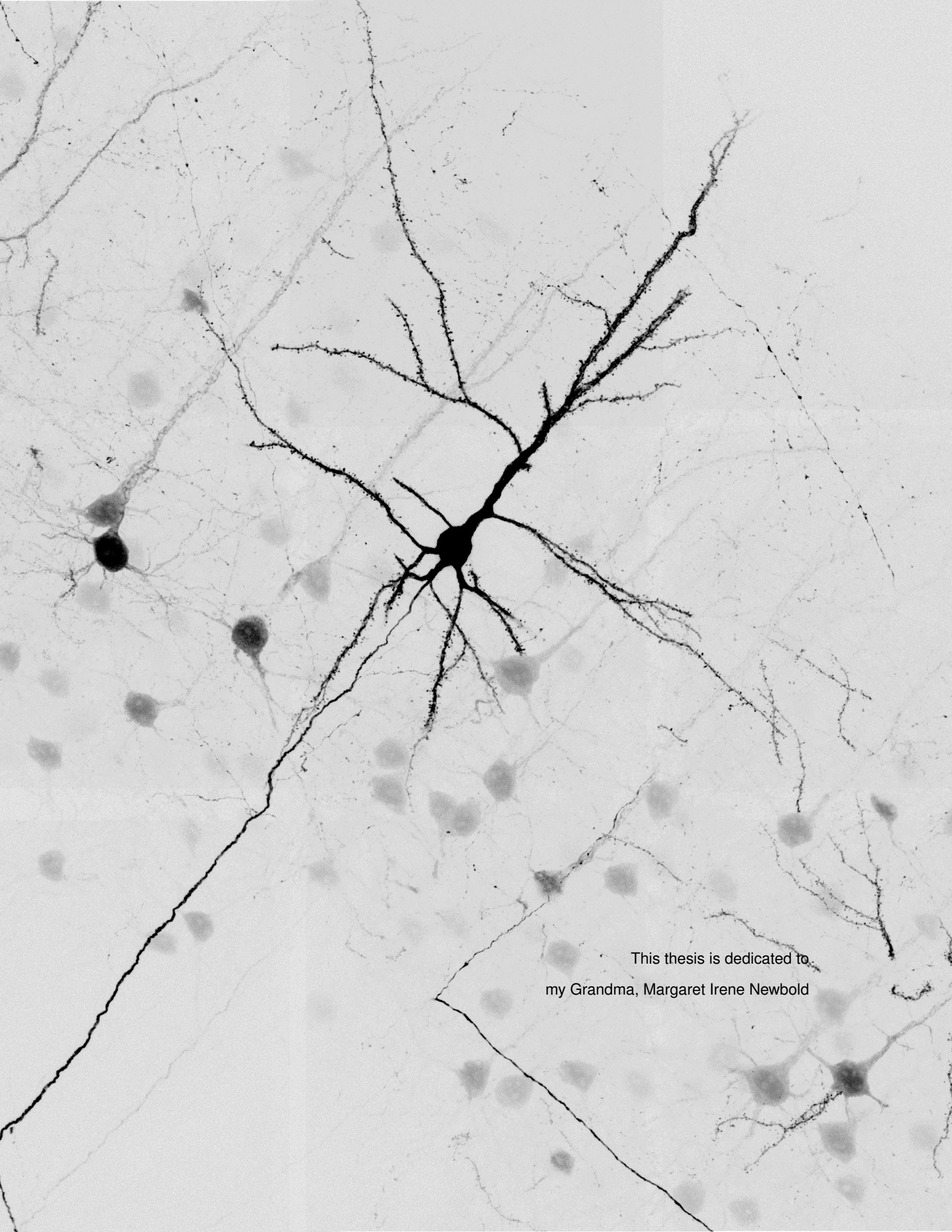
The Proteolytic Processing and Nuclear Function of Protocadherin-19

A thesis submitted in fulfilment of the requirement for the degree of Doctor of Philosophy in Integrative Neuroscience

Sylvia Adriana Newbold

Supervisor: Dr. Isabel Martinez-Garay
Co-Supervisor: Prof. Frank Sengpiel

School of Biosciences
March 2021



This thesis is dedicated to

my Grandma, Margaret Irene Newbold

Acknowledgments

First of all I would like to thank my supervisors, Dr. Isabel Martinez Garay and Prof Frank Sengpiel. Isabel, thank you for the constant support, encouragement, and great advice at every step of the way. I am truly grateful for the invaluable opportunity of a PhD in your lab and everything that I have learned from you along the way. Thank you for instilling in me the motivation to push through. Frank, thank you for your enthusiasm in the project, your positive attitude and the helpful advice throughout my PhD.

I would also like to say a special thank you to Prof Yves Alain Barde. I am grateful for all the advice and kind words I have received from you throughout the years, and thank you to Dr. Stephane Baudouin, Dr. David Petrik, Dr. Francesco Bedogni and Dr. Erik Mire for valuable input during our common lab meetings.

Thank you to the funder of this project, Wellcome Trust, which made the research possible. And a special mention to the PCDH19 Alliance which also provided funding for consumables related to this project.

I would like to thank all the people that have helped me with the technical aspects of this project, which would have not been possible otherwise. Dr. Xinsheng Nan, for valuable expertise in embryonic stem cell culture and genetic engineering, Angela Marchbank, from Cardiff University Genome Hub for advise on and execution of the RNA sequencing, Dr Sumukh Deshpande and Dr

Daniel Pass for bioinformatic support. Thanks to Kate Heesom, from the Proteomics Facility of the University of Bristol for the mass spectrometry analysis. Thanks to Marc Bishop for the FACS sorting. Thank you to Dr. Cristina Llinares-Benadero for expertise and execution of the *in utero* electroporations. I would also like to extend my special thanks to Fangli Chen, for expertise in cranial-window implantation, 2-photon imaging and being amazing to work with. Thank you to Prof. Bart De Strooper for providing ADAM10 KO and control mouse embryonic fibroblasts.

Thank you to all my extended fellow lab mates, past and present, for providing an amazing enviroment to work in, which has grown to feel like home over the years. Erin, Sarah, Laura, Shireene, Ellen, Sven, Xinsheng, Spyros, Hayley, Blanca, Sara and Rosie. You are all great! Thank you to all my fellow IMG lab members, a little ever-growing family. Natalia, Ian and James thank you for being great lab partners and a pleasure to work with. Thank you to my buddy Jess, for being a constant source of inspiration and joy, and for everything you have taught me along the way! Muchas gracias to my friend Cris, for always being there for me, helping me out and always having my back. I will miss you all.

Thank you to my amazing PhD student community of Cardiff friends. Monika, Zoe, Chiara, Ele, Aine, Sophie and Zanetta for making this journey fun, the best memories I have of my time in Cardiff are with you and you know that! Thank you to my Univerisity of Bath buddies, for all the great times and a special shout out to my gals Amy and Valerie. Thank you to my forever friends back in Italy, Cava and Laura, far only in space :) Ci vediamo presto!

Thank you to my parents, for all their love, and to my brother, Thomas, always present for the most important steps of my life!

Finally, last but definitely not least, thank you to my better half, Jay - this would not have been possible without you. Thank you for all the tech support, teaching me the ways of LaTeX, for your amazing cooking, for making me laugh and keeping me sane all the way to the end! <3

Abstract

Mutations in the X-linked gene *PCDH19* lead to epilepsy with cognitive impairment in heterozygous females and post-zygotic mosaic males. The disorder phenotype is currently explained by cellular mosaicism of PCDH19 expressing and non-expressing PCDH19 cells in the brain, which leads to defective cell-cell communication and circuits. Although the gene codes for a cell adhesion protein belonging the cadherin superfamily localized at the cell membrane, recent reports have implicated PCDH19 in the regulation of gene expression and have identified the protein in the nucleus. Despite this, the nuclear function of PCDH19 in neurons and the potential proteolytic processing of PCDH19 have not been investigated yet. This thesis focussed on the proteolytic processing of PCDH19 as a potential mechanism of membrane-to-nucleus signalling in neurons. mESC-derived neurons were used to test the involvement of different proteases in the processing of PCDH19 and it was established that PCDH19 can undergo activity-dependent proteolysis. As the cytoplasmic domain of PCDH19 was found to localise to the nucleus, the nuclear function of the generated fragment was investigated. To determine potential transcriptional targets of the PCDH19 cytoplasmic domain, a mouse embryonic stem cell line that overexpresses the cytoplasmic domain of PCDH19 from the *Rosa26* locus (PCDH19-CYTO), and an isogenic PCDH19-knockout line (PCDH19-KO) were generated. The transcriptional profile of embryonic stem cell-derived progenitors and neurons was obtained via RNA sequencing. The results of the analysis suggest a role for the cytoplasmic domain of PCDH19 in modulating expression of genes related to neuronal circuit assembly and synaptic function. Further analysis will be necessary to determine the relevance of these findings in the context of *PCDH19*-epilepsy.

Symbols

α	Alpha
β	Beta
δ	Delta
ϵ	Epsilon
γ	Gamma
kDa	Kilodalton (s)
μg	Microgram (s)
μl	Microliter (s)
μM	Micromolar

Abbreviations

Aβ	Amyloid-Beta
ADAM10	Disintegrin and Metalloproteinase domain-containing protein 10
ADAM17	Disintegrin and Metalloproteinase domain-containing protein 17
APP	Amyloid Precurson Protein
CAM	Cell-Adhesion Molecule
cKO	Conditional Knock Out
CPM	Counts Per Million
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats

CTF1	C-terminal fragment 1
CTF2	C-terminal fragment 2
DAPI	4'6-diamidino-2-phenylindol dihydrochloride
DMEM	Dulbecco's Modified Eagle Medium
FBS	Fetal Bovine Serum
GABA	γ-aminobutyric acid
GO	Gene Ontology
HEK293T	Human Embryonic Kidney Cells experiment 293 transformed with large T antigen
HET	Heterozygous
I-CLiPs	Intramembrane-cleaving protease(s)
ICD	Intracellular Domain
ID	Intellectual Disability
IP	Intermediate Progenitor
KO	Knock Out
KPNA1	Karyopherin subunit alpha-1
LC-MS	Liquid-Chromatography Mass-Spectrometry
MEF	Mouse embryonic fibroblast
mESC	Mouse embryonic stem cell
MGE	Medial Ganglionic Eminence
MT-MMP	Membrane-Type Matrix Metalloprotease
NEAA	Non Essential Amino Acids
NLS	Nuclear Localisation Sequence

NMDA	N-Methyl-D-aspartic acid
NONO	Non-POU-domain-containing octamer binding protein
NTF	N-terminal fragment
PAM	Protospacer Adjacent Motif
PCDH	Protocadherin
PCDH19	Protocaderhin 19
PSN	Presenellin-1/Presenellin-2
RGC	Radial Glial Cell
RIP	Regulated Intramembrane Proteolysis
RT	Room Temperature
scRNAseq	Single-cell RNA-sequencing
SVZ	subventricular zone
VZ	ventricular zone
WAVE	Wiskott-Aldrich syndrome protein
WIRS	WAVE interacting receptor sequence
WRC	Wave Regulatory Complex
WT	Wild Type
XCI	X-Chromosome Inactivation
ZFN	Zinc Finger Nuclease

Contents

1 Introduction **1**

1.1 The cadherin superfamily 1

1.1.1 Classical cadherins 2

1.1.2 Protocadherins 2

1.2 Protocadherin-19 7

1.2.1 *Pcdh19*/PCDH19 structure 7

1.2.2 Molecular interactions and function of PCDH19 8

1.2.3 *PCDH19*-epilepsy 11

1.2.4 Models to study *PCDH19*-epilepsy 16

1.3 Proteolytic processing 19

1.3.1 Step 1: ectodomain shedding 21

1.3.2 Step 2: intramembrane-cleaving by I-CLiPs 23

1.3.3 Proteolytic processing of cadherins 24

1.3.4 Known functions of cleavage fragments 25

1.3.5 Activity-dependent processing 27

1.4 PCDH19 and the cortex 28

1.4.1 *Pcdh19* expression in the cortex 30

1.4.2 PCDH19 in neurogenesis and migration 33

1.4.3 Axon, dendrite and synapse morphogenesis 38

1.5 Aims and hypothesis . . . 40

2 Materials and Methods 41

2.1 Animals . . . 41

2.1.1 Husbandry . . . 41

2.1.2 *Pcdh19* KO mouse line . . . 41

2.1.3 Genotyping . . . 42

2.1.4 *In utero* electroporation . . . 42

2.2 Tissue processing . . . 44

2.2.1 Perfusion . . . 44

2.2.2 Sectioning . . . 44

2.2.3 IHC and ICC . . . 44

2.2.4 Microscopy . . . 44

2.3 Cell culture . . . 45

2.3.1 Mycoplasma testing . . . 45

2.3.2 Mouse embryonic fibroblasts . . . 45

2.3.3 HEK293 cells . . . 45

2.4 Neuronal differentiation of mESCs . . . 46

2.4.1 Culture of mESCs . . . 46

2.4.2 Growing mESCs on feeder cells . . . 46

2.4.3 Formation of cellular aggregates . . . 47

2.4.4 Poly-DL-ornithine/laminin coating of plates . . . 47

2.4.5 Dissociation and neuronal differentiation . . . 49

2.5 Genetic engineering of ESCs . . . 51

2.5.1 Cloning of targeting vectors . . . 51

2.5.2 Nucleofection . . . 54

2.5.3 Antibiotic selection and colony picking . . . 57

2.5.4 Cell sorting . . . 57

2.5.5 DNA extraction and genotyping . . . 58
2.5.6 Karyotyping . . . 59
2.6 Drug treatments on cells . . . 59
2.7 Cell lysis for RNA/protein extraction . . . 60
2.8 Immunoprecipitation . . . 60
2.9 Membrane fraction enrichment . . . 61
2.10 Western blotting . . . 61
2.11 Statistical Analysis . . . 62
2.12 LC-MS . . . 64
2.13 RNA sequencing . . . 64
2.14 R packages for plotting . . . 65

3 Proteolytic processing of PCDH19 66

3.1 Introduction . . . 66
3.1.1 Aims . . . 67
3.2 Results . . . 69
3.2.1 PCDH19 nuclear localisation sequences . . . 69
3.2.2 Nuclear accumulation of PCDH19 cytoplasmic domain . . . 69
3.2.3 Predicting the size of PCDH19 cleavage fragments . . . 72
3.2.4 PCDH19 processing generates NTF, CTF1 and CTF2 . . . 74
3.2.5 ADAM10 cleaves PCDH19 in mouse embryonic fibroblasts . . . 76
3.2.6 Shedding by ADAM10 in mESC-derived neurons . . . 81
3.2.7 PS1/γ-secretase cleaves PCDH19 . . . 83
3.2.8 Activity-dependent processing of PCDH19 . . . 85
3.3 Discussion . . . 87
3.3.1 Conclusion . . . 90

4 PCDH19 interactors 91

4.1 Introduction . . . 91

4.1.1 Aims . . . 92
4.2 Results . . . 94
4.2.1 Experimental setup . . . 94
4.2.2 PCDH19 peptides . . . 96
4.2.3 PCDH19 interactors . . . 98
4.2.4 Gene ontology analysis . . . 101
4.2.5 Interesting potential novel interactors of PCDH19 . . . 104
4.2.6 KPNA1 interacts with PCDH19-CYTO . . . 108
4.3 Discussion . . . 110
4.3.1 Conclusion . . . 112

5 Generation of an *in vitro* model to study the function of PCDH19 cytoplasmic domain 113
5.1 Introduction . . . 113
5.1.1 Aims . . . 116
5.2 Results . . . 117
5.2.1 Generation of *Rosa26*-PCDH19-ICD-HA targeting vector . . . 117
5.2.2 Targeting mESCs for generation of PCDH19-CYTO mESC line . . . 119
5.2.3 Design and cloning of the gRNAs against *Pcdh19* . . . 122
5.2.4 Generation of the PCDH19-KO mESCs . . . 124
5.2.5 Karyotyping of PCDH19-CYTO and PCDH19-KO mESCs . . . 127
5.2.6 Neuronal differentiation of PCDH19-CYTO and PCDH19-KO mESCs . . . 128
5.2.7 PCDH19 expression in PCDH19-CYTO and PCDH19-KO neurons . . . 131
5.3 Discussion . . . 133
5.3.1 Conclusion . . . 135

6 Transcriptome analysis of PCDH19-CYTO and PCDH19-KO cells 136
6.1 Introduction . . . 136
6.1.1 Aims . . . 138
6.2 Results . . . 140

6.2.1 Experimental set up and quality control . . . 140
6.2.2 Principal Component Analysis and clustering . . . 143
6.2.3 *Pcdh19* expression . . . 146
6.2.4 Expression of key marker genes . . . 149
6.2.5 Differential Expression Analysis . . . 153
6.2.6 Enrichment Analysis with DEGs . . . 158
6.2.7 Gene Set Enrichment Analysis . . . 163
6.2.8 Transcriptional changes induced by neuronal activity . . . 173
6.2.9 *In vivo* overexpression of PCDH19-CYTO . . . 178
6.3 Discussion . . . 182
6.3.1 Conclusion . . . 187

7 General discussion **189**
7.1 Main findings on proteolytic processing of PCDH19 . . . 189
7.2 Main findings on nuclear function of PCDH19 . . . 190
7.3 Proposed function of PCDH19-CTF2 . . . 191
7.4 Findings in the context of *PCDH19*-epilepsy . . . 193
7.5 Limitations of mESC-derived neuron culture . . . 196
7.6 Future directions . . . 198
7.7 Open questions on PCDH19 . . . 204
7.8 Concluding remarks . . . 204

8 Appendix **244**
8.1 PCDH19 immunoprecipitation . . . 245
8.2 PCDH19 interactors . . . 247
8.2.1 Overlapping MS proteins between adult cortex and E11.5 forebrain . . . 247
8.2.2 MS proteins in adult cortex . . . 248
8.2.3 MS proteins in E11.5 forebrain . . . 265
8.3 Differentially expressed genes . . . 272

8.3.1 CYTO vs WT . . . 272
8.3.2 KO vs WT . . . 303
8.3.3 CYTO vs KO . . . 313
8.3.4 DIV12 BIC *vs* DIV12 . . . 314
8.4 Over-representation analysis with DEGs . . . 328
8.4.1 CYTO *vs* WT CA8 GO:ALL . . . 328
8.4.2 CYTO *vs* WT DIV8 GO:ALL . . . 333
8.4.3 CYTO *vs* WT DIV12 GO:ALL . . . 334
8.4.4 KO *vs* WT CA8 GO:ALL . . . 335
8.4.5 KO *vs* WT DIV8 GO:ALL . . . 337
8.4.6 KO *vs* WT DIV12 GO:ALL . . . 337
8.4.7 CYTO DIV12 BIC*vs* CYTO DIV12 GO:ALL . . . 338
8.5 Gene Set Enrichment Analysis . . . 339
8.5.1 CYTO *vs* WT CA8 GO:ALL . . . 339
8.5.2 CYTO *vs* WT DIV8 GO:ALL . . . 343
8.5.3 KO *vs* WT CA8 GO:ALL . . . 347
8.5.4 KO *vs* WT DIV8 GO:ALL . . . 348
8.5.5 CYTO *vs* WT CA8 KEGG . . . 349
8.5.6 CYTO *vs* WT DIV8 KEGG . . . 349
8.5.7 KO *vs* WT CA8 KEGG . . . 349
8.5.8 KO *vs* WT DIV8 KEGG . . . 350
8.5.9 CYTO DIV12 BIC *vs* CYTO DIV12 KEGG . . . 350
8.6 DEGs with p. adjusted cut-off < 0.01 . . . 351
8.7 DEGs with p. adjusted cut-off < 0.05 . . . 354

List of Figures

1.1 The cadherin superfamily 3
1.2 *Pcdh19*/PCDH19 structure and function 9
1.3 The cellular interference hypothesis in *PCDH19*-epilepsy 15
1.4 Mouse models of *PCDH19*-epilepsy 17
1.5 Proteolytic processing of N-Cadherin 26
1.6 *Pcdh19* expression in the glutamatergic and GABAergic clusters determined by Tasic *et al.* scRNAseq study 31
1.7 Mouse corticogenesis 37

2.1 mESC neuronal differentiation protocol 48
2.2 Flow diagram of mESC engineering via nucleofection 56

3.1 PCDH19 cytoplasmic domain localises to the nucleus in HEK293 cells 70
3.2 PCDH19 cytoplasmic domain localises to the nucleus in mESC-derived neurons . . 71
3.3 Functional domains and putative cleavage sites of PCDH19 73
3.4 Detection of PCDH19-CTF1, CTF2 and NTF in different cellular fractions 75
3.5 Stimulated shedding of N-Cadherin by ADAM10 78
3.6 Stimulated shedding of PCDH19 by ADAM10 79
3.7 PCDH19 CTF1 production in WT MEFs 80
3.8 Stimulated shedding of N-Cadherin and PCDH19 by ADAM10 in mESC-derived neurons 82

