neurons of day 20 and 35 (Fig. 5.3 A, D, E, H). Overall, these results are consistent with an increased activation of βCAT signalling in CYFIP1tg neural cells and a downregulation of βCAT signalling in CYFIP1ko cells, compared to their respective controls.
Western blots for NCAD, phosphorylated (p-S33/37 and p-S552), total βCAT and GAPDH in CYFIP1tg (A) and CYFIP1ko (E) cultures at three time points. These were defined as NPCs, early and mature neurons and corresponded to day 18, 35 and 45 for CYFIP1tg cells and to day 12, 25 and 35 for CYFIP1ko respectively. Bar graphs representing the quantification of NCAD (B, F), p-βCAT (S33/36) over the total βCAT (C, G) and p-βCAT (S552) over the total βCAT (D, H). The intensity of each band was measured with ImageJ and normalised to the intensity of GAPDH band on the same blot. Data represent the mean value ± SD from two technical replicates of two independent experiments. Expression values were compare via two-way ANOVA, followed by Tukey post-hoc test (*p<0.05).
Manipulation of N-CADHERIN-mediated activation of β-catenin rescues the differentiation of CYFIP1tg progenitors

The data presented above demonstrate a higher activation of the NCAD-βCAT signalling axis in CYFIP1tg neural cells than in those derived from the parental line H7. According to mouse developmental data, this alteration could be responsible for the abnormal expansion of PAX6⁺ progenitors in the overexpressing cultures. To verify this, late progenitors of day 30 CYFIP1tg and H7 cultures were exposed to several treatments aimed at reducing βCAT signalling activity and, ultimately, rescuing the delayed neurogenesis. These treatments included direct blockage of NCAD activity or, downstream of this, inhibition of WNT and AKT pathways, alone or in combination. The small molecules used for WNT and AKT inhibition were XAV939 and API2, respectively (Huang et al., 2009; Yang et al., 2004). NCAD blocking was achieved by incubating the cells with a specific anti-NCAD antibody, which was used for the same purpose by several other studies on cultured mouse and human neural cells (Iefremova et al., 2017; Zhang et al., 2013). A control group exposed only to the vehicle DMSO was also included for both CYFIP1tg and H7 cells. After the treatments, the cells were left to mature for 5 additional days, before being fixed and stained for PAX6. The experimental scheme is illustrated in figure 5.4 A and B.

The analysis of DMSO-treated cultures at the end of the experiment showed that 40% of NPCs in CYFIP1tg cultures were PAX6⁺, compared to less than 20% in the control, confirming previous results indicating an expansion of the PAX6⁺ fraction in CYFIP1tg cultures. The percentage of PAX6⁺ progenitors was still higher also in CYFIP1tg cells treated with XAV939 or API2, than in control cells exposed to the same molecules, showing that inhibition of WNT or AKT signalling alone had only a small effect on the delayed neurogenesis. However, the combination of these two treatments effectively lowered the number of PAX6⁺ CYFIP1tg NPCs to a level that was not statistically different from that obtained from control cells that had received the same inhibitors. Moreover, the percentage of PAX6⁺ cells present in overexpressing cultures treated with XAV939 and API2 was comparable to the amount of PAX6⁺ cells present in H7 cultures in control conditions. Finally, also the application of NCAD-blocking antibody resulted in a reduction of PAX6⁺ progenitors in CYFIP1tg cultures to below 30%, with no significant difference from the H7 cultures exposed to the same antibody (Fig. 5.4 C, D). In summary, the use of the NCAD-blocking antibody and of XAV939 and API2 together were the only two treatments able to normalise the levels of PAX6⁺ NPCs present in CYFIP1tg cultures, with the combination of WNT and AKT inhibition being the most successful.

Dysregulation of the basal level of WNT signalling has been associated with an alteration of cells fate during in vitro differentiation, reflected in an imbalanced expression of dorsal and ventral telencephalic markers (Srikanth et al., 2015). Moreover, incubation of stem cells-derived neural
progenitors with the WNT inhibitor XAV939 is used by several protocols with the aim to ventralise cells during the derivation of GABAergic neurons (Maroof et al., 2013; Nicoleau et al., 2013). Therefore, to ascertain that the observed reduction of PAX6+ cells was not due to a switch from dorsal to ventral cortical fate following XAV939 and API2 treatment, immunostaining for the neurotransmitter GABA was performed to quantify the number of inhibitory cells. No significant differences were found in the number of GABA+ cells in the vehicle treated control and CYFIP1tg cultures with or without WNT and AKT inhibition (Fig. 5.5).