3.9 PCDH19 can be processed by PS1/γ-secretase 84
3.10 PCDH19 processing is triggered by activity of the NMDA Receptor 86

4.1 Immunoprecipitation of PCDH19 for LC-MS analysis 95
4.2 Distribution of detected peptides along the PCDH19 protein 97
4.3 Subcellular distribution of detected proteins 100
4.4 Dotplot of GO enrichment analysis for adult cortex 102
4.5 Cnetplot of GO enrichment analysis for adult cortex 103
4.6 Dotplot of GO enrichment analysis E11.5 forebrain 105
4.7 Cnetplot of GO enrichment analysis E11.5 forebrain 106
4.8 PCDH19-CYTO-HA interacts with KPNA1-MYC 109

5.1 Geneting engineering of ESCs 115
5.2 Zinc-finger nuclease targeting of the *Rosa26* locus 118
5.3 Generation of the *Rosa26*-floxedNeo-PCDH19-ICD-HA targeting vector 120
5.4 PCDH19-ICD-HA expression in targeted ESC subclones 121
5.5 CRISPR/Cas9 targeting of the *Pcdh19* locus 123
5.6 *Pcdh19* KO sequencing 125
5.7 Lack of PCDH19 in PCDH19-KO clones 126
5.8 Representative karyotype 127
5.9 PCDH19-KO and PCDH19-CYTO ESC-derived cellular aggregates gave rise to equal numbers of progenitor cells compared to WT 129
5.10 PCDH19-KO and PCDH19-CYTO ESCs can differentiate into cortical-like neurons . 130
5.11 PCDH19 expression in WT, KO and CYTO neurons 132

6.1 RNA-seq experimental design 141
6.2 RNA samples passed RNA Integrity Number (RIN) quality control 142
6.3 Principal Component Analysis (PCA) 144

6.4 Correlation matrix showing hierarchical clustering of samples based on expression of all genes . . . 145
6.5 *Pcdh19* expression in CYTO, KO and WT cells . . . 147
6.6 *Pcdh19* exon usage . . . 148
6.7 Expression of selected neuronal marker genes . . . 151
6.8 Expression of selected cortical marker genes . . . 152
6.9 Vulcano plots showing differentially expressed genes for CYTO *vs* WT, KO *vs* WT and CYTO *vs* KO at CA8, DIV8 and DIV12 . . . 155
6.10 Number of differentially expressed genes . . . 156
6.11 Top differentially expressed genes . . . 157
6.12 Overrepresentation Analysis of Gene Ontology terms in the DEG list . . . 160
6.13 DEGs at CA8 involved in axon and dendrite morphogenesis . . . 161
6.14 DEGs involved in neuronal differentiation and function . . . 162
6.15 Gene Set Enrichment Analysis with Gene Ontology terms . . . 165
6.16 GO enrichment plots . . . 166
6.17 Gene Set Enrichment Analysis with KEGG . . . 167
6.18 KEGG enrichment plots . . . 168
6.19 Expression of selected genes at CA8, DIV8 and DIV12 . . . 172
6.20 Bicuculline treated and untreated DIV12 samples . . . 175
6.21 Differentially expressed genes for bicuculline treated (DIV12 + BIC) *vs* untreated (DIV12) neurons for each cell line (CYTO, KO, WT) . . . 176
6.22 Gene Set Enrichment Analysis with KEGG for CYTO DIV12 + BIC *vs* DIV12 . . . 177
6.23 PCDH19-CYTO-HA overexpression *in vivo* . . . 180
6.24 PCDH19-CYTO overexpression effect on morphology of layer II/III neurons *in vivo* . 181

7.1 PCDH19 processing and gene regulatory function in neurons . . . 194

8.1 PCDH19 IP for LC-MS on cortex samples, uncropped . . . 245
8.2 PCDH19 IP for LC-MS on E11.5 samples, uncropped . . . 246

8.3 Vulcano plots showing DEGs for CYTO *vs* WT, KO *vs* WT and CYTO *vs* KO at CA8, DIV8 and DIV12 at p.adjusted < 0.01 351
8.4 Number of DEGs at p.adj < 0.01 352
8.5 DEGs for bicuculline treated (DIV12 + BIC) *vs* untreated (DIV12) neurons for each cell line (CYTO, KO, WT) with p.adjusted cut-off <0.01 353
8.6 Vulcano plots showing DEGs for CYTO *vs* WT, KO *vs* WT and CYTO *vs* KO at CA8, DIV8 and DIV12 at p.adj < 0.05 354
8.7 Number of DEGs at p.adj < 0.05 355
8.8 DEGs for bicuculline treated (DIV12 + BIC) *vs* untreated (DIV12) neurons for each cell line (CYTO, KO, WT) with p.adjusted cut-off <0.05 356

List of Tables

2.1 Pcdh19 genotyping primers 42
2.2 *Pcdh19* plasmids for IUE and cell-culture experiments 43
2.3 Complete medium components and concentrations for 100X stock preparation . . . 50
2.4 Preparation of fresh complete medium 51
2.5 In-Fusion cloning primers 53
2.6 CRISPR guides 53
2.7 Plasmids for ZFN targeting of *Rosa26* locus 55
2.8 Plasmids for CRISPR/Cas9 targeting of *Pcdh19* 55
2.9 Genotyping primers for *Rosa26* PCDH19-CYTO targeting 58
2.10 Sequencing primers for *Pcdh19* KO targeting 58
2.11 List of compounds used for cell treatments 59
2.12 List of Primary Antibodies 63
2.13 List of Secondary Antibodies 63

4.1 PCDH19 peptides identified by LC-MS at 1% FDR 97
4.2 Known PCDH19 interactors identified by LC-MS 99
4.3 Nuclear transporters identified by LC-MS 107

Chapter 1

Introduction

1.1 The cadherin superfamily

Adhesion molecules are essential for brain development, enabling communication and connectivity between neighbouring cells (Takeichi, 1977), (Takeichi, 2007), (Suzuki and Takeichi, 2008), (Hirano and Takeichi, 2012). The cadherin superfamily is one of the major cell-adhesion family of proteins, comprising more than 110 members further classified into subfamilies, which include classical cadherins, clustered protocadherins and non-clustered protocadherins (Hulpiau and van Roy, 2009), which will be described in further detail below. Cadherins are calcium dependent cell adhesion molecules, and members of the cadherin superfamily are so classified by the presence of extracellular cadherin repeats (EC). EC repeats contain Ca^{2+} binding sequences such as AXDXD, LDRE, and DXNDN. Binding of Ca^{2+} confers rigidity to the extracellular domain and allows homophilic and heterophilic interactions (Shapiro et al., 1995), (Chappuis-Flament et al., 2001).

1.1.1 Classical cadherins

Classical cadherins are type I transmembrane proteins, meaning they have a single-pass transmembrane domain (TM), and have extracellular oriented N-terminus and a cytosolically oriented C-terminus. They have 5 EC repeats and a conserved cytoplasmic domain that contains a β-catenin binding site and a p120-catenin binding site (**Figure 1.1a**). The adhesion mechanism of classical cadherins is mediated by their EC1 (Nose et al., 1990), via strand swap (**Figure 1.1b**). Classical cadherins are divided into type I (CDH1 (E-Cadherin), CDH2 (N-Cadherin), CDH3 (P-Cadherin), CDH4 (R-Cadherin), CDH15 (M-Cadherin)) and type II (CDH5 (VE-Cadherin), CDH6 (K-Cadherin), CDH7, CDH8, CDH9 (T1-Cadherin), CDH10 (T2-Cadherin), CDH11 (OB-Cadherin), CDH12 (N2-Cadherin), CDH18, CDH19, CDH20, CDH22, CDH24), based on sequence homology (Sotomayor et al., 2014).

Cadherins play a crucial role in morphogenesis in many different tissues. In the nervous system, they are crucial throughout development: from establishment of the neuroepithelium, to cell migration, axonal pathfinding and synapse formation (Hirano and Takeichi, 2012), (Kadowaki et al., 2007). Later, they continue to play a crucial role in synaptic transmission and plasticity (Arikkath and Reichardt, 2008). Cadherins are also a key component of adherens junctions (AJ). AJs, cell-cell adhesion complexes required for tissue morphogenesis and homeostasis, are formed by *trans* interactions of cadherins which intracellularly interact with p120-catenin and β-catenin. β-catenin, in turn, binds α-catenin (Hirano et al., 1992), forming the cadherin-catenin complex (Takeichi, 2014). Both β-catenin and p120-catenin are also key components of the Wnt signalling pathway.

1.1.2 Protocadherins

Protocadherins are expressed throughout the central nervous system both during development and in adulthood and play a pivotal role in the establishment and maintenance of neuronal circuits (Hayashi and Takeichi, 2015). Protocadherins, so called for their similarity to the classic cadherins, contain either 6 or 7 extracellular cadherin repeats (Sano et al., 1993). Protocadherins lack the

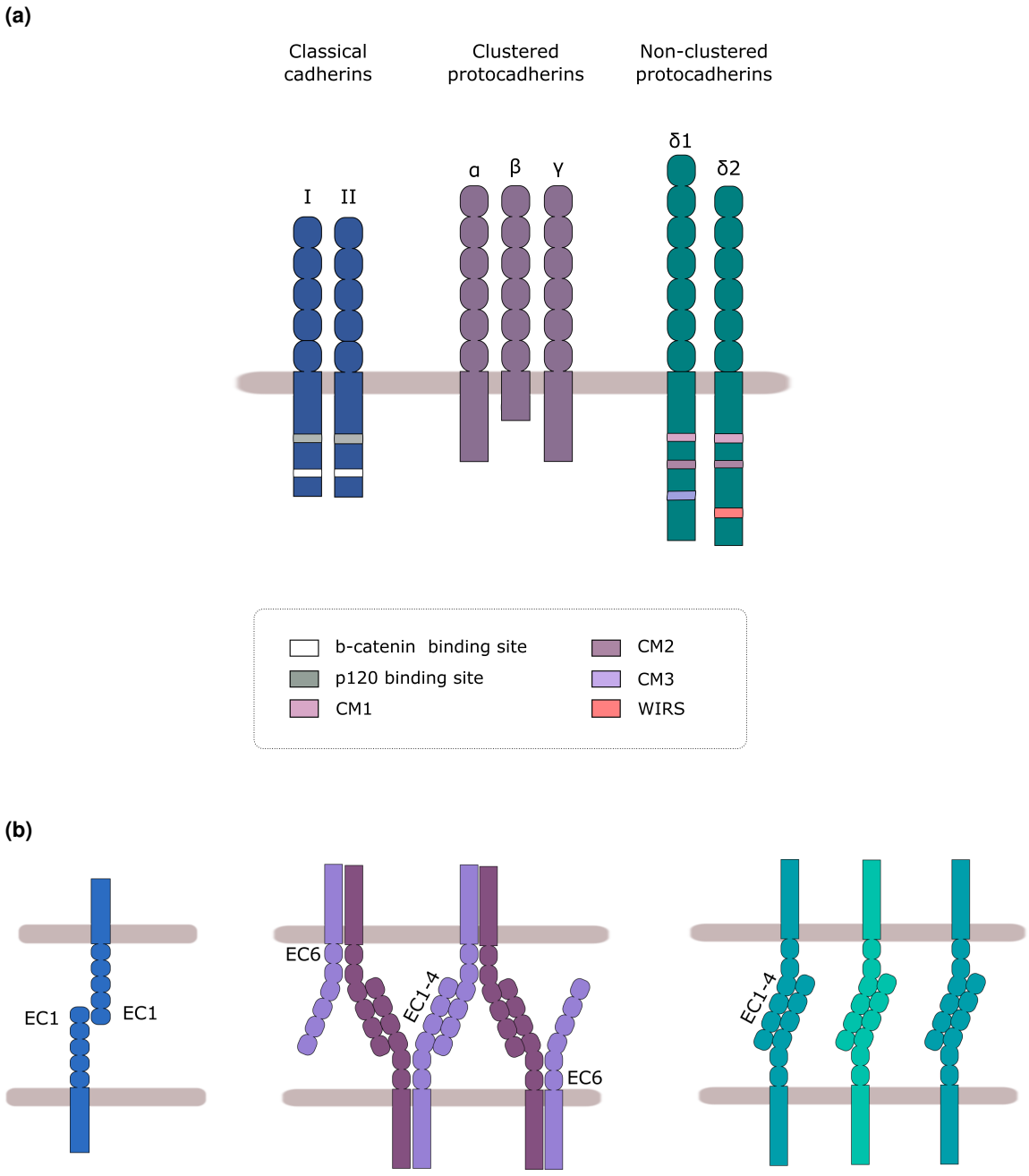


Figure 1.1: The cadherin superfamily. (a) Subfamilies of the cadherin superfamily of relevance to this thesis: the classical cadherins, the clustered protocadherins and the non-clustered δ protocadherins. Conserved cytoplasmic domains of each subfamily are highlighted. **(b)** Adhesion mechanisms: classical cadherins interact via EC1, clustered protocadherins interact in *cis* via EC6 and in *trans* via heterophilic EC1-4 interaction; non-clustered δ protocadherins interact homophilically via EC1-4.

β-catenin binding site and have different adhesion properties than classical cadherins, overall adhesiveness of protocadherins is weaker. Based on their position in the genome, protocadherins are divided into clustered and non-clustered protocadherins. In some classifications the protocadherin subfamily also includes molecules with different numbers of cadherin repeats and more than one transmembrane domain. Amongst these we find the FATs and the seven pass transmembrane flamingo/CELSR (Takeichi, 2007), (Redies et al., 2005), (Kim et al., 2011), (Hulpiau and van Roy, 2009), but the following sections will focus specifically on clustered and non-clustered protocadherins in the more narrow sense.

Clustered protocadherins

Located on mouse chromosome 18 and human chromosome 5q31 (Wu et al., 2001), the clustered protocadherins are found in three genomic clusters, the α, the β and the γ clusters, which together comprise almost 60 different isoforms (*Pcdhα1-12, αc1-c2; Pcdhβ1-22; PcdhγA1-12, γB1-7, γC3-5*) (Wu and Maniatis, 1999). The *Pcdhα* cluster has 14 large variable exons, whilst the *Pcdhβ* and *Pcdhγ* clusters have 22. Each variable exon encodes for 6 extracellular cadherin repeats, the transmembrane domain, and the variable part of the cytoplasmic domain (Peek et al., 2017). *Pcdhα* and *Pcdhγ* clusters have also 3 small constant exons that encode the constant part of the intracellular domain. Unlike *Pcdhα* and *Pcdhγ*, the *Pcdhβ* cluster does not have any constant intracellular domain (Wu et al., 2001) (**Figure 1.1a**).

The variable exons of each of the 3 clusters are stochastically expressed; therefore many different isoforms can be generated via both alternative promoter choice and alternative splicing. With the exception of *Pcdhγ C3, -C4 and -C5,* which are ubiquitously and biallelically expressed, all the other isoforms are sparsely and monoallelically expressed. For instance, a single Purkinje neuron can express up to 15 different isoforms in a completely random fashion (Esumi et al., 2005). Random expression therefore contributes to the creation of a cell surface unique barcode, that differs from neuron to neuron and allows self-recognition of neurites.

Clustered protocadherins can form heterophilic *cis* dimers and homophilic *trans* interactions (Schreiner and Weiner, 2010), (Brasch et al., 2019). In more detail, *trans* binding is mediated via antiparellel interaction of EC1-EC4, in a ”head-to-tail” conformation whilst EC6 is needed for *cis* dimerisation (Rubinstein et al., 2015), (Goodman et al., 2017) (**Figure 1.1b**). EC6 is also needed for the delivery of Pcdhα to the cell surface, which is mediated by Pcdhγ (Thu et al., 2014). When neurites of the same neuron come into contact, Pcdhs on the surface form ”zipper-like” arrays, now visualised by cryo-EM (Brasch et al., 2019) that are necessary for neuronal self-recognition and triggering of self-avoidance signalling (**Figure 1.1b**). Self-avoidance of neurite by Pcdhγ was first demonstrated in mouse retinal starburst amacrine cells (SACs) and in cerebellar Purkinje cells (Lefebvre et al., 2012).

Non-clustered protocadherins

Non-clustered protocadherins, so defined in opposition to the clustered protocadherins, are found scattered throughout the genome. They are subdivided in δ1 and δ2, which will be the focus of the next section, and solitary Pcdhs (-12, -15, -20 and -21) based on phylogenetic analysis (Redies et al., 2005).

δ protocadherins

Non-clustered δ Pcdhs can be further subdivided into δ1 and δ2 based on the number of extracellular cadherin repeats and the presence/absence of specific intracellular domains (Kim et al., 2011). δ1 Pcdhs (Pcdh-1, -7, -9, -11X/Y) have 7 EC repeats and in the cytoplasmic region they have the conserved motifs CM1, CM2 and CM3. δ2 Pcdhs (-8, -10, -17, -18, -19) have 6 EC repeats and in the cytoplasmic region conserved motifs CM1 and CM2. In addition, δ2 Pcdhs have a Wave Regulatory Complex (WRC) Interacting Receptor Sequence (WIRS), which binds the WRC (Chen et al., 2014) (**Figure 1.1a**). CM3, which is unique to δ1-Pcdh, is a binding site for protein phosphatase-1α (PP1α) (Vanhalst et al., 2005).

Non-clustered δ Pcdhs are all expressed in the central nervous system and exhibit region specific patterns of expression (Kim et al., 2007). Interestingly, some Pcdh have been shown to have complementary expression patterns, for example *Pcdh10* and *Pcdh17* have complementary expression in the corticobasal ganglia circuits (Hoshina et al., 2013). These observations led to the hypothesis that also δ Pcdhs might play a role in establishment of neuronal circuits. In fact, many δ Pcdhs have a role in axonogenesis and axonal pathfinding in different circuits (Light and Jontes, 2017). PCDH7 is important in axonal guidance of retinal ganglion cells (Leung et al., 2013) by interaction with axonal guidance cues such as Netrin1 and Sema3A. δ2 Pcdhs also play an important role in axon elongation and pathfinding via modulation of the actin cytoskeleton via the WIRS domain. For instance PCDH18 is important for motor axon elongation (Biswas et al., 2014), whilst PCDH10 and PCDH17 have been shown to be necessary for extension of thalamocortical projections (Hoshina et al., 2013), (Hayashi et al., 2014) (Uemura et al., 2007). *Pcdh10* heterozygous mutant mice also have altered circuits in the amygdala and synaptic dysfunction (Schoch et al., 2017). During zebrafish neuronal development, δ2 Pcdhs have been shown to modulate progenitor proliferation via interaction with the Wnt receptor Ryk (Biswas et al., 2020).

δ Pcdhs form homophilic *trans* interactions via EC1-EC4 antiparallel interaction (**Figure 1.1b**), but unlike clustered Pcdhs, no *cis* interactions have been shown, whilst looking at adhesion via the extracellular domain (Harrison et al., 2020). Nonetheless, it cannot be excluded that *cis* interactions between δ-Pcdh could be happening via the transmembrane or cytoplasmic region. Moreover, there is no *trans* interaction between δ1 and δ2 subfamilies, whereas within both the δ1 and δ2 subfamily, weak heterophilic *trans* interactions can occur (Harrison et al., 2020). Similarly to clustered protocadherins, δ Pcdhs can be combinatorially expressed by individual neurons, to form an ”adhesion code”, in fact a single olfactory sensory neuron can express up to 7 different δ Pcdhs (Bisogni et al., 2018).

Involvement of δ protocadherins in human disorders

Different *PCDHs* have been identified as risk genes for the development of psychiatric disorders. For instance, a genome-wide association study (GWAS) found a single nucleotide polymorphism (SNP) in *PCDH9* to be a risk factor for major depressive disorder (MDD) (Xiao et al., 2018). A SNP in *PCDH17* has also been associated with MDD and subjects presented with decreased volume and increased activity of the amygdala (Chang et al., 2018). Mutations in *PCDH10* and *PCDH8* are associated with Autism Spectrum Disorders (Morrow et al., 2008), (Butler et al., 2015). Finally, *PCDH19* mutations are the cause of early-onset epilepsy with autism features (Juberg and Hellman, 1971), (Dibbens et al., 2008), (Depienne et al., 2009). As PCDH19 is the focus of this thesis, *PCDH19*-epilepsy will be discussed in further detail later (**Subsection 1.2.3**).

1.2 Protocadherin-19

1.2.1 *Pcdh19*/PCDH19 structure

Protocadherin-19 is a δ2 non-clustered protocadherin. The structure of the protein is typical of the δ2 protocadherin subfamily – with 6 extracellular cadherin repeats (EC1-6), a single-pass transmembrane domain, a cytoplasmic domain containing the conserved domains CM1 and CM2 and the WIRS domain (**Figure 1.2a**). The extracellular cadherin repeats are responsible for mediating adhesion and binding Ca^{2+} ions. The human *PCDH19* gene is located at Xq22.1 and the mouse *Pcdh19* gene is also located on the X-chromosome. Importantly, both human and mouse *PCDH19/Pcdh19* is subject to X-chromosome inactivation (XCI). The gene is formed by 6 exons encoding a 1148 amino acid sequence. The first exon is the largest and codes for the whole extracellular portion of the protein plus the transmembrane domain. The smaller exons 2, 3, 4, 5 and 6 encode the cytoplasmic portion of the protein (**Figure 1.2a**). Exon 2 is subject to alternative splicing (Dibbens et al., 2008). As for other δ Pcdhs, PCDH19 can interact homophilically in *trans* with a PCDH19 molecule on a neighbouring cell via anti-parallel interaction. The crystallographic structure of the minimal adhesive fragment of zebrafish PCDH19 extracellular domain showed

how adhesion happens via a ”forearm handshake” which involves EC1-4 (EC1:EC4; EC2:EC3; EC3:EC2; EC4:EC1) (Cooper et al., 2016) (**Figure 1.2b**). *Pcdh19* is expressed in different tissues, both during development and in adulthood. Detailed description of *Pcdh19* expression can be found later (**Subsection 1.4.1**).

Pcdh19 **isoforms**

As mentioned, *Pcdh19* can be alternatively spliced. There are therefore several different isoforms of *Pcdh19*. In the mouse, there are 4 known isoforms: Q80TF3 is the longest isoform, containing exon 2. The other 3, E9Q5E1, Q147Z9 and A2AGW4 lack exon 2. The isoforms differ by the existence of two possible acceptor sites for intron 4 (which can add an extra residue at the beginning of exon 5). These isoforms exist in databases but their *in vivo* expression and function remains to be elucidated at the molecular level. The human *PCDH19* gene has similar isoforms: Q8TAB3-1, Q8TAB3-2 (missing exon 2) and Q8TAB3-3 (missing exon 2 and the first residue of exon 5).

1.2.2 Molecular interactions and function of PCDH19

Although no comprehensive study has been done *in vivo* in a mammalian system to determine protein interactions of PCDH19, some interesting and well-characterized interactors exist, which will be summarized below.

As previously discussed, crystallography studies have shown that PCDH19 homophilically binds to itself *in trans* (Cooper et al., 2016), confirming previous studies that showed this at the cellular level (Emond et al., 2009). PCDH19 can also form a strong complex with N-Cadherin in *cis* (**Figure 1.2b**) (Biswas et al., 2010) (Emond et al., 2011). The PCDH19/N-Cadherin complex has much stronger adhesion properties than PCDH19 alone. This *cis* complex, mediated by the EC repeats, is important during zebrafish neurulation as PCDH19/N-Cadherin act together to control cell-movement (Emond et al., 2009) (Biswas et al., 2010). Interestingly, in the PCDH19/N-

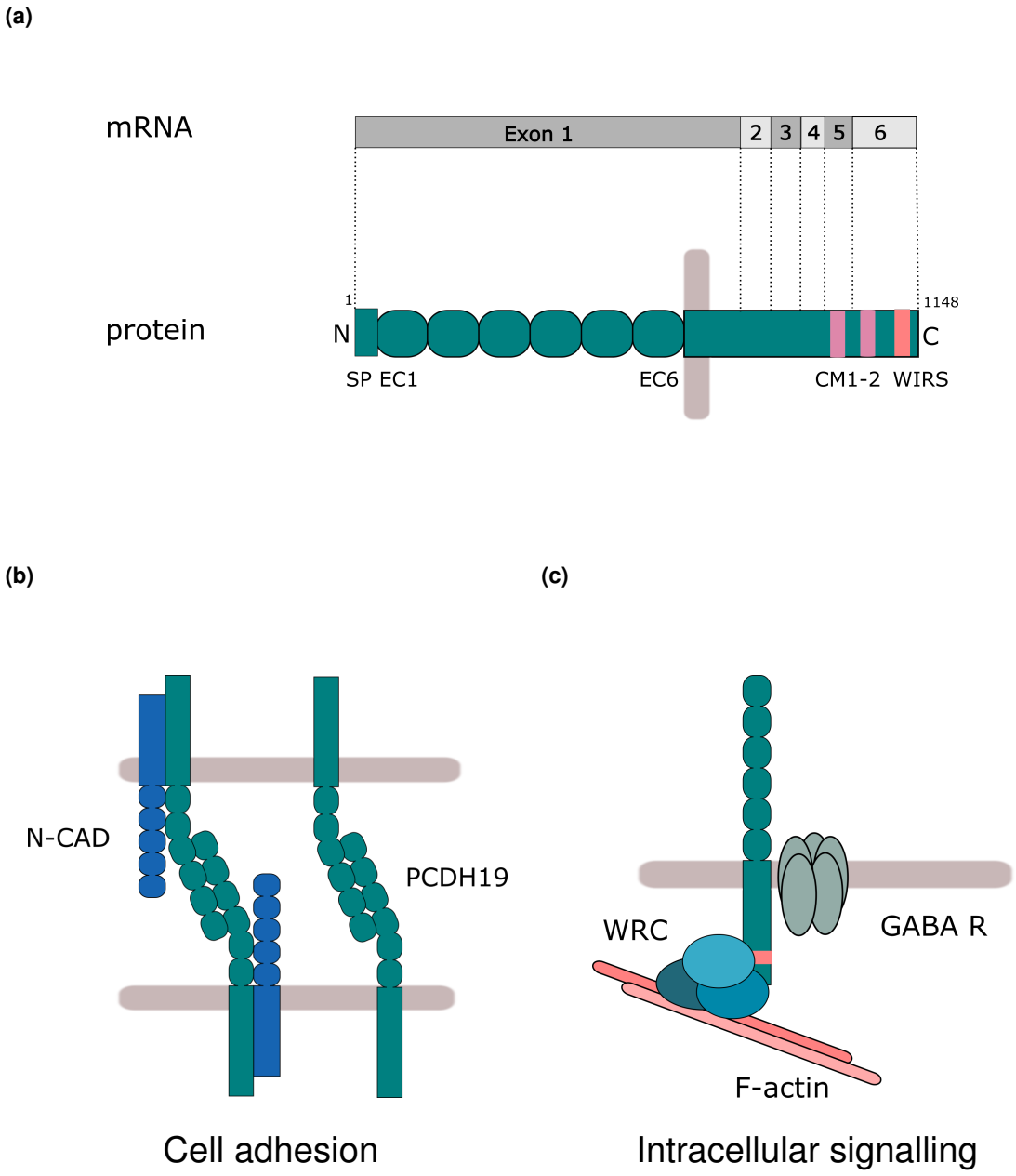


Figure 1.2: *Pcdh19*/PCDH19 structure and function. **(a)** *Pcdh19* exons and corresponding PCDH19 protein domains. SP; signal peptide; EC1-6 extracellular cadherin repeat, CM1-2, conserved motif 1, 2; WIRS, WAVE regulatory complex (WRC) Interacting Receptor Sequence. **(b)** PCDH19 mediates cell adhesion by homophilic *trans* binding of PCDH19 and via formation of PCDH19/N-Cadherin *cis* complexes. **(c)** PCDH19 interacts with the WRC to influence cytoskeletal remodelling. PCDH19 also interacts with the GABA R to modulate intracellular signalling.

Cadherin complex, N-Cadherin acts as a cofactor and *trans* adhesiveness happens via PCDH19.

PCDH19 interacts with the WAVE regulatory complex (WRC) proteins cytoplasmic FMR1-interacting protein 2 (CYFIP2) and Nck-associated protein 1 (NAP1) (Tai et al., 2010) (**Figure 1.2c**). CYFIP2 and NAP1 can also interact with other members of the δ2 Pcdh family, such as PCDH10 and PCDH17 (Nakao et al., 2008) (Hayashi et al., 2014). NAP1 and CYFIP2 are members of the WAVE regulatory complex (WRC). The WRC promotes actin nucleation via the ARP2/3 complex and therefore is an enhancer of actin cytoskeletal dynamics. This suggests that PCDH19 could be playing a role in regulating actin dynamics making PCDH19 a candidate for a role in synapse formation and axon elongation.

Supporting this theory, new interactors were identified by proximity-dependent biotinylation followed by mass spectrometry, a first attempt to define a PCDH19 interactome, in HEK293 cells. Concordantly to the previously described functions of δ Pcdhs in regulating the cytoskeleton, many of the newly identified proteins were actin and microtubule binding proteins and regulators of Rho GTPases. In particular, interaction with DOCK7, RAC-GEFs (guanine nucleotide exchange factors) and with NEDD1 (Neural precursor cell expressed developmentally down-regulated protein 1) a microtubule associated protein, were further confirmed *in vitro* (Emond et al., 2021).

PCDH19 C-terminus interacts with a conserved motif of the GABA-A α1 receptor subunit, which is shared between α1-3,5 (Bassani et al., 2018) (**Figure 1.2c**). Expression levels of PCDH19 alter surface availability of GABA receptors, possibly through endocytosis, and impact on miniature inhibitory post-synaptic currents (mIPSCs). Migration, orientation and dendritic arborisation are also affected in CA1 hippocampal rat neurons following IUE with *Pcdh19* shRNAs. Further characterisation in primary hippocampal neurons shows that PCDH19 can regulate tonic current (Serratto et al., 2020).

PCDH19 has been shown to bind to non-POU-domain-containing octamer binding protein

(NONO)/P54nrb, also known as NONO, a paraspeckle, DNA-/RNA- binding nuclear protein (Pham et al., 2017). The interaction between PCDH19 and NONO is quite surprising, considering that PCDH19 is anchored to the membrane via the transmembrane domain whilst NONO would be found in the nucleus, in paraspeckles. Paraspeckles are membraneless subnuclear structures formed by aggregation of ”scaffolding” long-non-coding RNAs (lnRNAs) and proteins (Fox et al., 2002). Paraspeckles can be identified in the nucleus of cells by staining for NONO and NEAT1 and seem to play a role in the cell- stress response and, more generally, in the retention of transcripts in the nucleus and consequently in the regulation of gene expression (Bond and Fox, 2009). The interaction between NONO and PCDH19 leads to increased transcription of ERα dependent genes (Pham et al., 2017).

PCDH19 expression is regulated by miR-484, which targets the 3’ region of *Pcdh19* mRNA, inhibiting translation. miR-484 could be the driver of 16p13.11 microduplication syndrome via the effect on *Pcdh19*. miR-484 promotes neurogenesis by inhibiting *Pcdh19* (Fujitani et al., 2017). *Pcdh19* expression is also regulated by T-box transcription factor 2 (TBR2), which directly binds two DNA sequences next to *Pcdh19* (Lv et al., 2019). Lv *et al.* showed that PCDH19 might be playing an important role during neurogenesis as suppression of *Pcdh19* via shRNA reduced neuronal output from a single radial glial cell and also caused lateral dispersion of neurons derived from the same radial glial cell possibly by changed adhesive properties of the progenitors (Lv et al., 2019).

1.2.3 *PCDH19*-epilepsy

Mutations in *PCDH19* are causative of an epilepsy syndrome known as Early Infantile Epileptic Encephalopathy 9 (EIEE9) (OMIM #300088), first described by Juberg and Hellman in 1971 as Epilepsy and Mental Retardation Limited to Females (EMRF) (Juberg and Hellman, 1971). In 2008, Dibbens *et al.* identified a mutation in the *PCDH19* gene as causative for the disease

(Dibbens et al., 2008). Since then, *PCDH19* has become the second most relevant gene in epilepsy after *SCNA1*, which is associated with Dravet Syndrome (Depienne and Leguern, 2012).

PCDH19-epilepsy presents with early onset seizure clusters, often triggered by fever (Dibbens et al., 2008). In addition, patients suffer from varying degrees of autism and intellectual disability (ID). The timing of disease onset correlates with the degree of severity, and seizure onset before 12 months is associated with the most severe intellectual disability (Kolc et al., 2019). The disorder also presents with a variety of different seizure types including tonic-clonic, tonic, clonic and focal. The most recent reports have described schizophrenia and, more generally, psychosis, as a later-onset feature of the disease (Vlaskamp et al., 2019). Abnormal cortical folding and cortical thickening by MRI (Pederick et al., 2018), differences in gyrification index of limbic structures such as the parahippocampal cortex and changes in parameters from DTI tractography have also been reported (Lenge et al., 2020).

The vast majority of reported mutations are found in the extracellular domain of the protein (86%) and of these, almost half are located in EC3 and EC4, which are the EC domains that mediate *trans* adhesiveness in PCDH19. Almost half of the reported mutations are missense mutations, the rest are frameshifts and non-sense mutations (Kolc et al., 2019). As more cases are reported, some recurrent variants are beginning to emerge, namely p.Asn340Ser, located between EC3 and EC4 and p.Tyr366Leufs, in EC4 (Kolc et al., 2019).

Interestingly, unlike most X-linked disorders, *PCDH19*-epilepsy affects heterozygous females but not hemizygous males (Dibbens et al., 2008). When first discovered, it was hypothesised that the phenotype was rescued in hemizygous males by the presence of a compensatory factor, such as *PCDH11Y*, located on the Y chromosome, as the female version *PCDH11X* has a slightly different pattern of expression in the brain (Dibbens et al., 2008). This hypothesis was soon disproved by the report of a male presenting with the disorder, carrying a mosaic mutation in *PCDH19*, and replaced by the cellular interference hypothesis (Depienne et al., 2009).

Cellular interference hypothesis

As most X-linked genes, *PCDH19* is subject to X-chromosome inactivation (XCI), an epigenetic modification that occurs in all mammalian female cells to silence transcription from one of the two X chromosomes, at random. The aim of XCI is dosage compensation in transcripts originating from the X chromosome between females (XX) and males (XY) (Panning, 2008). The result of X-inactivation in the brain of *PCDH19*-epilepsy patients is the generation of two separate populations of cells: one that expresses the normal copy of *PCDH19* (PCDH19+), the other expressing the mutant copy of *PCDH19* (PCDH19-). The resulting phenomenon is known as ”cellular mosaicism”, whereby cells expressing different versions of PCDH19 co-exist in the same tissue. The co-existence in the mosaic brain of *PCDH19*-WT and *PCDH19*-mutant cells is hypothesised to be the driving cause of *PCDH19*-epilepsy by a phenomenon known as ”cellular interference”. This theory suggests that the two populations are interfering with each-other due to altered cell-cell communication (Depienne et al., 2009), (Depienne and Leguern, 2012). In particular it is hypothesised that a ”scrambling” of connections would happen due to impaired recognition between cells that would normally be wired together, leading to defects in circuitry, which ultimately could explain both the behavioural and the epileptic phenotype of the disease. The cellular interference hypothesis would also explain the lack of symptoms in PCDH19-KO carrier males, which carry and express only one mutant copy of PCDH19 from all cells (**Figure 1.3**). As mentioned above, most of the identified mutations to date map to the extracellular domain of the protein, to key residues for homophilic binding of PCDH19 (Homan et al., 2018). Therefore, the compromised adhesive function of PCDH19 and its ability to interact in *trans* with other PCDH19 molecules seems to be a key driver of the disease. *PCDH19*-epilepsy can present with an extremely heterogeneous phenotype, that could be reflecting variability in the extension of the mosaicism in the brain, which differs in every individual and occurs at random. Interestingly, there have been increasing but rare reports of males presenting with the disorder and it has been determined that the mutations in these individuals were postzygotic somatic mutations (de Lange et al., 2017), (Kolc et al., 2020). Moreover, an individual with Klinefelter Syndrome,

characterised by the presence of an extra X-chromosome in males (47, XXY), has been shown to also display *PCDH19*-epilepsy (Romasko et al., 2018). Together, these findings strengthen the cellular interference hypothesis. Whilst this theory is currently widely accepted, the precise mechanisms driving the disease have yet to be elucidated. Cellular mosaicism is not unique to *PCDH19* as it can happen with any gene subjected to X-chromosome inactivation, as seen for example in the case of cyclin-dependent kinase-like 5 (CDKL5) deficiency disorder (CDD), or craniofrontonasal syndrome (CFNS) (Hanson and Madison, 2007), (Wieland et al., 2004).

Steroid hormone hypothesis

PCDH19-epilepsy patients have altered levels of neurosteroids. In particular, they present with allopregnanolone deficiency, measured in the blood (Tan et al., 2015) and corresponding reduced levels of aldo-keto reductase 1C (AKR1C), the enzyme responsible for production of allopregnanolone. Neurosteroids have a modulatory effect on neuronal excitability by action through the GABA-A receptor and can modulate both phasic and tonic inhibition (Reddy, 2014). In fact, allopregnanolone is a positive allosteric modulator of the GABA-A receptor, and has natural anticonvulsant properties in the brain. Interestingly, in addition to allopregnanolone other neurosteroids are also altered in *PCDH19*-epilepsy patients, including pregnanolone sulfate,17OH-progesterone and cortisol (Trivisano et al., 2017). Because of these findings, ganaxalone, a synthetic analogue of allopregnanolone, is currently being tested as a therapeutic strategy to counteract symptoms of *PCDH19*-epilepsy. Ganaxalone is now in a Phase 2 clinical trial with Marinus Pharmaceuticals (Violet Study), which should conclude in 2021. Marinus Pharmaceuticals are also trialling allopregnanolone for cyclin-dependent kinase-like 5 (CDKL5) deficiency disorder (CDD), which is also characterized by early-onset epilepsy. The mechanism by which PCDH19 heterozygosity leads to neurosteroid imbalances has yet to be determined, but it is clear that altered levels of neurosteroids can also trigger epileptic episodes.

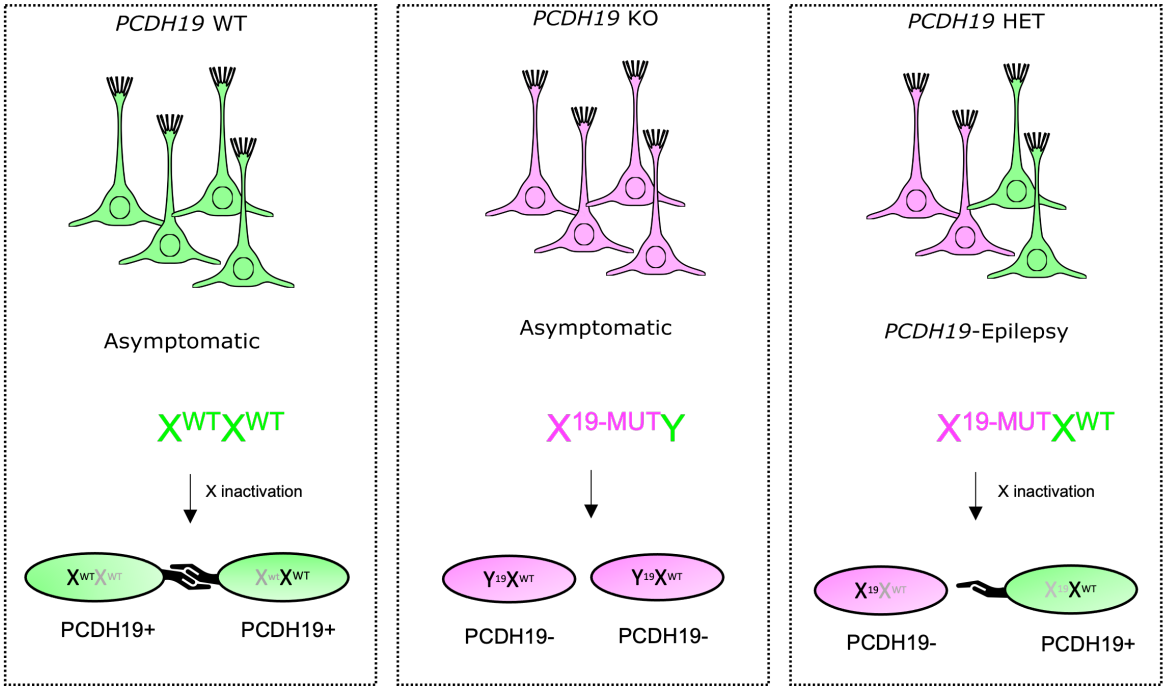


Figure 1.3: The cellular interference hypothesis in *PCDH19*-epilepsy. Disease arises in the context of *PCDH19* heterozygosity because of the cellular mosaic of *PCDH19*-WT and *PCDH19*-mutant cells co-existing in the same brain. Inspired by (Depienne et al., 2009).