In conclusion, these data provide evidence that the inhibition of WNT and AKT signalling, can efficiently rescue the over-proliferation of PAX6+ cortical progenitors caused by high levels of CYFIP1, without altering the identity of the treated cells.
Figure 5.4 Reduction of βCAT activity downstream of NCAD rescues the delayed neurogenesis of CYFIP1tg cultures.

(A) Scheme of βCAT stabilisation in VZ progenitors and the molecules used to modulate this signalling. NCAD stimulates AKT, which phosphorylates βCAT at the position S552, and WNT signalling, which blocks GSK3β preventing its phosphorylation of βCAT. These pathways can be altered by blocking NCAD function with a specific antibody or by using AKT and WNT antagonists (API2 and XAV respectively). Image modified from Zhang et al., 2013. (B) Experimental scheme illustrating the application of the different treatments between day 30 and 33 of differentiation on CYFIP1tg and H7 control cultures. (C) PAX6 staining, in red, at day 38, 5 days after the treatments, in each condition and cell line. Nuclei were counterstained with dapi (blue). (D) Quantification of the percentage of PAX6+ cells against the total of dapi+ nuclei. Data represent the mean value ± SD of three technical replicates from one experiment. The mean value for each treatment and cell line were compared with two-way ANOVA, followed by Tukey post-hoc test (*p<0.05, **p<0.01, ***p<0.001).
Figure 5.5 Quantification of GABAergic cells in CYFIP1tg and control cultures.
(A) Staining for the neurotransmitter GABA at day 38, at the end of the rescue experiment. The staining was carried out in control cells in control condition (DMSO) and in CYFIP1tg cells exposed to control (DMSO) and XAV+API2 treatments. (B) Quantification of GABA$^+$ cells in the same experimental conditions. Data represent the mean value ± SD of three technical replicates from one experiment. No significant differences were found.

5.2.3 Mitochondrial alterations in CYFIP1tg and CYFIP1ko neural cells

(I) Mitochondria-related genes affected by CYFIP1 changes

The KEGG pathway enrichment analysis presented in paragraph 5.2.1 highlighted a significant alteration of mitochondrial function in neural cells derived from CYFIP1tg and CYFIP1ko hESCs. To explore this aspect further, the DEGs related to mitochondrial function were analysed in more detail at the NPCs and neuronal stage.

Many of the differentially expressed mitochondria-related genes, which were present in both the overexpression and knock-out datasets, code for enzymes forming the respiratory complexes. These include NADH dehydrogenase, also known as respiratory complex I (NDUFA genes), Succinate dehydrogenase complex subunit D (SDHD) that constitute part of Complex II, Ubiquinol-Cytochrome C Reductase or Complex III (UQCR genes), the Cytochrome C oxydase or Complex IV...
(COX genes) and the ATP synthase or Complex V (ATP genes). Most of these transcripts show opposite direction of FC in CYFIP1tg and CYFIP1ko at the NPCs stage (Fig 5.5 A), a pattern that was not maintained at a later time (Fig. 5.5 B). The genes showing an opposite FC between CYFIP1tg and CYFIP1ko lines, at both time points, were the mitochondrial encoded genes $MT\text{-}ND2/3/4/5$, $MT\text{-}CO2$ and $MT\text{-}APT6/8$, which code for subunits of the respiratory complex I, II and V, respectively. Other transcripts that appeared to be affected by CYFIP1 levels were related to mitochondrial dynamics, such as fission, fusion and transport. These included genes for Fission, Mitochondrial 1 ($FIS1$), Mitofusin 1/2 ($MFN1/2$), Misato 1 ($MSTO1$), Mitochondrial fission regulator 2 ($MTFR2$) and Mitochondrial dynamin like GTPase ($OPA1$).

Together, these data show that the expression of many genes involved in the regulation of mitochondria metabolism and dynamics in neural cells is significantly affected by changes in CYFIP1 levels.
Figure 5.6 Mitochondria-related genes affected by abnormal CYFIP1 levels.
Bar graphs illustrating the DEGs associated with oxidative phosphorylation and mitochondria dynamics in CYFIP1tg and CYFIP1ko NPCs (A) and neurons (B). The height of the bars is proportional to the gene expression fold change (FC) in respect to the appropriate control parental line. The colour represents the significance of the FC, expressed as -\(\text{LOG}_{10}\) of the \(p\) adjusted value (padj). Genes were taken from the Human MitoCarta 2.0 (Calvo et al., 2016).
Data presented above suggest the presence of significant changes in mitochondrial dynamics caused by altered levels of CYFIP1. To validate this, CYFIP1tg and CYFIP1ko NPCs and neurons were incubated with Mitotracker, a reagent that selectively labels the mitochondria (Cataldo et al., 2010; Leonard et al., 2015) (Fig. 5.6 A). The cells were then fixed and stained with Dapi and analysed on a high-content screening (HCS) platform for detection of intracellular “spots”, which represent Mitotracker-labelled mitochondria (Ghosh et al., 2005; Leonard et al., 2015). Representative images of Mitotracker-stained cells elaborated by the HCS software can be found in figure 5.6 B and C. This approach revealed the presence of significant differences in the number and size of spots. In particular, both the spots count and area in CYFIP1tg NPCs and neurons showed a tendency to be higher than those of the parental control cells, although statistical significance was only observed for the spots number at the progenitor stage (Fig. 5.6 D). Conversely, CYFIP1ko NPCs had less and smaller Mitotracker+ spots than the parental control group, but this difference was no longer detected in neuronal cells (Fig. 5.6 D).