1.2.4 Models to study *PCDH19*-epilepsy

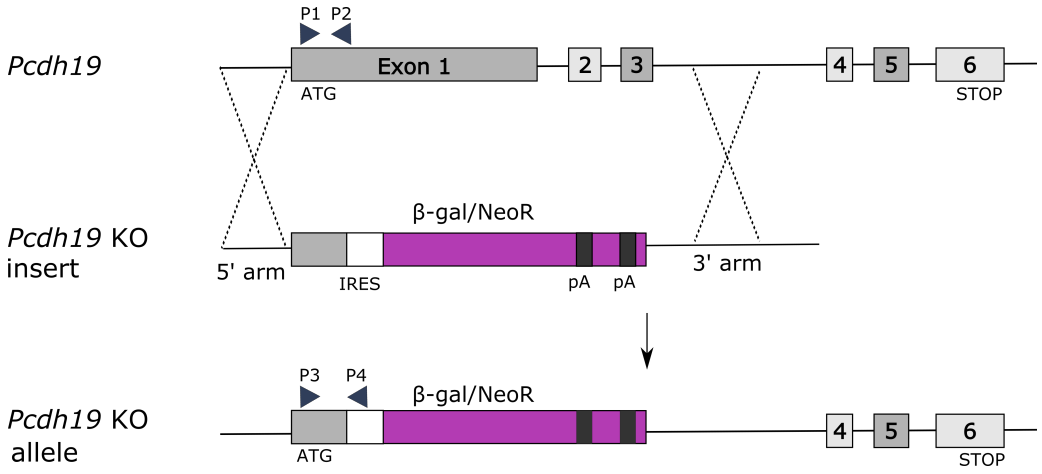
In vivo models of *PCDH19*-epilepsy

Taconic Biosciences generated a *Pcdh19*-null mouse model (*Pcdh19*$^{\beta-Geo}$) by replacement of *Pcdh19* exons 1, 2 and 3 with a β-galactosidase-neomycin fusion cassette (**Figure 1.4a**). In this model, it is possible to track expression of the mutant allele thanks to the β-galactosidase reporter, which is under control of the *Pcdh19* promoter. The first characterisation of this mouse model was done by Pederick *et al.*, which validated expression of WT and mutant *Pcdh19* but did not observe any gross abnormalities in the brains of mutant animals, as measured by cortical thickness. The only differences were seen *in vitro* when culturing KO and WT cortical explants, with KO cells showing increased migration (Pederick et al., 2016). The Taconic mouse model was investigated in further detail by Galindo-Riera *et al.*, by characterisation of the number and distribution of different cortical layer markers (CUX1, CTIP2, RORB, SATB2, TBR1) and inhibitory markers (Calbindin, Calretinin, Parvalbumin, Somatostatin) in HET females, KO males and WT animals. Confirming previous reports, cortical lamination was not altered in these animals (Galindo-Riera et al., 2020).

A second *Pcdh19* mutant mouse model was generated and validated by Hayashi *et al*. This is a KO model, with a *LacZ* insertion in the *Pcdh19* locus (Hayashi et al., 2017). Hayashi *et al.* analysed expression of *Pcdh19* in mutant animals via X-gal staining and also performed a behaviour characterisation. Wild type and mutant animals (KO males and HET females) were subjected to different tests and HET females showed hyperactivity under stress conditions and a decreased fear response (Hayashi et al., 2017).

Pederick *et al.* created a new mouse model in which the C-terminus of *Pcdh19* is tagged with an hemagglutinin (HA)-FLAG epitope sequence, making it possible to directly observe expression of the endogenous protein *in vivo*. Crossing of the *Pcdh19*$^{HA-FLAG}$ mouse and the *Pcdh19* $^{\beta-Geo}$ mouse lines enabled the tracking of mutant and WT cells in the context of the heterozygous brain.

(a)



(b)

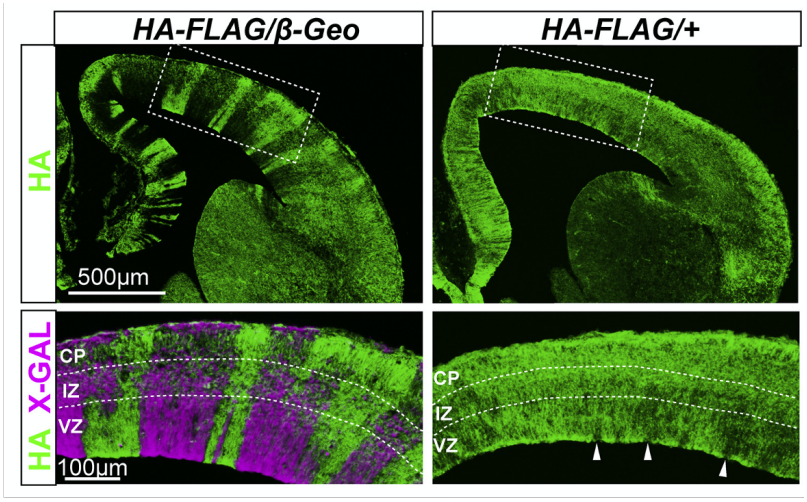


Figure 1.4: Mouse models of *PCDH19*-epilepsy. **(a)** Schematic representing the Taconic *Pcdh19* KO mouse model (*Pcdh19* $^{\beta-Geo}$) mouse model (Figure adapted from (Pederick et al., 2016). P1-P2, primers for genotyping WT allele; P3-P4; primers for genotyping KO allele, IRES, internal ribosome entry site; pA, polyA. **(b)** *Pcdh19*$^{HA-FLAG/\beta-Geo}$ mouse model shows cell-sorting behaviour during cortical development (Pederick et al., 2018).

In this *Pcdh19*$^{HA-FLAG/\beta-Geo}$ mouse model, Pederick *et al.* observed a striking cell-sorting behaviour of progenitor cells in the cortex, which started around E10.5, only 24 hours after onset of *Pcdh19* expression (Pederick et al., 2018) (**Figure 1.4b**). The authors confirmed that the column-like phenotype was not due to random X-inactivation, but to co-existance of PCDH19 expressing and non-expressing cells, in line with the cellular interference hypothesis (Pederick et al., 2018).

An acute model of *PCDH19*-epilepsy has also been generated via *in utero* electroporation in the rat (Bassani et al., 2018), (Cwetsch et al., 2020). In this model shRNAs were used to downregulate *Pcdh19* expression during development, targeting the hippocampus. Because *in utero* electroporation will only target a subset of cells, it effectively recreates a cellular mosaic in the targeted area. These models identified defects in migration and behavioural deficits linked to autism (Cwetsch et al., 2020), and morphological defects of pyramidal neurons in the hippocampus (Bassani et al., 2018). Recently, a new tool was developed to investigate cellular mosaicism *in vivo* based on tunable, amplified Cre recombinase activity, which could be applied to recreate a model of *PCDH19*-epilepsy (Trovato et al., 2020).

In vitro models of *PCDH19*-epilepsy

An *in vitro* attempt to recreate the cellular mosaic present in the PCDH19 heterozygous brain has been carried out by mixing neural stem cells derived from PCDH19 KO and WT E14.5 cortices in a 1:1 ratio to generate neurospheres (Homan et al., 2018). Homan also reprogrammed skin fibroblasts from *PCDH19*-epilepsy patients and controls into stable induced pluripotent stem cells (iPSCs) lines. When subjected to neuronal differentiation *in vitro*, both human and mouse cells carrying *PCDH19* mutations showed increased neurogenesis. Moreover, the human neuronal progenitors displayed changes in apical/basal polarity of the neuronal rosettes (Homan et al., 2018).

Recently, another *in vitro* model was generated by culturing of primary hippocampal neurons

derived from PCDH19 KO and PCDH19-Tag (*Pcdh19*$^{HA-FLAG}$) mouse. Interestingly, mosaic cultures show disrupted neuronal morphology and network activity, with reduced synaptic contact between PCDH19-KO and PCDH19-Tag neurons (Mincheva-Tasheva et al., 2021).

1.3 Proteolytic processing

A recent report has shown that PCDH19 can be found in the nucleus, where, together with NONO, it can modulate expression of ERα-dependent genes (Pham et al., 2017). As PCDH19 is a transmembrane protein, we hypothesised that PCDH19 could be proteolytically cleaved at the membrane with the resulting release of an intracellular fragment. This section of the introduction will present the main steps of proteolytic processing with relevant examples and a focus on the proteases ADAM10 and γ-secretase. A summary of what is known on cadherin processing and function of cadherin cleavage fragments will be presented. Finally, because of it's relevance to cortical development and function, a brief explanation of activity-dependent processing will be provided.

All cells need to interact with their environment and be able to respond to signals and stimuli. In many cases, this response will involve changes in gene expression, brought about by signalling pathways originating at the cell's membrane. Many membrane proteins function as receptors and relay signals downstream to inform and direct appropriate changes in gene expression. One example of a process that relies heavily on signal transduction between cell surface and nucleus is the wiring of neuronal circuits and refinement of synaptic contacts during cortical development. There are different ways in which signals can be transduced and many include phosphorylation cascades that activate certain proteins and allow them to translocate to the nucleus. Another molecular mechanism that allows signalling to occur efficiently is proteolytic processing or cleavage of membrane bound receptors and guidance molecules. The intracellular domains of those proteins are then released into the cytoplasm and can start signalling pathways or even enter the nucleus to interact with other proteins to regulate gene expression. Proteolytic

processing is the underlying mechanism of many core signalling pathways critical for appropriate development, ranging from cell-fate specification to axonal pathfinding (Coleman et al., 2010) and establishing of brain circuits, as it provides a simple, direct and irreversible mechanism of cell-surface to nucleus communication (Bai and Pfaff, 2011).

In the context of neuronal development, one of the best-known examples of proteolytic processing which results in membrane-to-nucleus signal transduction is the Notch signalling pathway (Bray, 2006). Upon binding to its ligand Delta on the surface of a neighbouring cell, the *trans*-activated receptor Notch undergoes a conformational change that exposes a previously buried cleavage site (Gordon et al., 2007). A proteolytic cascade leads to the release of the Notch intracellular fragment (NICD) (Pan and Gerald M, 1997), (Schroeter et al., 1998), (De Strooper et al., 1999), (Qi et al., 1999). The cleaved fragment then translocates to the nucleus where it interacts with the DNA-binding protein CSL and induces transcription of target genes (Jarriault S et al., 1995). Activation of the Notch pathway is key in maintaining the balance between neural progenitor proliferation and differentiation during development (Kawaguchi et al., 2008).

Because of its role in many physiological and pathological processes, Notch is the most studied example, but many other membrane proteins and cell-adhesion molecules such as Netrin receptors (Bai et al., 2011), Ephrins (Tomita et al., 2006), Neurexins (Bot et al., 2011), Neuroligin receptors (Sardi et al., 2006) and Cadherins (Marambaud et al., 2003) have been shown to be regulated by proteolytic processing in a similar way and are capable of transducing information across the membrane.

For certain types of proteins, such as type I membrane proteins (single TM, cytosolically oriented C-terminus), including Notch, proteolytic processing often occurs in a two-step process, referred to as Regulated Intramembrane Proteolysis (RIP). Upon an initial trigger, a first cleavage event, known as ”ectodomain shedding”, results in the release of the extracellular portion of the protein. Shedding is then followed by a second cleavage event, which releases the cytoplasmic

fragment, otherwise still attached to the membrane, into the cytosol. These two steps and the key-players involved in these processes are explained in further detail in the following sections.

1.3.1 Step 1: ectodomain shedding

Ectodomain shedding is the first event in the RIP proteolytic processing cascade (Lichtenthaler et al., 2018). Shedding can be triggered by direct ligand binding to the substrate (as with Delta binding of Notch Receptor (Gordon et al., 2007) or BDNF/NGF binding of p75NTR (Frade, 2005)) or by increased concentration of Ca^{2+} in the cell, as happens with neuronal activity (Suzuki et al., 2012). Shedding can have different functional consequences: activation or silencing of signalling pathways, triggering of further downstream proteolysis, altering strength of cell-cell or cell-matrix interactions or simply be a mechanism of protein turnover. Shedding can occur both at the plasma membrane or at any point of the endocytic pathway (Buchanan et al., 2010). Shedding can also be regulated by other factors that influence accessibility or conformational state of the substrate protein such as oligomerisation (Bonn et al., 2007) or N-glycosylation (Gowrishankar et al., 2004), (May et al., 2003). There are several families of proteases that mediate this process including the ADAMs, the BACEs and the MT-MMPs, described below.

ADAMs

The A-disintegrin and Metalloprotease (ADAM) family of proteases is one of the key families of proteases responsible for ectodomain shedding. The ADAMs are membrane-bound, zinc-dependent metalloproteases. The signature catalytic domain of the ADAM family is HEXGHXXGXXHD, in which the histidine residues (H) bind Zn^{2+} and the glutamic acid helps with the catalytic event (Bode et al., 1993), (Hsia et al., 2019).

Many ADAMs are expressed in the mammalian nervous system (ADAM8, -9, -10, -12, -15, -17 and -19) but amongst them ADAM10 and ADAM17 are the best characterised. ADAM10 has

gained much attention in the fields of neurodevelopment and neurodegeneration because of the relevance of its substrates, Notch and the amyloid precursor protein (APP), in these two areas (Hartmann et al., 2002), (van Tetering et al., 2009), (Lammich et al., 1999). Other substrates of this protease include members of the cadherin superfamily, ephrins, neurexins and other CAMs, making this protein crucial for normal development (Janes et al., 2005), (Borcel et al., 2016). ADAM10 KO mice die at E9 due to a defective Notch/Delta signalling pathway (Hartmann et al., 2002). Interestingly, ADAM10 cKO animals induced with Nestin-Cre have abnormal cortical development with premature neuronal differentiation and defective neuronal migration (Jorissen et al., 2010). Postnatal disruption of ADAM10 via CaMKII-Cre conditional ADAM10 KO leads instead to defects in synaptic function and seizures (Prox et al., 2013).

BACE1 and BACE2

The β-site APP cleaving enzymes (BACE1 and BACE2) are sheddases most studied for their role in Alzheimer's Disease (AD). BACE1 is an aspartic acid protease known to generate the pathological amyloid-beta (Aβ) fragments from the amyloid precursor protein (APP) via the amyloid β-secretase pathway (BACE1/β-secretase cleavage + γ-secretase cleavage) (Vassar et al., 1999). Other substrates are being continuously identified including protocadherins, ephrins and semaphorins (Hemming et al., 2009).

Membrane-Type Matrix Metalloproteases

Part of the larger Matrix Metalloprotease family, the Membrane-type Matrix Metalloproteases (MT-MMPs) are a 6 member subfamily of transmembrane (MT1-MMP, MT2-MMP, MT3-MMP, MT5-MMP) and GPI-anchored proteases (MT4-MMP, MT6-MMP) (Itoh, 2015). MT-MMPs primary substrates include components of the extracellular matrix (laminin, fibronecin, collagen, gelatin) but interestingly, MT5-MMP, can also cleave N-Cadherin. N-Cadherin cleavage happens in the stem-cell niche of the subventricular zone and regulates proliferative behaviour of adult neuronal stem cells (Porlan et al., 2014). MT-MMPs can undergo shedding themselves and be released from the

cell surface into the ECM as a means of regulating their activity (Itoh, 2015).

1.3.2 Step 2: intramembrane-cleaving by I-CLiPs

The second step of RIP of membrane proteins is carried out by a type of proteases often referred to as intramembrane-cleaving proteases (I-CLiPs) (Beel and Sanders, 2008). During this process, a membrane protein is cleaved within its transmembrane domain and a soluble C-terminal fragment, often with signalling function, is released within the cytoplasm. Ectodomain shedding is in most cases a prerequisite for I-CLiP function, as these will often recognize their substrate by the cleaved extracellular stub left by the sheddase, with length of the stub being a key to substrate recognition (Fleck et al., 2016). Amongst the I-CliPs there are the metalloproteases signal peptide peptidases (S2P), the serine proteases of the rhomboid family and the well-known γ-secretase complex.

The γ-secretase complex

The γ-secretase complex is responsible for proteolysis of both Notch and APP (De Strooper et al., 1999), (De Strooper et al., 1998) and also cleaves many other type I single TM-proteins with cytosolically oriented C-terminus, including N-Cadherin (Uemura et al., 2006b). In some cases, the γ-secretase complex can cleave substrates without previous ectodomain shedding, as in the case of p75-CTF generation, which is independent from p75-NTF production (Frade, 2005).

The γ-secretase is a complex formed of 4 essential protein components: presenilin (PSN), nicastrin (Nct), anterior pharynx-defective 1 (APH-1) and presenilin enhancer 2 (Pen-2), with a stoichiometry of 1:1:1:1 (Sato et al., 2007), (Lu et al., 2014). Presenilin constitutes the core catalytic subunit and exists in two different isoforms, PSN1 (the major isoform) and PSN2. Presenilin is a 9-pass TM aspartyl protease containing the signature catalytic motifs YD and LGxGD (on TM 6 and 7 respecitvely), each containing one of the two active site aspartates (D) (Wolfe et al., 1999). Nicastrin is necessary for substrate recognition and binding (Bolduc

et al., 2016), APH-1 is necessary for assembly and trafficking of the complex (Niimura et al., 2005) and Pen-2 is required for endoproteolysis of Presenilin, which yields a C-terminal and an N-terminal fragment (Brunkan et al., 2005) - thus each component is necessary for the function of the complex (Beel and Sanders, 2008). Mutations in PSN1 are the most common cause of familial Alzheimer's disease (FAD) (Kelleher and Shen, 2017). As mentioned, PSN1 can in fact cleave APP to generate the pathogenic amyloid-beta (Aβ) fragments after BACE1 cleavage but the mechanism of how loss-of-function mutations lead to FAD is still a matter of debate.

1.3.3 Proteolytic processing of cadherins

Several examples of proteolytic processing have been identified for the cadherin superfamily: the classical cadherins N-Cadherin (Reiss et al., 2005), (Uemura et al., 2006b) and E-Cadherin (Marambaud et al., 2002), several clustered protocadherins belonging to the α and γ subfamilies (PcdhγA3 (Haas et al., 2005), PcdhγC3, (Haas et al., 2005) (Reiss et al., 2006), Pcdhα4 and PcdhγA1 (Bonn et al., 2007), and the solitary PCDH12 (Bouillot et al., 2011) have been reported so far. Processing occurs mostly by RIP, with the extracellular portion of the protein being released by a sheddase, often ADAM10, and subsequently the cytoplasmic domain being released in the cytosol by the γ-secretase complex.

As cadherins engage in cell-adhesion, their proteolytic processing can first of all impact the strength of contact between cells. For example, processing of E-Cadherin can lead to disassembly of adherens junctions via release of the E-Cadherin-CTF2/β-catenin complex to the cytoplasm (Marambaud et al., 2002). Similarly, processing of N-cadherin at neuronal synapses can have an effect on size and morphology of dendritic spines and their plasticity (Malinverno et al., 2010).

Not much is known yet about the mechanisms that regulate processing of cadherins, although it has been shown that endocytosis of α Pcdhs via interaction with the endosomal sorting complex

required for transport (ESCRT-0) is a prerequisite to processing during neuronal differentiation (Buchanan et al., 2010). It has also been determined in primary cortical neurons that combinatorial expression of α- and γ-Pcdhs can alter their susceptibility to Presenilin-dependent processing (Bonn et al., 2007).

1.3.4 Known functions of cleavage fragments

Initially, it was believed that processing of cadherins served merely as a mechanism to regulate the adhesion properties of cadherin complexes or as a mechanism to remove and degrade these proteins, but it has been established that both the extracellular, N-terminal fragment (NTF) and the intracellular, C-terminal fragment (CTF) of processed cadherins can retain biological activity (McCusker and Alfandari, 2009).

The extracellular fragment of N-Cadherin (N-Cad-NTF), for example, can stimulate neurite outgrowth of cerebellar neurons in culture via action on the FGF receptor (Utton et al., 2001). These neurite promoting function of N-Cadherin NTF has also been seen during retinal development in the chick (Paradies and Grunwald, 1993). Similarly, Cadherin-11 NTF is necessary for cranial neural crest migration in *Xenopous* by antagonistic binding to full-lenght Cadherin-11 (McCusker et al., 2009).

In parallel, the released cytoplasmic fragment of N-cadherin (N-Cad-CTF2) sequesters β-catenin, preventing its degradation and subsequently enhancing transcription of the downstream targets *cyclinD1*, *c-myc* and *c-jun* (Reiss et al., 2005), (Shoval et al., 2007). N-Cadherin CTF2 can also bind the CREB binding protein (CBP). CBP is a co-activator of the transcription factor CREB (cAMP response element-binding protein), which controls expression of key genes for synaptic plasticity and neuronal function (Marambaud et al., 2003) (**Figure 1.5**). Processing of N-Cadherin results in the release of cadherin binding proteins and the activation of β-catenin

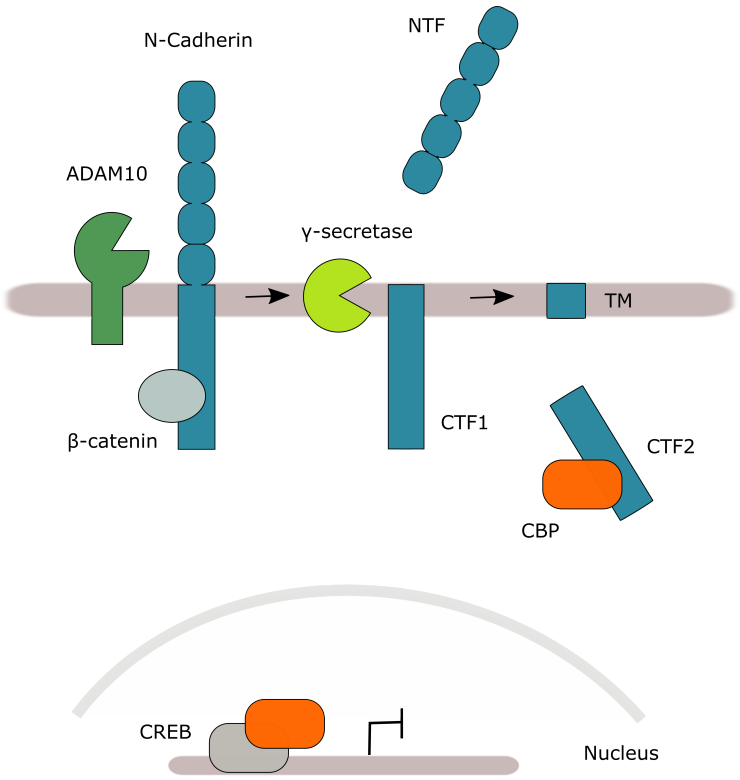


Figure 1.5: Proteolytic processing of N-Cadherin. N-Cadherin is cleaved in two steps by ADAM10 and γ-secretase. The first step releases the extracellular domain NTF, whilst the second cleavage results in the production of CTF2. CTF2 can associate with CBP and modulate CREB dependent gene expression.

target genes. Similarly, E-cadherin CTF also plays a role in the cell nucleus by modulating the p120-Kaiso-mediated signalling pathway (Ferber et al., 2008).

PcdhγC3 and Pcdhα3 have been shown to be processed and their CTFs have been found in the nucleus (Haas et al., 2005). In addition, CTFs of γ-Pcdhs promote locus expression of γ-*Pcdhs*, so that processing of the protein results in increased transcription of the gene itself in an auto-regulatory manner (Hambsch et al., 2005).

1.3.5 Activity-dependent processing

As mentioned earlier, proteolytic processing in neurons often occurs in response to neuronal activity (Malinverno et al., 2010) (Suzuki et al., 2012), (Alberi et al., 2011). In fact, the sheddases ADAM10, ADAM17 and MT5-MMP can localize to synapses in both pre-synaptic and post-synaptic compartments (Restituito et al., 2011). γ-secretase has also been identified both in pre- and post-synaptic compartments (Schedin-Weiss et al., 2016). Hence the whole machinery, substrate, sheddases and I-CliPs can be found in synapses.

Activity dependent processing starts during synaptogenesis and helps shape forming synapses. This is seen for example with cleavage of ICAM5, which serves as a maturation-inhibiting signal until it is cleaved by MMP-9 (Tian et al., 2007). Neuroligin-1 processing, on the other hand, inhibits further maturation once a synapse is formed, hence fine-tuning synaptic maturation (Suzuki et al., 2012). Activity-dependent processing continues in adulthood and plays a role in modulating synaptic transmission, affecting both long term potentiation (LTP) and long term depression (LTD). Processing of adhesion molecules at the synapse can alter the strength of synaptic contacts, remodel spines, and it generally reduces synaptic transmission (Restituito et al., 2011). In fact, PS1^{L286V} mutant mice show defects in spine morphology and synaptic plasticity, showing increased spine density and increased LTP responses to NMDA-R-mediated activity (Auffret et al., 2009). Similarly, as mentioned, ADAM10 postnatal conditional KO leads to impaired LTP (Prox

et al., 2013). In fact, ADAM10 is endocytosed in response to LTP and reinsterted in the membrane following LTD, and is important for LTD maintenance (Marcello et al., 2013), (Musardo et al., 2014).

Interestingly, a cleavage-resistant N-Cadherin mouse model (N-Cadherin^{R714G,I715D}) was recently generated by Asada-Utsigi *et al* (Asada-Utsugi et al., 2021). The missense mutation confers resistance to ADAM10 cleavage. These mutant mice show higher synaptic density and complexity of dendritic tufts in the CA3 region of the hippocampus and better performance on the radial maze test, confirming *in vivo* how N-Cadherin cleavage is important in excitatory synapse function and plasticity (Asada-Utsugi et al., 2021).

Processing can immediately alter synaptic strength through structural changes, but there is some evidence that ICDs can also undergo regulated nuclear import from the synapse by long-distance retrograde transport, in order to modulate transcription (Sachse et al., 2019), (Bao et al., 2004) (Hallaq et al., 2015). In fact, nuclear transport proteins, known as importins, have been found in synapses (Jeffrey et al., 2009). The NR1 subunit of the NMDA receptor contains a nuclear localisation sequence (NLS) which is recognised by importin α (Jeffrey et al., 2009). The NR1-importin α interaction holds importin α at the post-synaptic density until, in response to a neuronal activity, PKC phosphorylates NR1 within the NLS and releases importin α, making it free to bind and transport other cargo (Jeffrey et al., 2009). Synapse-to-nucleus signalling is important for triggering long-lasting transcription-dependent forms of synaptic plasticity (Thompson et al., 2004).

1.4 PCDH19 and the cortex

The cerebral cortex, responsible for both sensory information processing and higher cognitive function, is arguably the most interesting and complex structure of the mammalian brain. Meticulous mapping of the cytoarchitecture of the human cortex, done more than 100 years ago by Broadmann, led to classification of more than 50 areas, still relevant today, based on anatomical

and functional organisation (Brodmann, 1909).

Radially the cerebral cortex is organised into 6 distinct cell layers (*laminae*): layer I (the molecular layer); layer II (the external granular layer); layer III (the external pyramidal layer); layer IV (the internal granular layer); layer V (the internal pyramidal layer, subdivided into outer layer, Va, and inner layer, Vb) and layer VI (the multiform layer). Each layer has a unique cellular composition and connections. Laminar cytoarchitecture is not homogeneous across the cortex but reflects functionality of different areas (Scala et al., 2019). For example, in the somatosensory cortex (which classifies as heterotypical granular), layer IV, the major input layer, is thick at the expense of layer V, the major projection layer. The situation is reversed in the motor cortex (heterotypical agranular), with a thick projection layer and thin input layer (Brodmann, 1909), (Harris and Shepherd, 2015).

There is extraordinary cell diversity in the cortex, with hundreds of different neuronal and glial cell types (Tasic et al., 2018). In the broadest classification there are two major neuronal populations in the cortex, inhibitory, γ-aminobutyric acid (GABA) expressing interneurons, representing about 20% of cortical neurons, and excitatory, glutamatergic neurons, accounting for the remaining 80%. GABAergic neurons are locally projecting and provide the inhibitory input to the excitatory neurons. GABAergic neurons have remarkably diverse morphologies such that not only are they classified based on expression of certain molecular markers (Parvalbumin (PV), Somatostatin (SST), Ionotropic Serotonin Receptor (5HT3a)), but also on the shape created by their dendritic arbour and their electrical activity. GABAergic interneurons thus include basket cells, chandelier cells and Martinotti cells, amongst others (Huang and Paul, 2019).

The vast majority of glutamatergic neurons are classified as pyramidal neurons. Pyramidal neurons have a long apical dendrite and wide basal dendrites, and can be further classified in tufted and non-tufted based on the shape of the apical dendrite. Both soma size and dendritic size varies across brain regions (Benavides-Piccione et al., 2006) and across layers. Pyramidal

neurons have diverse connectivity. For example layer II/III pyramidal neurons mainly project within the cortex, whilst layer V mainly project subcortically (Harris and Shepherd, 2015). Molecular markers such as CTIP2, OTX1, TBR1, SATB2 or CUX2 distinguish pyramidal neurons based on their laminar identity and projection patterns (Arlotta et al., 2005), (Nieto et al., 2004) (reviewed by (Molyneaux et al., 2007)). Another, less prominent class of glutamatergic neurons, is the spiny stellate neurons, that populate layer IV in sensory areas (Saint Marie and Peters, 1985) and are so called because of the many dendritic spines.

1.4.1 *Pcdh19* expression in the cortex

Pcdh19 is expressed in several tissues including heart, kidney and lung, but expression is most prominent in the brain (Wolverton and Lalande, 2001). In the the mammalian brain, *Pcdh19* expression starts very early in development and is detectable in the mouse forebrain as early as embryonic day 9 (E9) (Gaitan and Bouchard, 2006). At the onset of cortical neurogenesis, *Pcdh19* is strongly expressed by rapidly dividing neural progenitor cells, radial glial cells that are located in the ventricular zone. Interestingly, expression of PCDH19 is complementary to the neurogenic gradient, declining after E13.5 (Dr. Jessica Griffiths, unpublished). Postnatally, *Pcdh19* can be detected in several brain regions. Between P0 and P7 it is strongly expressed in the hippocampus, in the CA1, CA2 and CA3 fields, but surprisingly it only appears in the dentate gyrus during adulthood (Schaarschuch and Hertel, 2018) (Bassani et al., 2018). Strong expression is also seen in several nuclei of the amygdala, in the anterior hypothalamus and the anterior thalamus (Hertel et al., 2008). In the adult cortex, *Pcdh19* expression is restricted to cortical layers II/III and Va (Pederick et al., 2016).

Combined *in situ* hybridisation with immunohistochemistry has started to tease apart specific subpopulations of neurons in the cortex that express *Pcdh19*. RORB, SATB2, CTIP2, TBR1 and PVALB were used as markers and show partial co-localization with *Pcdh19*, with strongest

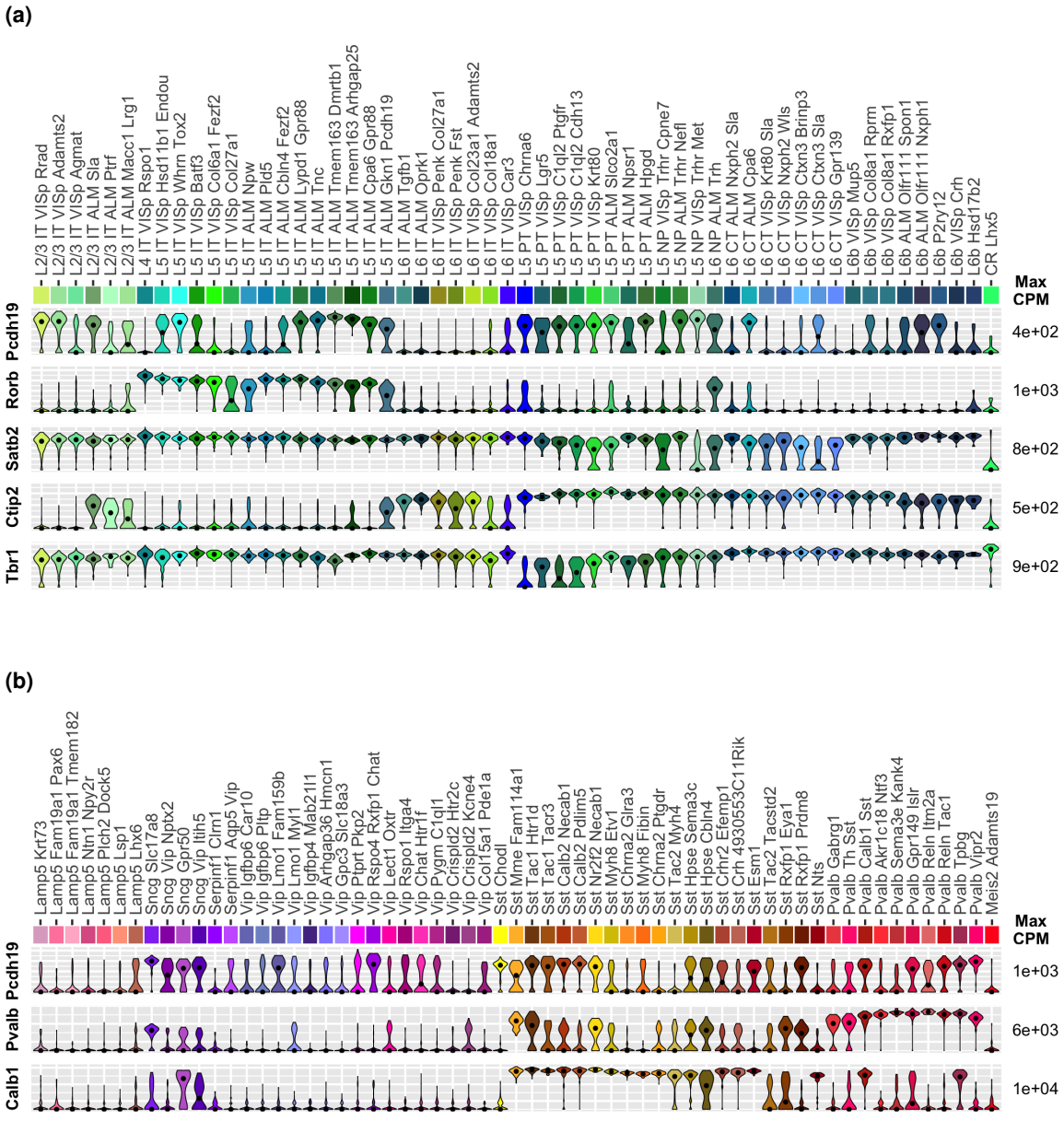


Figure 1.6: *Pcdh19* expression in the glutamatergic and GABAergic clusters determined by Tasic *et al.* scRNAseq study. Black dots are medians. Counts are normalised to CPM and displayed on a $\log_{10}$ scale (Galindo-Riera et al., 2020).

co-localisation seen in the SATB2+ callosal projection neurons (Galindo-Riera et al., 2020). This study shows for the first time that *Pcdh19* is expressed not only by excitatory, but also by inhibitory neurons. Single-cell RNA-sequencing (scRNA-seq) technology has allowed to interrogate the transcriptional signature of individual cells and is therefore especially valuable in unravelling cellular transcriptional profiles in highly complex tissues such as the brain. Deep sequencing of 23,822 cells from the adult mouse primary visual (VISp) and anterior lateral motor (ALM) cortex led to the identification of 133 distinct transcriptomic cell types, of which 56 are excitatory neurons, 61 inhibitory and 16 non-neuronal (Tasic et al., 2018). Whilst GABAergic neuronal clusters are found equally in both brain regions within the excitatory population of neurons there are region specific clusters (Tasic et al., 2018). The Tasic data set is available via the Allen Brain Institute (Tasic et al., 2018). Re-analysis of this data set, in order to unravel what specific neuronal cell-types express *Pcdh19*, confirms what has been previously reported but also provides new interesting insights (Galindo-Riera et al., 2020).

In VISp and ALM *Pcdh19* is expressed in both GABAergic and glutamatergic neurons and has very low expression in non-neuronal cells. Within the glutamatergic population, *Pcdh19* is found primarily in L2/3 intratelencephalic (IT), L5 IT, L5 pyramidal tract (PT) and near-projecting (NP) neurons. Neurons in L2/3 display lower expression compared to L5 neurons . L5 ALM Tmem163 Dmrtb1 have the highest expression. Remarkably, one cluster of L5 IT ALM neurons is defined by expression of *Pcdh19* and *Gkn1* (**Figure 1.6a**). Within the GABAergic population, *Pcdh19* is mainly expressed by *Pvalb*, *Sncg* and some *Sst* interneurons. *Pcdh19* is only weakly found in *Serpin1* and *Vip* neurons. Chandelier and Basket cells are the interneuron populations that show highest expression (**Figure 1.6b**). Interestingly, re-analysis of data derived from human *post-mortem* and surgical samples reveals a similar picture, with *PCDH19* being expressed by both glutamatergic neurons and GABAergic interneurons. Chandelier cells are the cell-type with highest expression in both human and mouse (Galindo-Riera et al., 2020).

1.4.2 PCDH19 in neurogenesis and migration

Corticogenesis in the mouse begins between embryonic day 9 (E9) and 10, corresponding to human gestation week (GW) 5-6 with the dramatic expansion of the neuroepithelial sheet that lines the rostral end of the neural tube, the telencephalon. Amazingly, from this single-cell layer of neuroepithelial (NE) cells the whole diversity of neocortical projection neurons will be generated in a regulated and ordered fashion. The NE cells undergo rapid proliferative division and gradually lose their epithelial features, such as tight junctions (Aaku-Saraste et al., 1996), and become radial glial cells (RGC), recognizable by expression of transcription factor *Pax6* (Gotz et al., 1998). RGC are highly polarized progenitor cells, with a long process that spans the whole width of the developing cortex and end-feet structures that anchor them at the ventricular surface via adherens junctions. Like the NE cells, RGCs undergo rapid expansion by symmetric divisions, whilst going through cycles of interkinetic nuclear migration (INM), dividing at the ventricular surface and going through S-phase at the basal surface (Sauer, 1935), (Subramanian et al., 2017) (**Figure 1.7**). The layer of pseudostratified epithelium that results from INM of RGCs, is called the ventricular zone (VZ).

After expansion, around E11.5 in the mouse, RGC switch to a new mode of cell division: neurogenic asymmetric division, whereby each RGC produces one daughter RGC and one daughter neuron (Noctor et al., 2001) (**Figure 1.7**). Asymmetry is also found in the inherited component of each daughter cell, such that, for example, one of the two cells will inherit the radial fiber and/or the apical domain with the adherens junction (Miyata et al., 2001). Asymmetric divisions can also give rise to an other type of progenitor, the intermediate progenitor (Noctor et al., 2004); (Miyata et al., 2004). Intermediate progenitors (IP) detach from the apical surface and form the subventricular zone (SVZ), where they produce neurons via indirect neurogenesis (**Figure 1.7**). Notch (Kawaguchi et al., 2008), Wnt (Munji et al., 2011), (Hirabayashi et al., 2004), (Harrison-Uy and Pleasure, 2012), Shh (Komada et al., 2008) and Fgf (Yoon et al., 2004) are some of the signalling pathways that can influence progenitor behaviour and the outcome of divisions.

As neurogenesis progresses, progenitors become gradually restricted in their competence state (Molyneaux et al., 2007).

As mentioned, PCDH19 is highly expressed in the VZ by rapidly dividing RGCs (Fujitani et al., 2017) (Dr Jessica Griffiths, unpublished) for a transient period, and PCDH19 expression begins to decline in intermediate progenitors. In regions populated by IPCs, *Pcdh19* mRNA expression is downregulated, relative to the high levels expressed in regions populated by RGCs. Along with the data that indicate *Pcdh19* mRNA expression is complementary to/ opposing the neurogenic gradient, these observations indicate that *Pcdh19* may be playing a role in regulating the symmetric proliferative divisions occuring in early RGCs. Already at E11.5 in the *Pcdh19* heterozygous brain, a striking cell sorting of progenitors occur, with the formation of PCDH19-expressing (WT) and PCDH19-non expressing (KO) mosaic columns (Pederick et al., 2018). RGC are anchored to the apical surface of the ventricular zone via strong adherens junctions, which contain the adhesion molecule N-Cadherin (Kadowaki et al., 2007) (Gil-Sanz et al., 2014). As mentioned, PCDH19 forms a complex with N-Cadherin, which has different adhesion properties than N-Cadherin or PCDH19 alone. Differential adhesion via lack of this complex in PCDH19 KO progenitors in the heterozygous brain could be explaining the cell-sorting behaviour. These PCDH19 WT and KO progenitors $\underline{within}$ the heterozygous brain behave differently: KO progenitors have higher mitotic index and higher quitting fraction than WT progenitors, hence PCDH19 KO progenitors tend to divide asymmetrically more often than WT progenitors, which is reflected by a reduction in PAX6+ PCDH19 KO cells (Dr. Jessica Griffiths, unpublished). Therefore PCDH19 WT and PCDH19 KO cells within the heterozygous brains have a differential neurogenic behaviour, suggesting a possible regulatory mechanism is at play.

The cortex is generated “inside-first, outside-last”, whereby neurons that occupy the deep layers are born first, and neurons that occupy the upper layers are born later (Rakic, 1974). Later born neurons therefore need to migrate through the deep-layer neurons in order to populate their correct laminae (Angevine and Sidman, 1961). Initially, newly born neurons can do so

autonomously by somal translocation, but as the width of the developing cortex expands, newly born neurons use the scaffolding provided by the basal process of the radial glial cell to migrate radially (Rakic, 1972) towards their laminar position in a process known as glial guided locomotion (Nadarajah et al., 2001) (**Figure 1.7**). An exception to this rule is represented by Cajal-Retzius cells which originate in several embryonic structures including the caudomedial wall of the telencephalon and migrate tangentially to form the marginal zone (MZ) (Takiguchi-Hayashi et al., 2004) (**Figure 1.7**). From the MZ, Cajal-Retzius cells secret the extracellular matrix protein Reelin which is recognised by newly-born projection neurons and is necessary for their correct migration (Ogawa et al., 1995), (Franco et al., 2011), (Gil-Sanz et al., 2013). Defects in neuronal migration can lead to dramatic disorganisation of cortical layer formation as exemplified by the semi-inverted lamination phenotype of the Reelin mutant mouse (D'Arcangelo et al., 1995), (Caviness, 1976). CAMs are crucial during the migration phase of corticogenesis: newly-born neurons crawl up the basal process of RGC thanks to tensile forces generated by constant remodelling of adhesion points (Martinez-Garay et al., 2016).