Moreover, the high-content analysis highlighted the presence of significant differences in the size of nuclei, with an opposite trend, in CYFIP1tg and CYFIP1ko cells compared to their respective controls. CYFIP1tg nuclei were significantly smaller than those of the parental control neuronal cells, while CYFIP1ko cell nuclei were smaller than the control at the NPCs stage but bigger at the neuronal stage. Overall, these results demonstrate significant alterations to mitochondrial dynamics and nuclear morphology in neural cells with abnormal CYFIP1 expression levels.
Figure 5.7 Alterations of mitochondria dynamics in CYFIP1tg and CYFIP1ko NPCs and neurons.

(A) High magnification image of hESCs-derived NPCs stained with Mitotracker (red) and dapi, marking the nuclei (blue). (B) Representative field of mitotracker-stained cells elaborated by the HCS platform. Single nuclei are identified as primary object and the area surrounding each nucleus (delimited by green lines) is used for the detection of mitotracker+ spots. (C) Mitotracker+ spots (in yellow) detected in single cells. (D) Quantification of parameters relative to spots count, spots area and nuclei area, in arbitrary units provided by the HCS platform. Data are expressed as mean ± SD (N=3). Significance was analysed by t-test (*p<0.05).
5.3 Discussion

Whole-genome transcriptomic analysis was used to identify CYFIP1-mediated molecular changes in hESC-derived cortical cells with the aim to decipher the mechanisms responsible for the differentiation phenotype observed in CYFIP1tg and CYFIP1ko cells. However, the interpretation of these results should take into account the possibility that some gene changes may also reflect the different composition of the sample population, due to the altered differentiation kinetics of CYFIP1tg and CYFIP1ko lines presented in the previous chapter. A way to overcome this complication could be represented by sorting the cells samples prior to RNA-seq analysis, in order to separate progenitors and neuronal populations, or, after the RNA-seq, normalising mRNA changes to the expression levels of some stage-specific markers, such as Nestin and MAP2. Despite this, the transcriptomic analysis presented here helped to identify some important and novel pathways in which CYFIP1 is involved and provided important support to the phenotypic characterisation of CYFIP1-modified lines.

First of all, CYFIP1 and CYFIP2 were the most significant DEGs, showing up- and downregulation in NPCs and neurons derived from CYFIP1tg and CYFIP1ko ESCs respectively. These results provided independent confirmation of the increased CYFIP1 levels in the overexpressing cells and of the decreased CYFIP1 expression in the knock-out ones, in line with the characterisation of these engineered lines reported in Chapter 3. With regard to CYFIP2, the regulation of its expression is not well characterised and previous studies report different results on the levels of this gene following knock-out or knock-down of CYFIP1. The reduction in CYFIP2 transcript levels in CYFIP1ko neural cells is in agreement with the results reported by Abekhoukh et al., showing a downregulation of Cyfip2 mRNA in mouse primary neurons with knock-down of Cyfip1 and in the blood of patients carrying 15q11.2 deletion (Abekhoukh et al., 2017). CYFIP2 has been reported to be one of the top-ranked FMRP targets, while CYFIP1 itself did not appear in the same list (Darnell et al., 2011). This suggests that CYFIP2 translation could be regulated by the FMRP-CYFIP1 complex and that changes in CYFIP2 transcript may be a consequence of alterations of its protein levels, as part of a feedback mechanism. However, other published studies did not find any change in CYFIP2 expression following CYFIP1ko in human NPCs or in the brain of Cyfip1−/− mice (Bozdagi et al., 2012; Nebel et al., 2016). Future work on the regulatory network involving FMRP, CYFIP1 and CYFIP2 may help to explain these discrepancies.