PCDH19 involvement in neuronal migration has been investigated by *in utero* electroporation of both reporter plasmids in *Pcdh19* mutant animals, and of shRNAs against *Pcdh19* in WT animals (Dr. Galindo Riera, unpublished). These experiments revealed defective neuronal migration of both early and late born neurons in PCDH19 KO animals. In line with these findings, increased migration of PCDH19 KO cortical explants has been reported (Pederick et al., 2016). Intriguingly, these defects in migration are not reflected by cortical layer composition as a detailed investigation into the distribution of different cortical layer markers in PCDH19 mutant mice did not reveal any difference in lamination (Galindo-Riera et al., 2020). Nonetheless, in this analysis, the two populations within the heterozygous brain, PCDH19 WT and PCDH19-KO, were not distinguished and a subtle effect in opposing directions might have been undetected.

Many of the basic mechanisms described are conserved between mammalian species, but the human cortex is far more complex than the mouse. Human cortical expansion is based on

the emergence of an extra germinal zone, known as the outer SVZ, in which different types of progenitors, such basal radial glia, give rise to many more neurons in humans (Hansen et al., 2010) resulting in cortical folding (gyrenchephalic *vs* lissencephalic species) (Llinares-Benadero and Borrell, 2019). A recent report investigating cortical structure in *PCDH19*-epilepsy patients by MRI revealed that gyrification is reduced in the limbic areas of *PCDH19*-epilepsy patients, indicating that PCDH19 might be involved in cortical folding in the human brain (Lenge et al., 2020).

Unlike excitatory neurons, cortical GABAergic interneurons are generated in the ganglionic eminences. From there interneurons migrate tangentially towards the cortex in different migratory streams (Anderson et al., 1997) (Marín and Rubenstein, 2001). Interneuron migration is an intermittent process as neurons move in jumps, regulated by cell-intrinsic cues (Silva et al., 2018) (Silva et al., 2019). After birth, interneurons migrate radially to invade the cortical plate and integrate in the correct laminae. Differently than for excitatory neurons, interneuron identity is acquired later, but the mechanism which lead to specification of interneuron diversity are still being elucidated (Lim et al., 2018). Interneurons continue to develop postnatally to integrate into the nascent cortical circuits and at least 30% of interneurons undergo programmed cell-death in the process.

At the end of neurogenesis, around E17-E18 in the mouse, GW20 in humans, gliogenesis begins, with RGCs giving rise to different types of glia: astrocytes, oligodendrocytes, ependymal cells (Malatesta et al., 2000). Interestingly microglia, the brain resident immune cell are originated in the yolk-sac but invade the brain before the onset of neurogenesis (Menassa and Gomez-Nicola, 2018).

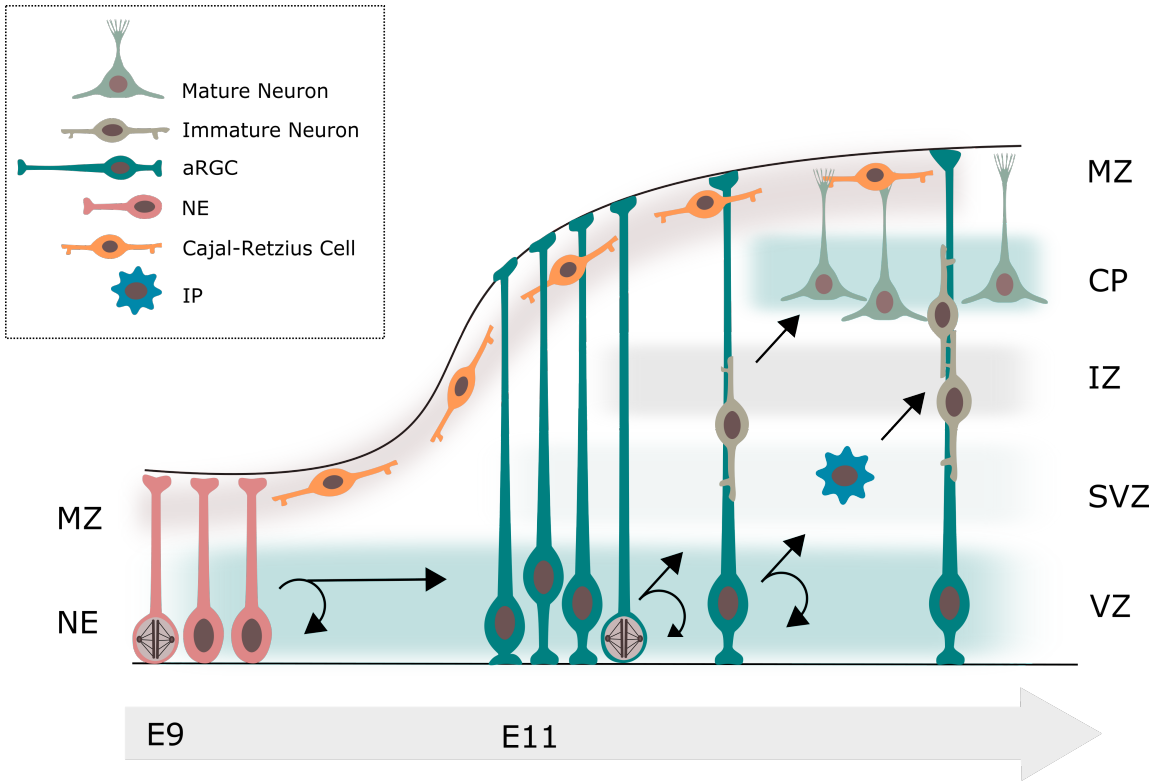


Figure 1.7: Mouse corticogenesis. Summary of the main stages of mouse cortical development. aRGC, apical radial glial cell; NE, neuroepithelial cell; IP, intermediate progenitor; MZ, marginal zone: NE, neuroepithelium; VZ, ventricular zone; SVZ, subventricular zone; IZ, intermediate zone, CP, cortical plate.

1.4.3 Axon, dendrite and synapse morphogenesis

As mentioned earlier, cortical pyramidal neurons are highly polarised with very intricate morphologies, and, as observed *in primis* by Ramon y Cajal, beautifully complex arborisations. After a brief multipolar phase, polarisation of neurons, which entails definition of the axon and dendrite compartments, is initiated during cortical migration. For pyramidal neurons the leading process will become the apical dendrite, whereas the trailing process will become the axon (Hatanaka and Murakami, 2002), (Schwartz et al., 1991). The basal and lateral dendrites eventually emerge from the cell body. Both structural and functional (output *vs* input) differences arise during polarisation. For instance, axons and dendrites are enriched in different microtubular proteins (Mapt in axons and Map2 in dendrites) (Matus et al., 1981). Dendrites also contain both plus- and minus-end microtubules (as opposed to only plus-end in axons) which are necessary for assembly of dendritic branch points (Baas et al., 1988).

So, as projection neurons migrate to their laminar destination, they extend their axons behind (Schwartz et al., 1991). A whole range of extracellular guidance cues, such as semaphorins and netrins, helps the axon navigate to find its synaptic target in the extraordinary process of axonal pathfinding (Polleux et al., 1998) (Serafini et al., 1996). Cues can be extracellular secreted factors or cell-adhesion molecules, and can be both attractive and repulsive. Axons follow the path created by pioneering axons and use the growth cone as a cue-sensing motor structure in order to elongate. Actin-microtubule interactions and cytoskeleton remodelling are key for axon extension (Schaefer et al., 2002). PCDH19 could be playing a role during this phase through interactions with the WAVE regulatory complex via the WIRS domain. PCDH17, for instance, was shown to mediate collective axon extension via the recruitment of the WRC (Hayashi et al., 2014). When axons move in a similar direction they fasciculate and form bundles, eventually giving rise to the main axonal tracts in the cortex: the *corpus callosum*, the corticospinal tract and the corticothalamic tract (Leyva-Díaz and López-Bendito, 2013).

Following extension of the axon, the input receiving part of the neuron develops by progressive outgrowth and branching of the dendritic arbour. Several molecular cues modulate the process and, interestingly, molecules that are chemorepulsive for axons can be chemoattractant for dendrites. Slit-Robo interactions, for example, are both necessary for dendrite morphogenesis and axon extension, (Whitford et al., 2002) and Sema3a functions both as a chemoattractant for the apical dendrites of pyramidal neurons (Polleux et al., 2000) and as a repellent for axons (Polleux et al., 1998). Dendrites extend and branch out to create non-redundant coverage of a specific receptive field. Self-avoidance of dendrites belonging to the same neuron is key to this process, so that dendrites cover a wide field. Dendrites also avoid dendrites of neighbouring neurons, in a process known as tiling (Jan and Jan, 2010).

Finally, dendrites develop spines, small protrusion of different shapes, ranging from thin and long filopodia-like" to short and stubby "mushroom-like". Spines are highly specialised cellular compartments packed with receptors and the necessary machinery to modulate synaptic transmission and are where the vast majority of excitatory synapses are formed (Gray, 1959). Spines first appear postnatally and keep increasing in number for at least one month, followed by a dramatic loss of spines by pruning (Pan and Gan, 2008). Whilst the dendritic tree is more or less stable, spines undergo constant remodelling throughout life based on synaptic activity (Lang et al., 2004). PCDH19, although not exclusively, is found in dendritic spines and a 30% of PCDH19 puncta were found to co-localise with pre or post synaptic markers, or both (Pederick et al., 2016), (Hayashi et al., 2017), (Bassani et al., 2018) (Mincheva-Tasheva et al., 2021). As synaptic contacts are reduced *in vitro* between PCDH19 WT and PCDH19 KO cells (Mincheva-Tasheva et al., 2021), PCDH19 seems to also be important for synapse formation.

1.5 Aims and hypothesis

As explained throughout this introduction, PCDH19 is thought to have a role throughout cortical development, starting at the onset of neurogenesis. And whilst complete lack of the protein is not detrimental, as seen by asymptomatic KO male carriers, cellular mosaic of PCDH19 expressing and non-expressing cells in the developing cortex leads to asynchrony in neurogenesis and abnormal, columnar segregation of progenitors. The role of PCDH19 in cortical migration is more controversial because although *PCDH19*-epilepsy patients can have morphological abnormalities such as cortical dysplasia, PCDH19 mutants animals have no defects in cortical lamination. But, as animal models of *PCDH19*-epilepsy have differences in behaviours relating to anxiety and response to novel environments, there might more subtle alterations at the circuit level, suggesting a yet undescribed role for PCDH19 in circuit formation and function. Finally, as discussed, PCDH19, seems to be able to modulate gene expression. Therefore, we hypothesised that PCDH19 could be subject to proteolytic processing and this could represent the first step in a signalling cascade. Downstream effects on gene expression could be altering neuronal morphology and/or function and contributing to the pathology. This could potentially be a mechanism of modulating plasticity throughout life. In order to address this hypothesis the following aims were set for this thesis, each point representing a result chapter:

(1) Investigate the proteolytic processing of PCDH19 and identifying the proteases involved.

(2) Determine novel interactors of PCDH19.

(3) Establish a cell-culture system to study the function of the cytoplasmic domain of PCDH19.

(4) Determine whether the cytoplasmic domain of PCDH19 can modulate or affect gene expression by analysing the transcriptome of transgenic progenitors and neurons.

Chapter 2

Materials and Methods

2.1 Animals

2.1.1 Husbandry

Animals were housed in cages on a 12 h light/dark cycle and *ad libitum* access to food and drink. All procedures were conducted in accordance with the Animals (Scientific Procedures) Act 1986 (amended 2012), under the Project Licence number PDDC89B6E. All animals were ear notched for identification purposes. Ear notches were stored at -20℃ until needed for genotyping. Animals were housed at a maximum of 5 animals per cage. When animals of a specific embryonic stage were required, plug checking was performed, with noon of the day the plug was found considered as E0.5. C57BL/6J WT mice were purchased from Charles River Laboratories.

2.1.2 *Pcdh19* KO mouse line

Pcdh19 knock-out (KO) mice (TF2108) were purchased from Taconic Biosciences. As previously explained in (**Figure 1.4a**), in this mouse line exons 1-3 of *Pcdh19* are replaced by a β-galactosidase (β-gal)/neomycin (neo) reporter cassette. The mouse line has been characterized anatomically and behaviourally (Pederick et al., 2016), (Galindo-Riera et al., 2020).

2.1.3 Genotyping

Genotyping was done using the Mouse Direct PCR kit (Biotool, cat no. B4001), following the manufacturer's instructions. Genomic DNA was extracted from earclips or post-mortem tailclips in 100 μl Buffer L with 2 μl of Protease K Plus at 55℃ for 30 minutes. Heat inactivation was done for 5 minutes at 95℃. DNA samples were analysed by PCR. Primer sequences and product sizes are listed below (**Table 2.1**).

Table 2.1: *Pcdh19* genotyping primers

Primer	Direction	Product Size.	Sequence (5'-3')	Annealing Temp
Pcdh19 WT	F	123 bp	TAGAGGTTCTTGCTGAAGACTTCC	56.5℃
Pcdh19 WT	R	123 bp	TCAACTGTTTCGATGAGACACTGC	56.5℃
Pcdh19 Mut	F	437 bp	GTGCGTACCAGGCGGGAGC	57.2℃
Pcdh19 Mut	R	437 bp	CCCTAGGAATGCTCGTCAAGA	57.2℃

2.1.4 *In utero* electroporation

C57BL/6J WT females were mated to C57BL/6J WT males by pairing overnight and separating once a plug occurred. Experienced mothers were used to increase chances of the litter being cared for, as tissue was to be examined postnatally. At E15.5, pregnant females were deeply anaesthetised and maintained at 2% isoflurane for surgery. Surgeries were performed by Dr. Cristina Llinares Benadero. The abdominal cavity was opened and the uterine horns, containing the embryos were temporarily taken out of the womb. Plasmid DNA was injected in the lateral ventricle using pulled glass capillaries. Current was used to transfer the plasmid DNA into the progenitor cells, with 50 ms electric pulses of 35-45 V. The plasmids used were pCIG and pCIG-*Pcdh19*-CYTO-HA, described in (**Table 2.2**), electroporated at a concentration of 1 μg/μl, diluted

in 1X PBS and coloured with 0.5% of Fast Green (Sigma-Aldrich). Following surgery, animals were monitored daily until birth. Litters were kept until P60, then perfused for tissue processing.

Table 2.2: PCDH19 plasmids for IUE and cell-culture experiments

Plasmid	Source	Function
pCIG-*Pcdh19*-CYTO-HA	IMG	Chicken-beta-actin promoter driven expression of the cytoplasmic domain of PCDH19, C' terminally HA-tagged, followed by IRES-GFP
pCIG-*Pcdh19*-HA	IMG	Chicken-beta-actin promoter driven expression of full-length PCDH19, C' terminally HA-tagged, followed by IRES-GFP
pCIG	IMG	Chicken-beta-actin promoter driven expression of enhanced GFP
pCBA-*Pcdh19*-CYTO-HA	IMG	Chicken-beta-actin promoter driven expression of the cytoplasmic domain of PCDH19, C' terminally HA-tagged
pCBA-*Pcdh19*-CYTO-HA	IMG	Chicken-beta-actin promoter driven expression of full-length PCDH19 C' terminally HA-tagged
pCBA	IMG	Empty backbone vector used as control
mbEGFP	IMG	Membrane-bound enhanced GFP

2.2 Tissue processing

2.2.1 Perfusion

Animals were injected with 100 μl of Euthatal (Merial, R02701A). Once the reflexes were gone, animals were transcardially perfused with 30 ml of 1X PBS, followed by 30 ml of 4$\%$ PFA. Brains were extracted by careful dissection and placed in 4$\%$ PFA overnight at 4℃, then washed in 1X PBS the following day and stored at 4℃ in the dark until sectioned.

2.2.2 Sectioning

Brains were embedded in 4% Top Vision low melting-point agarose (Thermo Fisher Scientific, R0801) in 1X PBS and cut on a vibrating blade microtome (Leica VT1000S). 350 μm sections were used for the dendritic arborisation and spine density analysis. Slices were collected, processed via IHC if needed, or counterstained directly with DAPI (1:4000 in 1X PBS) and mounted on glass slides for imaging.

2.2.3 IHC and ICC

Brain sections, or cells on glass coverslips, were washed in 1X PBS for a minimum of 3 times, followed by several washes in 0.25% PBS-T. Sections, or cells were incubated at RT for at least 3 hours in BSA/blocking solution in 0.25% PBS-T, then incubated with the primary antibodies overnight at 4℃ in the dark. The following day, sections or cells were washed, and incubated with appropriate fluorescently-conjugated secondary antibodies, washed and counterstained with DAPI (1:4000 in 1X PBS) and mounted with DAKO mounting media on glass slides. Anti-HA antibody (clone 3F10, Roche) was used at 1:500.

2.2.4 Microscopy

Brain sections or cells were imaged on the confocal microscope (Carl Zeiss, LSM 780) with Zen Black software (version 2.0, Carl Zeiss). For reconstruction of the neuronal morphology, 1 μm

spaced Z-stack tiles, including the whole neuron, were taken with a 40X water-immersion objective to be processed via Imaris Filament Tracer software. For spine analysis, representative segments of secondary order apical or basal dendrites belonging to the previously imaged neurons, were imaged with a 63X oil-immersion objective, as 1 μm spaced Z-stacks. Images of cultured cells, HEKs and mESC-derived neurons were also taken with a 63X oil-immersion objective, as 1 μm spaced Z-stacks.

2.3 Cell culture

2.3.1 Mycoplasma testing

Cells in culture were routinely tested for mycoplasma infection with Lookout Mycoplasma PCR detection kit (Sigma, MP0035), following the manufacturer's instructions. Unlike other infections that compromise health of cultured cells, mycoplasma is too small to be detectable by visual inspection using a microscope but can have massive effects on gene transcription. It is therefore vital to keep an active monitoring strategy by testing regularly.

2.3.2 Mouse embryonic fibroblasts

WT and ADAM10 KO mouse embyonic fibroblasts (MEFs), a generous gift of Prof. De Strooper's laboratory, were cultured in CA media (DMEM + 1% non-essential amino acids + 1% L-Glutamine + 10% FBS heat inactivated + 1.43 mM β-mercaptoethanol), on 0.2% gelatin-coated (60 mm, 100 mm; Nunc) plates. MEFs were split every 2-3 days depending on confluence. Early-passage stocks of cells were kept in liquid nitrogen until needed.

2.3.3 HEK293 cells

HEK293 cells were maintained in CA media (DMEM + 1% non-essential amino acids + 1% L-Glutamine + 10% FBS heat inactivated + 1.43 mM β-mercaptoethanol) on 100 mm (Nunc) dishes.

When needed for an experiment, they were split and plated in 12/24-well plates (Thermofisher Scientific) on autoclaved glass coverslips, pre-coated with 0.5 mg/ml poly-lysine.

2.4 Neuronal differentiation of mESCs

E14 male mouse embryonic stem cells (mESCs) used throughout this thesis were kindly provided by Dr. Xinsheng Nan (Cardiff University). Differentiation into cortical-like neurons was done following a well established protocol (Bibel et al., 2004) (Bibel et al., 2007), graphically summarised below (**Figure 2.1**). WT, PCDH19 KO and PCDH19 CYTO mESC were all differentiated following this protocol. Unlike other mESCs, E14 are feeder-independent, which means they don't require feeder cells to grow. Nonetheless, culturing with feeder cells can improve their quality. Therefore, in some instances, feeder passaging was added to the differentiation protocol.

2.4.1 Culture of mESCs

ESC vials were kept in liquid nitrogen. When needed, cells were quickly thawed by immersion in a 37℃ water bath and resuspended in fresh ESC medium (DMEM + 1% non-essential amino acids + 1% L-Glutamine + 10% FBS heat inactivated + 1,43 μM β-mercaptoethanol + 1000 U/ml LIF). Cells were grown on 0.2% gelatin-coated dishes (60mm, 100mm; Nunc) and passaged three times on a splitting ratio between 1:5 and 1:7 depending on rate of growth, before moving onto aggregate formation. Medium was never allowed to exhaust.