Another aspect providing additional support to the validity of our model is the significant enrichment for genes related to the regulation of actin cytoskeleton and axonal guidance in both CYFIP1tg and CYFIP1ko datasets (Fig. 5.1). Alterations to these pathways have also been reported
by other studies on in vitro and in vivo models of CYFIP1 alterations (Nebel et al., 2016; Oguro-Ando et al., 2014; Pathania et al., 2014). Other pathways highlighted by this analysis include the PI3K/AKT, cAMP and calcium signalling. PI3K/AKT alterations have been reported in FXS and SZ and are known to be important for dendritic branching, synaptogenesis and spine formation (Cuesto et al., 2011; Gross et al., 2010; Jaworski, 2005; Kalkman, 2006; Zheng et al., 2012). Moreover, modulation of this pathway has been proposed for pharmacological intervention for both ASD and SZ (Enriquez-Barreto & Morales, 2016). Calcium and cAMP signalling are also fundamental in neuronal cells, especially in the regulation of synaptic activity and plasticity mechanisms (reviewed by Bading, 2013). Together, these alterations confirm the relevance of the gene network affected by CYFIP1 changes in the context of neurodevelopmental psychiatric disorders.

In light of the phenotypic characterisation presented in Chapter 4, the signalling pathways involved in the regulation of neurogenesis were prioritised for a more in-depth investigation. Many of the DEGs affected by both the increase and loss of CYFIP1 included WNT signalling genes and adhesion molecules such as Cadherin proteins. These, together with AKT, form an interconnected molecular system that has been previously shown to regulate the balance between proliferation and differentiation in the mouse cortex. According to this model, NCAD regulates βCAT stability by stimulating both WNT and AKT signalling, leading to an increase in βCAT-dependent transcriptional activation (Zhang et al., 2010, 2013). The same signalling mechanism has been recently shown to be conserved in human cortex development using forebrain organoid differentiation from iPSCs derived from individuals with Miller-Dieker syndrome, which causes microcephaly and seizures amongst other symptoms (Iefremova et al., 2017). Defects in the distribution of NCAD at the ventricular surface and in the proliferation of NPCs in Miller-Dieker organoids were demonstrated to be specifically linked to decreased activity of the NCAD-βCAT signalling axis and could be rescued by treatment with a WNT agonist (Iefremova et al., 2017). The same pathways were found to be altered in CYFIP1tg and CYFIP1ko cells, as the phosphorylation pattern of βCAT by GSK3β and by AKT was consistent with a hyper-activation of βCAT in overexpressing cells and a downregulation in knock-outs. Moreover, treatment with a combination of WNT and AKT inhibitors was successful in rescuing the over-proliferation of PAX6 progenitors in CYFIP1tg cultures, confirming that this signalling cascade is specifically altered by the excess of CYFIP1.

A likely mechanism by which CYFIP1 may be disturbing this pathway is through translational regulation of NCAD, AKT3 and GSK3-β that have all been reported to be FMRP targets (Darnell et al., 2011; La Fata et al., 2014). However, the fraction of transcriptionally active βCAT can also be affected by alterations to the actin cytoskeleton caused by malfunction of the WAVE complex of
which CYFIP1 is part. For instance, F-Actin accumulation in the nucleus has been shown to enhance the translocation of active \( \beta \text{CAT} \) to this cellular compartment (Yamazaki et al., 2016).

Dysregulation of WNT signalling has been reported by several studies on \textit{in vitro} models of psychiatric disorders, including ASD, SZ and BD (Brennand et al., 2011; Srikanth et al., 2015; Topol et al., 2015; P. Wang, Mokhtari, et al., 2015). Together with this, another aspect linking CYFIP1 alterations to a broader spectrum of neurodevelopmental disorders is represented by changes in FOXG1 expression levels. FOXG1 plays a fundamental role in brain development and mutations in its locus are associated with severe mental retardation and microcephaly (reviewed by Florian, Bahi-Buisson & Bienvenu, 2012). In the context of \textit{in vitro} models of psychiatric disorders, cells carrying \textit{DISC1} mutation were shown to have higher baseline WNT, together with downregulation of \textit{FOXG1} and downregulation of ventral forebrain markers (Srikanth et al., 2015). Conversely, upregulation of \textit{FOXG1} was reported in cerebral organoids derived from ASD iPSCs, a defect specifically associated with a shift towards a more ventral identity of the differentiated cells (Mariani et al., 2015). Similarly, cortical progenitors with modified CYFIP1 levels showed an increase in the amount of \textit{FOXG1} transcript, a change that was stronger for the overexpressing line. However, this was not accompanied by differences in the percentage of \textit{FOXG1}\(^+\) cells or changes in the dorso-ventral identity of the differentiated neurons, as the number of GABAergic cells in CYFIP1tg cultures was the same as in the controls. A reason for this may be that the increase observed in NPCs was only transient, as it was no longer present at the neuronal stage, when only a small but significant downregulation was detected by the RNA-seq analysis (data not shown).

Finally, RNA-seq analysis revealed the presence of significant alterations in the expression of genes regulating mitochondria dynamics and oxidative phosphorylation. Overall, genes located in the nuclear DNA and coding for components of the respiratory chain showed similar levels and direction of FC, consistent with a co-regulation of their expression, a process controlled by signals originating from the mitochondria, such as the intracellular levels of \( \text{Ca}^{2+} \), reactive oxygen species (ROS) and ATP (Reinecke et al., 2009; van Waveren & Moraes, 2008).

Changes in CYFIP1 levels seem to very specifically affect the expression of mitochondrially-encoded genes (\textit{MT-ATP6/8}, \textit{MT-CO2}, \textit{MT-ND2/3/4/5}), as these show completely opposite FC pattern in overexpressing and knock-out cells at both the time points analysed. It has been suggested that the number of mitochondrial DNA (mtDNA) molecules per cell could play a role in regulating the transcription of mitochondrially-encoded genes. This indicates that the gene expression changes observed in CYFIP1tg and CYFIP1ko cells could be explained, at least in part, by different amounts of mitochondria or mtDNA (Reinecke et al., 2009). Consistently with this, HCS analysis revealed
significant differences in several parameters related to the size and number of mitochondria between control NPCs and neurons and those with abnormal CYFIP1 levels.

The mechanisms regulating the number of mitochondria or the number of mtDNA molecules are not well characterised (Lee & Wei, 2005; Reinecke et al., 2009). Nonetheless, it has been reported that mitochondrial dynamics play an important role during neural differentiation. In the mouse developing cortex, the commitment to neuronal differentiation is accompanied by mitochondrial fragmentation and by an increase in physiological levels of ROS, at the same time as the cells also undergo a metabolic switch from anaerobic to aerobic metabolism (Khacho et al., 2015). In addition, ROS levels have been shown to contribute to cellular signalling post-translationally, maintaining an appropriate level of PI3K-AKT and Notch activation, which are fundamental for self-renewal of neural stem cells (Khacho et al., 2015; Le Belle et al., 2011). Therefore, the mitochondrial alterations present in CYFIP1tg and CYFIP1ko cells could contribute to the dysregulation of neurogenesis observed in these lines.

Importantly, mitochondrial abnormalities were also reported in neurons derived from BD iPSCs, which displayed high membrane potential and decreased mitochondria size (Mertens et al., 2015) and in those from SZ iPSCs, which showed altered mitochondrial respiration and distribution (Robicsek et al., 2013). Mitochondrial homeostasis and metabolism were also affected in in vivo and in vitro models of tuberous sclerosis syndrome (Ebrahimi-Fakhari et al., 2016). This aspect represents another point of convergence between the phenotype caused by CYFIP1 mutations and other models of neurodevelopmental psychiatric disorders.

Finally, the HCS analysis allowed the discovery of the presence of significant differences in the size of the nuclei in CYFIP1tg and CYFIP1ko NPCs and neurons. Similarly, Nebel and colleagues reported that nuclear size was affected also by CYFIP1KD, a phenotype that is consistent with alterations to the actin cytoskeleton (Jevtić et al., 2014; Nebel et al., 2016).

In conclusion, the changes in gene expression revealed by the RNA-seq analysis confirmed the presence of a dysregulation of neurogenesis and supported a role for WNT and the NCAD/βCAT axis in the phenotype cause by CYFIP1 alterations. These, together with the changes in mitochondria dynamics, demonstrate the relevance of the phenotypes caused by CYFIP1 alterations in the context of neurodevelopmental psychiatric disorders.
6 General Discussion

6.1 Summary of findings

In this thesis, I investigated the role played by CYFIP1 during cortical differentiation of hESCs. CYFIP1tg and CYFIP1ko hESCs were derived to examine the consequences of increased and decreased levels of this gene, as both deletions and duplication of the genetic region where CYFIP1 is located are found in psychiatric patients. NPCs derived from both lines showed deficits in the formation of neural rosettes and in the balance between proliferation and differentiation. In particular, the excess and loss of CYFIP1 had opposite consequences on this process. High levels of this gene were associated with an increased proliferative ability of the cortical progenitor pool, while the reduction of CYFIP1 was linked to premature exit from the cell cycle and neuronal differentiation. With the help of RNA-seq analysis, it was possible to confirm that the expression of a high number of genes involved in the regulation of neurogenesis was significantly affected in CYFIP1tg and CYFIP1ko cells. Follow up experiments, including analysis of the phosphorylation state of β-CAT and manipulation of the differentiation conditions, allowed to specifically link the neurogenesis phenotype with a malfunction of the NCAD-β-CAT signalling axis. In addition, the transcriptomic analysis revealed the presence of alterations to the oxidative phosphorylation system and to the regulation of mitochondria dynamics, which were confirmed using a platform for automated high-content screening. These findings uncovered a novel mechanism by which CYFIP1 participates in the regulation of neuronal differentiation and revealed new biological functions in which this gene is involved.