2.4.2 Growing mESCs on feeder cells

WT MEFs, derived in our lab from C57BL/6J embryos by Dr. Jessica Griffiths, were used as feeder cells for E14 mESCs. MEFs were grown as previously described (**Section 2.3.2**). Once confluent, MEFs were inactivated by addition of 1 μg/ml of mitomycin C (Sigma). After 2 hours, the medium was changed and MEFs were left to recover for a minimum of 1 hour, before plating ESCs on top. ESCs were seeded at 3×10^6/60 mm dish. On alternate days, ESCs were passaged

at a splitting ratio between 1:7 and 1:10 in ES medium (DMEM + 1% non-essential amino acids + 1% L-Glutamine + 10% FBS heat inactivated + 1,43 μM β-mercaptoethanol + 1000 U/ml LIF) until a homogenous, rapidly proliferating population of defined colonies was seen, for about 5-6 passages. Following this, ESCs were passaged on gelatin for 2-3 passages, in order to remove MEFs before continuing to the formation of aggregates.

2.4.3 Formation of cellular aggregates

Using the NucleoCounter ® NC-100™ (Chemometec), 4×10^6 ESCs were counted and transferred to uncoated, non-adherent UV-sterilised bacteriological dishes for aggregation. As cells don't adhere to the dish, they rapidly form 3-dimensional structures, called cellular aggregates (CAs). CAs were grown in 15 ml of CA medium (DMEM + 1% non-essential amino acids + 1% L-Glutamine + 10% FBS heat inactivated + 1.43 mM β-mercaptoethanol), which was changed every two days, at CA2, CA4 and CA6. At timepoints CA4 and CA6, medium was also supplemented with 5 μM retinoic acid. About 2×10^7 cells should be obtained at CA8, when dissociating the aggregates.

2.4.4 Poly-DL-ornithine/laminin coating of plates

Depending on the experiment, different plate sizes were used. For protein and RNA extraction 12-well plates were used (Theromofisher Scientific), 4-well plates (Theromofisher Scientific) for transfections and immunocytochemistry. Plates were precoated starting at CA6: plates were incubated with poly-DL-ornithine (0.5 mg/ml in borate buffer (150mM BH_3O_3, pH 8.3 adjusted with NaOH), filtered and diluted 1:5 with dd$H_2$0) at 37℃ overnight. The next day, they were washed three times with distilled water and coated with 5 μg/ml laminin (Life Technologies) dissolved in 1X PBS at 37℃ overnight.

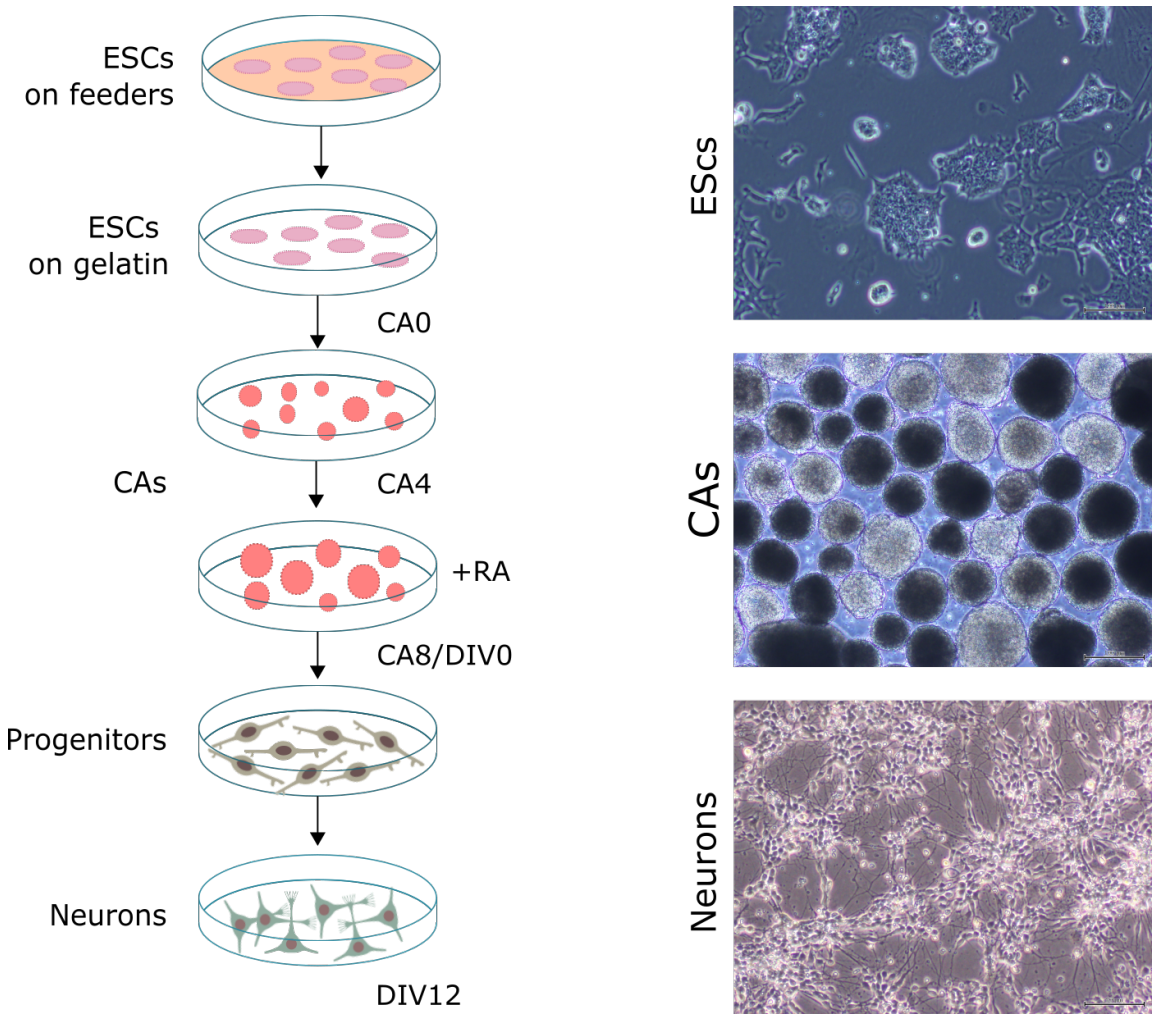


Figure 2.1: mESC neuronal differentiation protocol. mESCs are differentiated *in vitro* via formation cellular aggregates (CAs) and using retinoic acid (RA) to drive neuronal fate.

2.4.5 Dissociation and neuronal differentiation

At CA8, 8 days after the initial formation, each plate of aggregates was dissociated with 1 ml of freshly prepared 0.05% trypsin in 0.05% EDTA/PBS (3 minutes/37℃). Trypsin was then inactivated with 10 ml of CA medium and aggregates were dissociated by pipetting gently, passed through a 40 μm nylon strainer and centrifuged (5 minutes/180 rpm). Supernatant was carefully removed, and progenitors were resuspended in N2 medium (DMEM/F12 medium (Life Technologies) + 1% N2 supplement (Life technologies) + 1% non-essential amino acids + 1mM Glutamax, 2.5 μg/ml insulin (Sigma-Aldrich) + 100 μM 2-β-mercaptoethanol + 100 U/ml Pen/Strep (Life Technologies)). Cells were counted using NucleoCounter ® NC-100™ (Chemometec) and plated at different densities, depending on the downstream application, in a range between 7.5×10^5 and 1.5×10^6 cells per well. After counting, cells were plated on the precoated plates, taking care to not dry the laminin. After 24 hours, medium was changed to fresh N2 medium. After 48 hours, at DIV2, medium was changed to home-made complete medium. The complete medium main ingredients mix is prepared in 100x batches and stored until needed at -80℃ (**Table 2.3**). When needed, complete media is prepared by dissolving BSA, transferrin and insulin in DMEM (**Table 2.4a**), then combining them with the final components listed (**Table 2.4b**). Between DIV2 and DIV4, complete medium was additionally supplemented with 5 μg/ml 5-Fluorodeoxyuridine (5-FdU) (Sigma) to stop potential growth of non-neuronal cells. Complete medium was changed again at DIV4, DIV8 and DIV12.

Table 2.3: Complete medium components and concentrations for 100X stock preparation

Components	100X (μg/ml)	200 ml stock of 100X (mg)	Suspension	
L-Alanine	200.00	40.0000	Dissolved in 26.6ml ddH_20	Added to 172ml DMEM
Biotin	10.00	2.0000		
L-Carnitine	200.00	40.0000		
Ethanolamine	100.00	20.0000		
D-Galactose	1500.00	300.0000		
L-Proline	776.00	155.2000		
Putrescine	1610.00	322.0000		
Na-Pyruvate	2500.00	500.0000		
Na-Selenite	1.60	0.3200		
Vitamin B12	34.00	6.8000		
Zinc sulfate	19.40	3.8800		
Catalase	256.00	51.2000		
Glutathione	100.00	20.0000		
Linoleic acid	100.00	20.0000	Dissolved in 1.4ml EtOH	
Linolenic acid	100.00	20.0000		
Progesterone	0.63	0.1260		
All-trans retinol	10.00	2.0000		
Retinylacetate	10.00	2.0000		
Tocopherol	100.00	20.0000		
Tocopherolacetate	100.00	20.0000		

Table 2.4: Preparation of fresh complete medium

(a) BSA/Transferrin/Insulin mix in 30 ml of DMEM for making fresh 1X complete medium

Component	mg
BSA	1000.00
Transferrin	20.00
Insulin	16.00

(b) Final mix for complete medium preparation

Final Mix	ml
100X Complete medium stock	4.00
BSA/transferrin/insulin mix	30.00
2.5 mg/ml Superoxidase dismutase	0.40
100X Pencillin/streptomycin	4.00
DMEM	358.00

2.5 Genetic engineering of ESCs

2.5.1 Cloning of targeting vectors

The *Rosa26-Pcdh19*-CYTO-HA targeting vector plasmid was generated starting from the pZDRosa-floxedNeo-IRES-EGFP plasmid (**Table 2.8**). 2 μg of vector was linearized and the IRES-EGFP fragment was excised by double restriction digestion with *BsrGI*-HF and *AscI* (4 hours/37℃). In parallel, the *Pcdh19*-CYTO-HA fragment was amplified from the pre-existing plasmid pCIG-*Pcdh19*-CYTO-HA (**Table 2.2**) using the *Rosa26*-CYTO-HA-F2 and *Rosa26*-CYTO-HA-R primers (**Table 2.5**) and Phusion polymerase. The resulting PCR product was cloned between the 3' and 5' arms of the *Rosa26* targeting vector using In-Fusion (Takara Bio) technol-

ogy. A 2:1 molar ratio of insert:vector was combined in a 10 μl reaction, following manufacturer's instructions. 5 μl of the reaction were then transformed in 50 μl of Stellar Competent cells as follows: cells were thawed 30 minutes on ice, heat-shocked for 45 seconds at 42℃, transferred back into ice for 2 minutes, diluted in 500 μl of SOC medium, shaken for 1 hour at 37℃ and finally plated on agar + ampicillin plates. 10 colonies were picked and regrown overnight in LB + ampicillin (37℃ /200 rpm). The following day, using the QIAprep Spin miniprep kit (QIAGEN, 27104) plasmid DNA was extracted, according to the manufacturer's protocol. To check if the plasmids contained the correct insert, they were analysed by *HindIII* restriction digestion and run on a 1% agarose gel. Confirmed plasmids containing the right insert were further verified by sequencing (Eurofins Genomics). One of the verified plasmids was re-transformed in DH5α cells for plasmid DNA extraction via Maxi Prep (QIAGEN, 12362).

Using the online CRISPR design tool CRISPOR (crispor.tefor.net, by Maximilian Haeussler and Jean-Paul Concordet (Concordet and Haeussler, 2018)), two gRNAs were designed to bind the beginning of exon 1 of *Pcdh19* (sequence shown in (**Table 2.6**)) The gRNAs were cloned separately into the Cas9-mCherry fusion plasmid (pU6-(*BbsI*)-CBh-Cas9-T2A-mCherry) (**Table 2.8**). In brief, the pU6-(*BbsI*)-CBh-Cas9- T2A-mCherry plasmid was linearized by *BbsI*-HF restriction digestion. The forward and reverse oligos encoding for the gRNAs were annealed and then phosphorylated with PNK. Finally the oligos were ligated into the vector via T4 DNA ligase. Colonies were screened by *SacI* digestion and sequenced with sg-seqR (5'-GTACCTCTAGAGCCATTTGTC -3') by Eurofins Genomics.

Table 2.5: In-Fusion cloning primers. Primers used to clone *Pcdh19*-CYTO-HA within the *Rosa26* targeting vector. The 15 bp in bold correspond to end of the vector.

Primer	Sequence (5'-3')
Rosa26-CYTO-HA-F2	**ACCTCGAGTGGCGCG**CCGCGCAGCCATGGCAATGGCAATCAAATGC
Rosa26-CYTO-HA-R	**CCGCTTTACTTGTAC**TCAAGCGTAATCTGGAACATCGTATG

Table 2.6: CRISPR guides. Sequences for sense and antisense oligos for the creation of gRNA1 and gRNA2. In green is the target sequence, the underlined C/G is added to increase transcription from the U6 promoter and in red the overhangs which are necessary for cloning.

Primer	Sequence (5'-3')
gRNA1 S	CACCGCGGGACGGTGATCGCTAACG
gRNA1 AS	AAACCGTTAGCGATCACCGTCCCGC
gRNA2 S	CACCGTTTCGCATTACGGCTCTCGA
gRNA2 AS	AAACTCGAGAGCCGTAATGCGAAAC

2.5.2 Nucleofection

E14 mESCs were passaged the day before nucleofection. On the day, they were trypsinised and counted with the NucleoCounter ® NC-100™ (Chemometec). 4×10^6 cells were used for one round of nucleofection. In brief, cells were pelleted and resuspended in 100 μl of P3 transfection solution (which contains 82 μl Amaxa Buffer and 18 μl P3 supplement; Lonza) and 10 μl of DNA mix. For zinc finger targeting the following concentrations of plasmids were used: 10 μg of the targeting construct (pZDRosa-floxedNeo-*Pcdh19*-CYTO-HA) and 1 μg each of the two zinc finger nucleases (pCMV-RosaR4 KKR mutations, containing the right ZFN (ZFN-R) and the pCMV-RosaL6 ELD mutations, containing the left ZFN-L (ZFN-L)) (**Table 2.7**). For CRISPR targeting the following concentrations of plasmids were used: 1 μg each of the two gRNA-Cas9 plasmids (pU6-gRNA1-CBh-Cas9-T2A-mCherry and pU6-gRNA2- CBh-Cas9-T2A-mCherry) (**Table 2.8**) in 10 μl. Cells were nucleofected using the 4D-Amaxa Nucleofector X-unit (Lonza) and the CG104 programme. Immediately after nucleofection cells were either plated at low density for antibiotic selection (ZFN targeting, (**Subsection 2.5.3**)) or individually sorted by FACS (CRISPR targeting, (**Subsection 2.5.4**)), procedures described below. For removal of the neomycin resistance cassette 10 μg of the pCIG-CRE plasmid (**Table 2.7**) was nucleofected, as described above. Flow diagram illustrating the two precedures is shown below (**Figure 2.2**).

Table 2.7: Plasmids for ZFN targeting of *Rosa26* locus. (#) indicated Addgene number. (*) indicates plasmid generated by Dr. Xinsheng Nan, Cardiff University. (**) indicates plasmid was generated by Dr. Cristina Gil-Sanz (University of Valencia).

Plasmid	Source	Function
pCMV-RosaR4 KKR mutations	#37199	Right zinc finger nuclease for *Rosa26* locus
pCMV-RosaL6 ELD mutations	#37198	Left zinc finger nuclease for *Rosa26* locus
pZDRosa-floxedNeo-IRES-EGFP	*lab-made	Targeting vector for *Rosa26* locus with EGFP
pZDRosa-floxedNeo-*Pcdh19*-CYTO-HA	lab-made	*Pcdh19*-CYTO targeting to *Rosa26* locus
pCIG-CRE	**lab-made	Chicken-beta-actin driven expression of CRE-IRES-EGFP

Table 2.8: Plasmids for CRISPR/Cas9 targeting of *Pcdh19*. (#) indicates Addgene number.

Plasmid	Source	Function
pU6-(*BbsI*)-CBh-Cas9-T2A-mCherry	#64324	Cas9 and mCherry expression with *BbsI* cloning site for gRNA
pU6-gRNA1-CBh-Cas9-T2A-mCherry	lab-made	*Pcdh19* gRNA1, Cas9 and mCherry expression
pU6-gRNA2-CBh-Cas9-T2A-mCherry	lab-made	*Pcdh19* gRNA2, Cas9 and mCherry expression

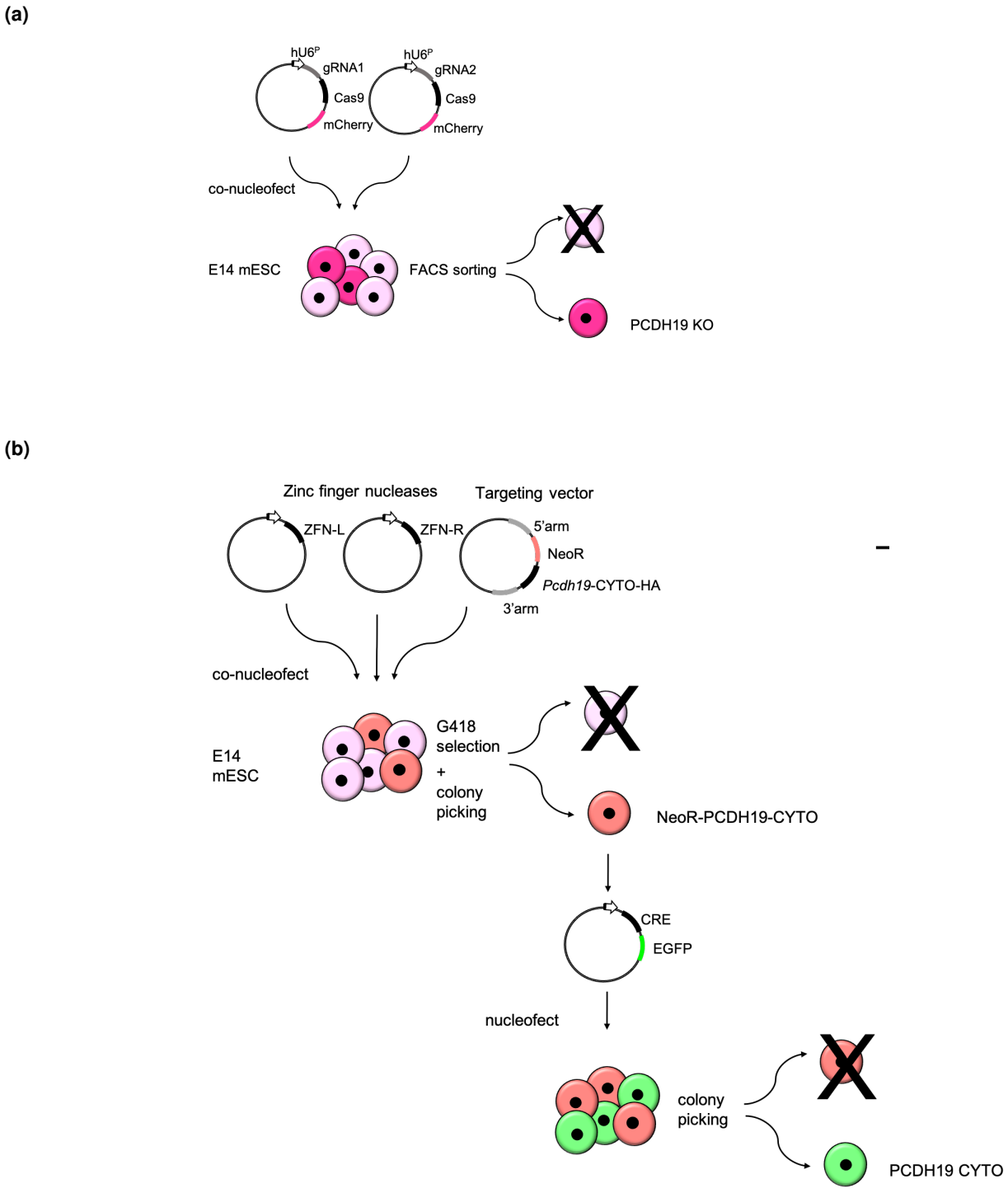


Figure 2.2: Flow diagram of mESC engineering via nucleofection. (a) Generation of PCDH19 KO cells via nucleofection of pU6-gRNA1-CBh-Cas9-T2A-mCherry and pU6-gRNA2-CBh-Cas9-T2A-mCherry and subsequent selection of mCherry+ cells via FACS. **(b)** Generation of PCDH19 CYTO cells via nucleofection of ZFN-R, ZNF-L and *Rosa26-Pcdh19*-CYTO-HA targeting vector followed by antibiotic selection. Cells are further nucleofected with pCIG-CRE in order to remove the selection cassette.