6.2 Convergent mechanisms in neurodevelopmental psychiatric disorders

6.2.1 Cadherins- and Catenins-related signalling

The analysis of the pathways dysregulated by altered levels of CYFIP1 provided interesting points of convergence with other in vitro and in vivo models of neurodevelopmental disorders. One of these aspects is represented by alterations to cadherin-dependent cell-cell adhesion. Mutations in genes of the Cadherins superfamily not only result in brain malformations, but they have also been found to increase the risk for several psychiatric diseases. For instance, deletions of the locus encoding CDH8 have been found in patients affected by ASD and learning disabilities (Pagnamenta et al., 2011) and SNPs in the proximity of CDH9 and CDH10 were identified in a ASD cohort by a GWAS study (Wang et al., 2009). In addition, SNPs within the sequence of CDH7 have been reported to increase the risk for SZ (Sklar et al., 2008; Soronen et al., 2010). These data have led, in the past,
to the formulation of a “cadherin hypothesis of schizophrenia”. According to this, the severe impact on brain development caused by alterations to the adhesion molecules system is responsible for an increased risk of psychiatric disorders (Yagi & Takeichi, 2000).

As discussed in previous chapters, disruption of WNT signalling has also been reported by several in vitro-based studies on psychiatric disorders (Brennand et al., 2011; Srikanth et al., 2015; Topol et al., 2015). In addition, several in vivo models of neurodevelopmental disorders have shown an association between alterations in WNT signalling and the presence of behavioural defects typical of ASD. A loss-of-function mutation in the dishevelled genes 1 and 3 (Dvl1−/−;Dvl3−/+) in mice led to early differentiation of basal progenitors and expansion of deep cortical layers and resulted in the development of repetitive behaviour and abnormal social interaction later in life (Belinson et al., 2016). Prenatal pharmacological activation of WNT signalling rescued both the structural and behavioural phenotypes (Belinson et al., 2016). Alterations of the balance between proliferation and differentiation in the embryonic cortex have also been reported in another mouse model, in which the ASD risk gene CDH8 was knocked-down (Durak et al., 2016). Histological and transcriptomic analysis of this model demonstrated the presence of cell cycle and differentiation defects in the developing cortex. Adult mice showed decreased exploratory behaviour and sociability. These defects were demonstrated to be caused by a reduction of WNT activity, as introduction of a stabilised form of β-Catenin in the embryonic cortex completely rescued the phenotype (Durak et al., 2016).

In humans, mutations involving WNT genes have been reported in a considerable number of psychiatric patients. These include SNPs in the sequence of WNTZ found in cohorts of ASD patients (Marui et al., 2010; Wassink et al., 2001) and CNVs involving the chromosomal locus containing FZD9, which are associated with serious developmental delay and ADHD (Merla et al., 2010; Sanders et al., 2011). A specific missense mutation in the WNT1 sequence, resulting in overactive WNT, has also been reported in ASD individuals (Martin et al., 2013). In addition, deleterious mutations of β-CAT have been found in SZ, ASD and ID patients (Levchenko et al., 2015; O’Roak et al., 2011; Tucci et al., 2014). Previous post-mortem studies also suggested alterations in the expression of WNT-related genes, such as WNT1 and β-CAT, in SZ brains (Cotter et al., 1998; Miyaoka, Seno, & Ishino, 1999).

A number of studies showed evidence for AKT alterations in psychiatric patients. Genetic variants of the AKT1 gene were found to be associated with SZ in populations of various ethnicities (Bajestan et al., 2006; Emamian et al., 2004; Ikeda et al., 2004; Schwab et al., 2005) and specific reduction of AKT1 was observed in the hippocampus and cortex of SZ patients (Emamian et al., 2004). More
recently, SNPs in the AKT3 locus were also reported to be significantly associated with SZ in two recent GWAS studies (Ripke et al., 2014).

Overall, the evidence from genetic and neuroanatomical studies, as well as from animal and in vitro models, converge to reinforce the hypothesis of the developmental origin of ASD and SZ and highlights the importance of WNT, AKT and Cadherin/Catenin signalling in their aetiology.

6.2.2 Mitochondria alterations

Mitochondria alterations represent another point of convergence between the phenotype observed in CYFIP1tg and CYFIP1ko neural cells and psychiatric disorders. Indeed, mitochondrial dysfunction has long been hypothesised to contribute to the pathology of ASD and SZ. Mitochondria-related phenotypes have been observed not only in several in vitro models of psychiatric disorders, as discussed in the previous chapter, but also in a number of animal models. For instance, two independent studies on a mouse model of Angelman syndrome showed abnormal levels of ROS and deficits in the structure of mitochondria, together with altered synaptic plasticity and performance in fear conditioning tasks (Santini et al., 2015; Su et al., 2011). Interestingly, the administration of a mitochondria-specific antioxidant in this model ameliorated both synaptic and memory phenotypes (Santini et al., 2015). Moreover, a large body of evidence suggests that altered oxidative metabolism is also present in ASD patients. Lower levels of antioxidants were found in the urine, blood and lymphoblastoid cells of ASD subjects, when compared to control samples, and a recent post-mortem study detected a decreased expression of antioxidant enzymes in ASD patients’ brains compared to age-matched controls (Damodaran & Arumugam, 2011; Frustaci et al., 2012; Gu, Chauhan, & Chauhan, 2013; James et al., 2009). In general, the prevalence of mitochondrial dysfunction in ASD has been estimated to be between 5% and 80%, depending on the measure used. Despite the variation, the level of prevalence is significantly higher than that of the general population (0.01%) (Giulivi et al., 2010; Rossignol & Frye, 2012). Moreover, the presence of mitochondrial disease in ASD patients correlates with the existence of other physiological abnormalities, such as motor delay and seizures (Rossignol & Frye, 2012).

A popular theory for the aetiology of ASD is represented by the imbalance between the excitatory and inhibitory neuron activities in the affected brain (Canitano & Pallagrosi, 2017; Marín, 2012). Interestingly, mitochondrial deficits could affect the development and function of inhibitory interneurons more seriously than excitatory neurons. A recent study showed that, in the mouse embryonic cortex, pharmacological or genetic disruption of the oxidative phosphorylation process has very severe effects on the tangential migration of MGE-derived cortical interneurons, while the radial migration of developing pyramidal neurons was not significantly affected (Lin-Hendel et al.,
Moreover, fast-spiking parvalbumin+ interneurons are particularly rich in mitochondria and have high energy requirements, which are considered critical to sustain the typical high-frequency generation of action potentials (reviewed by Kann, Papageorgiou, & Draguhn, 2014). In line with this, deletion of the cytochrome oxidase gene in mice resulted in severe electrophysiological abnormalities in these cells consistent with increased excitatory activity in the circuit. These mice also displayed social and behavioural abnormalities, similar to those observed in other SZ and ASD models (Inan et al., 2016).

Disruption of the excitation/inhibition balance has also been proposed to play an important role in the aetiology of SZ, making mitochondria dysfunction very relevant in the context of this disorder as well (Marin, 2012). In line with this, several post-mortem studies on SZ brains reported evidence of a reduction in mitochondrial number and/or functionality. The defects observed included a decreased number of mitochondria in layer 5/6 pyramidal neurons, altered expression of COX subunits in dopaminergic neurons and a significant alteration of mitochondria-related transcripts and proteins in the prefrontal cortex of SZ patients (Prabakaran et al., 2004; Rice et al., 2014; Roberts et al., 2015). Furthermore, evidence of decreased mitochondrial function has been found in the platelets of SZ subjects (Dror et al., 2002). In some cases, mitochondrial dysfunction could be secondary to other abnormalities, such as immune dysfunction or aberrant calcium homeostasis, nonetheless it represents a very common feature, present in a high number of ASD and SZ patients (Frye & Rossignol, 2011).

In conclusion, the phenotypes caused by altered levels of CYFIP1 appear to have a significant degree of overlap with several mechanisms at the origin of psychiatric disorders, supporting the view that abnormal levels of CYFIP1 expression play a major role in the clinical manifestations observed in 15q11.2 CNVs carriers.

### 6.3 Future directions

Disruption of the normal levels of CYFIP1 have been demonstrated to have a severe impact on the balance between proliferation and differentiation of cortical progenitor cells. However, alterations of the pathways regulating this process also have important implications at later stages of brain development, due to their involvement in neuronal migration and synapse formation. For instance, Cadherins are involved in synaptic adhesion and function. In particular, NCAD has been shown to be abundant at the contact sites between the filopodia, protruding from dendrites, and the axons (Hirano & Takeichi, 2012; Togashi et al., 2002). Moreover, NCAD–β-CAT complexes can bind to AMPA receptors in vivo and the amount of NCAD regulates the level of surface expression of these receptors (Nuriya & Huganir, 2006). In line with the strict connection between Cadherin-mediated
adhesion and WNT signalling, a number of WNT ligands has been demonstrated to be fundamental in the process of synaptogenesis, as well as for axonal and dendrite morphogenesis (Salinas & Zou, 2008). Together, these aspects suggest that the analysis of morphological and electrophysiological properties of CYFIP1tg and CYFIP1ko neuronal cells is also necessary and could provide additional information about the contribution of this gene to the development of psychiatric symptoms. The need for this analysis is further supported by the presence of morphological and synaptic deficits in mice with Cyfip1 mutations (Oguro-Ando et al., 2014; Pathania et al., 2014).

Another important aspect that should be examined using CYFIP1tg and CYFIP1ko hESCs is their differentiation into cortical inhibitory interneurons. Gene expression databases of mouse development show that CYFIP1 is expressed in the VZ of the MGE (Eurexpress.org), indicating that changes in the levels of this gene could potentially affect the development of this class of neurons. Moreover, as discussed above, cortical interneurons are more vulnerable to mitochondrial alterations, suggesting that the consequences of CYFIP1 loss-of-function could be even more dramatic than those observed in pyramidal cells.

Finally, to have a better understanding of CYFIP1’s role in the context of 15q11.2-related disorders, the investigation of neuronal cells derived from iPSCs carrying 15q11.2 CNVs is also necessary. Preliminary experiments on 15q11.2-deleted iPSCs showed that NIPA2 is expressed in developing neural cells and its level is reduced in neurons derived from 15q11.2del iPSCs. Therefore, NIPA2 could potentially play a role in brain development. The importance of this and the other 15q11.2 genes in the context of neurodevelopmental disorders can only be determined by analysing the phenotype of isogenic cell lines carrying mutations of one or multiple genes at a time. Experiments addressing these questions, including transcriptomic characterisation of 15q11.2-deleted neural cells and generation of lentiviral vectors for the knock-out of the genes located in this region, are currently ongoing in our group.

At the same time, we are carrying out an in-depth analysis of the RNA-seq data derived from CYFIP1tg and CYFIP1ko cells. This includes determining the degree of correlation between the transcriptome changes linked to CYFIP1 alterations and the network of genes affected by common variants in SZ patients, using the CLOZUK and SZ2 databases (Ripke et al., 2014). The same analysis will be done on the data set of DEGs found in 15q11.2-deleted neural progenitors and neurons. The comparison between the results obtained from patients’ iPSCs and CYFIP1 isogenic hESCs will provide significant information about the degree of CYFIP1 contribution towards the psychiatric phenotypes found in 15q11.2 CNVs carriers.

Lastly, we will be able to compare the transcriptomic signature of the cells used in this study with those obtained from other in vitro models of neurodevelopmental disorders developed in our
group. These include iPSCs carrying 22q11.2 deletion, a mutation strongly associated with an increased risk of SZ (Murphy, Jones, & Owen, 1999; Stefansson et al., 2008), and hESCs with knock-out of the Set Binding Protein 1 (SETBP1). Loss-of-function of this gene is strongly associated with ASD, intellectual disability and various congenital malformations (Coe et al., 2014; Roak et al., 2012). The investigation of altered mechanisms shared between the different models will help, with time, to gain a deeper understanding of the link between genetic risk factors and psychiatric phenotypes.
7 Bibliography


Boissart, C., Poulet, a, Georges, P., Darville, H., Julita, E., Delorme, R., ... Benchoua, a. (2013). Differentiation from human pluripotent stem cells of cortical neurons of the superficial layers amenable to psychiatric disease modeling and high-throughput drug screening. *Translational


Gratton, J., Wray, N. R., Keller, M. C., & Visscher, P. M. (2014). Large-scale genomics unveils the genetic


Ma, D., Salyakina, D., Jaworski, J. M., Konidari, I., Whitehead, P. L., Andersen, A. N., ... Pericak-Vance, M. A.


Marín, O., & Müller, U. (2014). Lineage origins of GABAergic versus glutamatergic neurons in the neocortex. Current Opinion in Neurobiology, 26(Figure 1), 132–141. http://doi.org/10.1016/j.conb.2014.01.015


Millar, J. K., Wilson-Annan, J. C., Anderson, S., Christie, S., Taylor, M. S., Semple, C. a, ... Porteous, D. J.


Pagnamenta, A. T., Khan, H., Walker, S., Gerrelli, D., Wing, K., Bonaglia, M. C., ... Monaco, A. P. (2011). Rare familial 16q21 microdeletions under a linkage peak implicate cadherin 8 (CDH8) in susceptibility to


Sandoe, J., & Eggan, K. (2013). Opportunities and challenges of pluripotent stem cell neurodegenerative