2.5.3 Antibiotic selection and colony picking

After ZFN targeting, nucleofected cells were suspended in 10 ml of ESC medium and plated at different densities (ranging between 0.625 and 2.5 ml/10 cm dish). Cells underwent a 10-day selection process with 250 μg/ml of G418 (Geneticin). Media was changed every two days, replenishing the antibiotics. After about 10 days, or when they were visible by the naked-eye, 100 colonies were manually picked. In brief, cells were incubated for a couple of minutes with 0.01% trypsin (0.05% trypsin, diluted in PBS) in order for colonies to detach from the plate but without dissociating. Colonies were then carefully transferred, using a 200 μl pipette tip, to individual wells in a 96-well plate, trypsinized with 0.05% trypsin, resuspended and transferred to a 24-well plate to grow. Clones were expanded for DNA and protein extraction and then frozen. For ESC subcloning, after nucleofection for removal of the selection cassette, cells were plated at a density of 300 cells/10 cm plate. 24 colonies originating from different clones were picked as described above. Once expanded, these clones were "reverse selected" to test out loss of antibiotic resistance, by plating clones in two wells, one with and one without antibiotics.

2.5.4 Cell sorting

After nucleofection with the CRISPR/Cas9 constructs, cells were plated, to allow for expression of the mCherry reporter. 24 hours post nucleofection, cells were trypsinised, counted and resuspended at a density of 1×10^6 cells/ml in ESC-medium with 1% FBS. Cells were transported to the Sir Martin Evans Builiding for sorting using the BD FACS Aria Fusion cell sorter (BD Biosciences). Operation of the cell sorting machinery was carried out by Mark Bishop, Cardiff University. Using a specific nozzle, mCherry expressing cells were individually sorted into 0.2% gelatin pre-coated and ESC medium-filled 96-well plates. Cells were returned to the incubator as quickly as possible and allowed to recover and grow into colonies, for about 10 days. Once colonies appeared, they were passaged into 24-well plates, and further amplified for extraction of genomic DNA and protein.

2.5.5 DNA extraction and genotyping

Cells were pelleted, resuspended in 500 μl of cell lysis buffer (10 mM Tris, pH 8.0, 1 mM CaCl2; 100 mM NaCl; 0.5% SDS; 5 mg/mL proteinase K (Promega)) and incubated overnight at 50℃ in a heat-block. The following day, 500 μl of 100% isopropanol and 50 μl of 3 M NaOAc were added to precipitate DNA. DNA was pelleted by centrifugation (15 minutes/top-speed), washed with 70% ethanol and resuspended in 30 μl of TE buffer (Qiagen). Clones were genotyped by PCR, using the long-range sequal prep PCR kit (Thermo Fisher Scientific). Primers were designed around the homology arms to test for correct insertion (**Table 2.9**), or around cut site, (**Table 2.10**). Sequencing of samples was done by Eurofins Genomics.

Table 2.9: Genotyping primers for *Rosa26 Pcdh19*-CYTO targeting

Primer	Location	Sequence (5'-3')
ReverseR26OUT2	5'arm genomic	CAAGCGGGTGGTGGGCAGGAATGCG
Neo-pR2	5' arm selection cassette	TCGGCAGGAGCAAGGTGAGATGAC
ForwardR26OUT2	3' arm genomic	ACCAGAAGAGGGCATCAGATCCCATTAC
gen19-ICDF2	3' arm intracellular domain	GCGTGAAGCGTCTGAAGGATATCGTTC

Table 2.10: Sequencing primers for PCDH19-KO targeting

Primer	Location	Sequence (5'-3')
gRNA1 F	gRNA1 site	CTCCAGCTCTATCTGTGCGG
gRNA1 R	gRNA1 site	TACTCGGTCGAAGAGGAGCA
gRNA2 F	gRNA2 site	CTCATCTGGATCGCTGGCAT
gRNA2 R	gRNA2 site	AGCTGGAGATCTCTGAGGCA

2.5.6 Karyotyping

ESC clones to be karyotyped were incubated with demecolcine solution (0.1 μg/ml) for two hours in the incubator. Cells were then trypsinised, pelleted and washed by centrifugation in 1X PBS twice. Cells were then resuspended in 2 ml of 1X PBS and 6 ml of hypotonic 0.0375 M potassium chloride solution and incubated for 12 min at 37℃. Cells were then pelleted, supernatant removed and a 3:1 volume:volume ratio of cold methanol/acetic acid mixture (-20℃) was added dropwise. Cells were incubated for 20 min at room temperature. Then cells were pelleted again, supernatant removed, and fresh methanol/acetic acid was added. Cells were centrifuged one last time, the supernatant was removed, this time leaving about 100 μl, in which the cells were resuspended. Finally cells were dropped from about 20 cm height on glass slides. Slides were left to dry, stained with DAPI (1:4000 in ddH_2O), coverslipped and imaged immediately. A minimum of 10 cells were imaged for each clone. Chromosomes were then counted using the ImageJ (Fiji) cell-counter plug-in.

2.6 Drug treatments on cells

Treatments used on MEFs and mESC-derived neurons are listed below. (**Table 2.11**).

Table 2.11: List of compounds used for cell treatments

Drug	Supplier	Cat. No.	μM	solubility
DAPT	Sigma	D5942	10	DMSO
GI254023X	Sigma	SML0789	10	DMSO
Ionomycin	Sigma	I3909	5	DMSO
Bicuculline	Sigma	14340	10	DMSO
NMDA	Sigma	M3262	50	Water
(+)-MK 801 maleate	Tocris	0924	1	Water

2.7 Cell lysis for RNA/protein extraction

For RNA extractions, neurons were rapidly washed twice with 1X PBS and lysed in 350 μl of cold RLT buffer (Qiagen) with 1% β-mercapthoethanol. Neurons were scraped off the well with a new scraper and collected in a clean eppendorf immediately on ice. For the RNA sequencing samples, RNA was rapidly transferred and stored at -80℃ until all samples for sequencing were collected. For protein extractions, cells were rapidly washed twice with 1X PBS and lysed in 100 μl of RIPA buffer freshly supplemented with protease and phosphatase inhibitors (50 mM Tris-HCl, 150 mM sodium chloride, 1 mM EDTA, 1% Triton X-100, 0.2% Sodium deoxycholate supplemented with: 1.5 mM aprotinin, 100 mM 1-10 phenantroline, 100 mM 6-aminohexanoic acid, 1% protease inhibitor cocktail (Sigma), 1% phosphatase inhibitor cocktail (Sigma)). Lysates were kept on ice for 30 minutes with brief vortexing every 5 minutes. Samples were then centrifuged at 14000g for 10 minutes and supernatant was transferred to a clean tube. Samples were aliquoted and stored at -80℃ until used for western blotting.

2.8 Immunoprecipitation

For immunoprecipitation, tissues or cells were lysed in freshly made IP lysis buffer (20 mM Tris-HCl, 150 mM NaCl, 1mM EDTA, 1% Triton X, 10 mM NaF, 1mM Na_3VO_4,1% protease inhibitor cocktail (Sigma), 1% phosphatase inhibitor cocktail (Sigma)). 10 μl of Protein G Sepharose beads were washed twice with 500 μl of cold 1X PBS by centrifugation (2000g, 2min, 4℃). In parallel, tissue or cell-lysate samples were centrifuged (14000g, 10 min, 4℃). Supernatant from the samples, corresponding to the protein fraction, was added to the washed beads for pre-clearing (removal of unspecific binding) (30 min, 4℃, constant rotating). Beads and unspecifically bound proteins were precipitated by centrifugation (2000g, 2 min, 4℃). 10% of sample supernatant was put aside and saved to be used as INPUT control. The remaining 90% of the supernatant was used for immunoprecipitation and added to 20 μl of pre-washed Protein G Sepharose beads with 2 μl of antibody of interest (PCDH19, HA or MYC; details of antibodies provided later (**Table 2.12**))

and incubated (2 hours, 4℃, constant rotating). Beads and antibody complex were precipitated by centrifugation (2000g, 2 min, 4℃) and washed with lysis buffer (3X 2000g, 2min, 4℃). Finally, samples were eluted in LDS buffer (for 500 μl: 250μl of 4X LDS, 50μl of DTT and 150 μl of water) and incubated for (10 min, 70℃) for releasing the antibody from the beads. Finally, the beads were removed by centrifugation (2000g, 5min, RT). Samples were stored at -80℃ until analysed by western blot (**Section 2.10**), or mass spectrometry (**Section 2.12**).

2.9 Membrane fraction enrichment

Cells lysates were processed using the Mem-PER™ Plus Membrane Protein Extraction Kit (Thermo Fisher Scientific, 89842) following the manufacturer's instructions.

2.10 Western blotting

Protein lysates were prepared by addition of LDS buffer and 10% 0.5 M DTT, and boiled at 70℃ for 10 min. Samples were then centrifuged at 14000 g for 10 min and loaded onto a NuPAGE Novex 4-12% Bis-Tris gel (Novex Life Technologies, WC1020) alongside a standard protein ladder, Novex Sharp pre-stained ladder (Invitrogen, LC5800) and run at 120 V for 90 minutes. Proteins were transferred to a nitrocellulose membrane with a 0.2 μm pore size (GE Healthcare Life Sciences, 10600001) by wet transfer at 100 V for 120 minutes. Ponceau stain was routinely used to ensure successful transfer before proceeding further. If the experiment required the use of different primary antibodies, membranes were carefully cut with a scalpel at this point, and processed separately. Membranes were incubated at RT shaking for 1 hour with 4% blocking solution (5% milk powder (BioRad) in TBS-T). Primary antibody incubation was done overnight at 4^{o}C shaking. Details of primary antibodies used are listed below (**Table 2.12**). The following day, membranes were then washed 3 times for 10 minutes in TBS-T and then incubated for one hour at RT with the appropriate secondary antibody (in 5% milk powder in TBS-T blocking). Details of secondary antibodies used are listed below (**Table 2.13**). Membranes were washed again 3 times for 10 minutes

in TBS-T. Blots were finally developed with 1 ml of WesternBright ECL substrate (Advansta) for 60 seconds and immediately imaged with the BioRad imaging machine, using the Image Lab software. Densitometric analysis of western blots was also done using Image Lab (v.6.0.1) (BioRad). Lanes and bands were drawn in the software and adjusted volume intensity of each band was extracted. When calculating proteolytic fragments, full-length protein and fragment band intensity were detected on the same blot. Intensity of the proteolytic fragment was always calculated as a ratio of the full-length protein.

2.11 Statistical Analysis

Statistical analysis was carried out on GraphPad Prism (version 8.4.2). Shapiro-Wilk test was used to test for normality of the data. As all data sets passed the normality test ($P > 0.05$) ANOVA was used for analysis. When one variable was to be compared for more than 2 groups, one-way ANOVA was used, followed by Tukey's correction for multiple comparisons. When comparing more two variables, two-way ANOVA was used, followed by Dunnett's multiple comparisons test. When comparing two groups, two-tailed un-paired t-test was performed.

Table 2.12: List of Primary Antibodies

Antibody	Host	Type	Supplier	Cat. No.	Dilution
Pcdh19 (C-terminal)	Rabbit	pAb	Bethyl	A304-468A	1:1000
Pcdh19 (N-terminal)	Rabbit	pAb	Biorbyt	orb312580	1:1000
N-Cadherin (3B9)	Mouse	mAb	Thermo Fisher	33-3900	1:1000
Adam10	Rabbit	pAb	Abcam	ab1997	1:1000
Pan-Cadherin	Rabbit	pAb	Abcam	ab6529	1:1000
HA	Rat	mAb	Roche	ROAHAHA	1:2000
Myc	Mouse	mAb	Thermo Fisher	MA1-980	1:2000
Histone H3	Rabbit	pAb	Abcam	Ab1791	1:5000
beta-Actin	Mouse	mAb	Abcam	ab8226	1:2000

Table 2.13: List of Secondary Antibodies

Antibody	Host	Type	Supplier	Cat. No.	Dilution
Anti-Rabbit-HRP	Goat	pAb	Promega	W4011	1:20'000
Anti-Mouse-HRP	Goat	pAb	Promega	W4021	1:20'000
Anti-Rat-HRP	Goat	pAb	R&D systems	HAF005	1:20'000

2.12 LC-MS

Liquid-chromatography mass spectrometry analysis was performed by Dr Kate Heesom, at the Bristol University Proteomics Facility. In brief, samples were run on an SDS-PAGE gel and then individual lanes were cut into 3 pieces. Each piece was digested via DigestPro automated digestion unit (Intavis Ltd.) and peptides were fractioned with the Ultimate 300 nano-LC system with an LTQ-Orbitrap Velos mass spectrometer (Thermo Scientific). Results were processed using the Proteome Discoverer software v1.4 (Thermo Scientific). Finally, results underwent a SEQUEST search against the mouse Uniprot database and were filtered at 1% false discovery rate (FDR).

2.13 RNA sequencing

RNA extraction was done with RNeasy Kit (Qiagen) in RNase-free conditions following the user's manual with DNase treatment (Qiagen). Quality control of the samples was done via Tapestation (Agilent Technologies) and RNA integrity number (RIN) was determined for all samples. Concentration of samples was measure by QUBIT. RNA sequencing was done at Cardiff University Genomic Hub. Library Preparation and sequencing was performed by Angela Marchbank. Libraries were prepared following Illumina's TruSeq Stranded mRNA sample preparation guide. In brief, mRNA was purified from total RNA using poly-T oligos, mRNA was then fragmented into smaller fragments and random priming was used for cDNA synthesis. The sequencing was carried out on a Illumina Nextseq 500 platform with 4 cartridges PE (2x75bp) sequencing on high-output 150 cycle V2.5 cartridges. 1% Phix was spiked into each run as per the Illumina recommendations. The samples were pooled to obtain equal reads for each sample with an aim of at least 44 M reads per sample. Sequencing was paired-end. Quality control of sequencing run, such as QC content and sequence duplication, was performed with FastQC by Dr. Daniel Pass, from the School of Biosciences, Cardiff University. Initial bioinformatics analysis, processing of FASTQ files, sequence alignment, generation of counts and differential expression analysis was performed by Dr. Sumukh Deshpande from the College of Biomedical Sciences, Cardiff Univer-

sity. Paired-end reads from Illumina sequencing were trimmed of adaptor sequences with Trim Galore and assessed for quality using FastQC, using default parameters. Reads were mapped to the mouse GRCm38 reference genome using STAR (Dobin et al., 2013) and counts were assigned to transcripts using featureCounts (Liao et al., 2014) with the GRCm38 Ensembl gene build GTF. Both the reference genome and GTF were downloaded from the Ensembl FTP site (http://www.ensembl.org/info/data/ftp/index.html/). Differential gene expression analyses used the DESeq2 package (Love et al., 2014), using the Benjamini-Hochberg correction for multiple testing. Differential gene splicing analyses used the DEXSeq package (Anders et al., 2012), (Reyes et al., 2013) also using the Benjamini-Hochberg correction for multiple testing.

2.14 R packages for plotting

All plotting of RNA sequencing and MS data was done on R (v.4.02) via RStudio (v.1.2.1335). Plotting was done using R package ”ggplot2” (v.3.3.2), in some cases supplemented by other packages such as ”EnhancedVolcano” (v1.6.0), for RNAseq vulcano plots or ”eulerr” (v.6.1.0), for plotting the proportional Venn diagrams. Over-representation analysis and Gene Set Enrichment analysis was done via ”clusterProfiler” (v.3.16.1) (Yu et al., 2012). Dotplots were made via ”DOSE” (v.3.14.0) (Yu et al., 2015). Cnet plots were made via ”enrichplot” (v.1.8.1). ”Pheatmap” (v.1.0.12) was used to draw heatmaps. ”UniprotR” (v1.4.0) was used to retrieve information from the Uniprot Database regarding proteins discovered in the MS analysis (Soudy et al., 2020). ”GO.db” (v.3.11.4) was used for simplification of Gene Ontology terms (Carlson, 2019).

Chapter 3

Proteolytic processing of PCDH19

3.1 Introduction

As described in (**Section 1.3**), many members of the cadherin superfamily have a signalling function, which is initiated by their proteolytic processing at the membrane. Proteolytic processing of cadherins often occurs in a two-step process whereby initial shedding of the extracellular fragment triggers the subsequent release of the intracellular domain. Release of the cytoplasmic domain leads to activation of downstream signalling pathways, indirectly by concurrent release of cadherin binding proteins such as β-catenin and p120 into the cytoplasm (Uemura et al., 2006a), or directly by the cytoplasmic fragment itself (Haas et al., 2005).

Despite several studies on proteolytic processing of members of all subfamilies of the cadherin superfamily (N-Cadherin (Reiss et al., 2005), (Uemura et al., 2006b), E-Cadherin and α, γ and δ protocadherins (Reiss et al., 2006), (Bonn et al., 2007), (Bouillot et al., 2011)), the possible proteolytic processing of PCDH19 has not yet been investigated. Only one study has demonstrated a nuclear function of PCDH19, which depends on the interaction between the paraspeckle protein NONO and PCDH19 (Pham et al., 2017). In that study, it was shown that the NONO-PCDH19 interaction leads to increased transcription of ER-α dependent genes.

Although nuclear localisation of PCDH19 was shown by Pham *et al.*, no processing mechanism was described. Proteolytic processing of (proto)cadherins is a highly conserved mechanism with recurring key proteases involved. For example, both N-Cadherin and PCDH12 are proteolytically cleaved in a two-step process in which ADAM10 releases the extracellular fragment and subsequently PS1/γ-secretase releases the intracellular fragment. This process directly alters cell-cell adhesion via removal of the extracellular domain, but also influences signalling via release of the cytoplasmic domain (Reiss et al., 2005) (Uemura et al., 2006a) (Bouillot et al., 2011).

Given the similarities highlighted between members of the cadherin superfamily and the recent report of a novel nuclear function of PCDH19 (Pham et al., 2017), it was decided to investigate if PCDH19 can be subjected to proteolytic processing and if its cytoplasmic domain has a signalling function. The role of both ADAM10 and PS1/γ-secretase in the potential cleavage of PCDH19 was also analysed, due to their involvement in the processing of other cadherins.

3.1.1 Aims

The aim of this chapter is to investigate the possibility that PCDH19, like other described (proto)cadherins, is subjected to proteolytic cleavage at the plasma membrane and, if so, to identify the protease(s) involved. This chapter also begins to investigate the functional significance of the generated fragments by determining their subcellular localisation. In summary this chapter will focus on the following questions:

- Can PCDH19, or part of it, be found in the nucleus?
- Is PCDH19 subjected to proteolytic processing and what fragments are generated?
- If PCDH19 is processed, which are the key proteases involved in the cleavage?

These questions are answered using *in vitro* treatment of mouse embryonic fibroblasts or mouse embryonic stem cell (mESC) derived neurons with different inhibitors and observing

PCDH19 fragments by western blot. As the processing mechanism for N-Cadherin has been extensively studied, N-Cadherin was used in several experiments throughout this chapter as positive control for the inhibitor treatments.

3.2 Results

3.2.1 PCDH19 nuclear localisation sequences

Nuclear localisation signals (NLSs) are recognised by nuclear transport proteins of the nuclear import pathway that carry cargo proteins into the nucleus. NLSs can be classified as monopartite (class I or class II) or bipartite, based on having either one or two clusters of basic amino acids (K/R). The basic sequence of monopartite NLSs is K(K/R)X(K/R), where X represents any amino acid. The basic structure of bipartite NLSs is (K/R)(K/R)X10-12(K/R)3/5, with X representing any amino acid and 3 out of 5 last amino acids being either K or R. The amino acid sequence of PCDH19 was inputted into the online NLS prediction tools cNLS mapper (Kosugi et al., 2009) and NLStradamus (Nguyen Ba et al., 2009) which predicted that the intracellular domain of PCDH19 contains a bipartite nuclear localisation signal (NLS) KRIAEYSYGHQKKSSKKKK (**Figure 3.3a**). PCDH19 ICD also contains two monopartite NL sequences, PTLKGKR and PGVKRLK (**Figure 3.3a**), which have been reported previously together with evidence of presence of PCDH19 in the nucleus (Pham et al., 2017). In that study a PCDH19 double-tagged construct was generated: PCDH19 was tagged with an N-terminal MYC-tag and a C-terminal FLAG-tag. It is important to note that tagging a transmembrane protein at the N-terminus, right before the signal peptide, could interfere with its correct translocation within the secretory pathway, hence subcellular localisation studies would have to be carefully interpreted.

3.2.2 Nuclear accumulation of PCDH19 cytoplasmic domain

Subcellular localisation of full length PCDH19, or of its cytoplasmic domain alone, was determined by transfection of C-terminal HA tagged PCDH19-HA or PCDH19-CYTO-HA, in HEK293T cells and mESC-derived neurons, plasmids described in methods (**Table 2.2**). mbEGFP plasmid was used as a reporter for transfection and to visualise the outline of the cells, which is especially useful when working with neuronal cells. Immunohistochemistry was used to detect

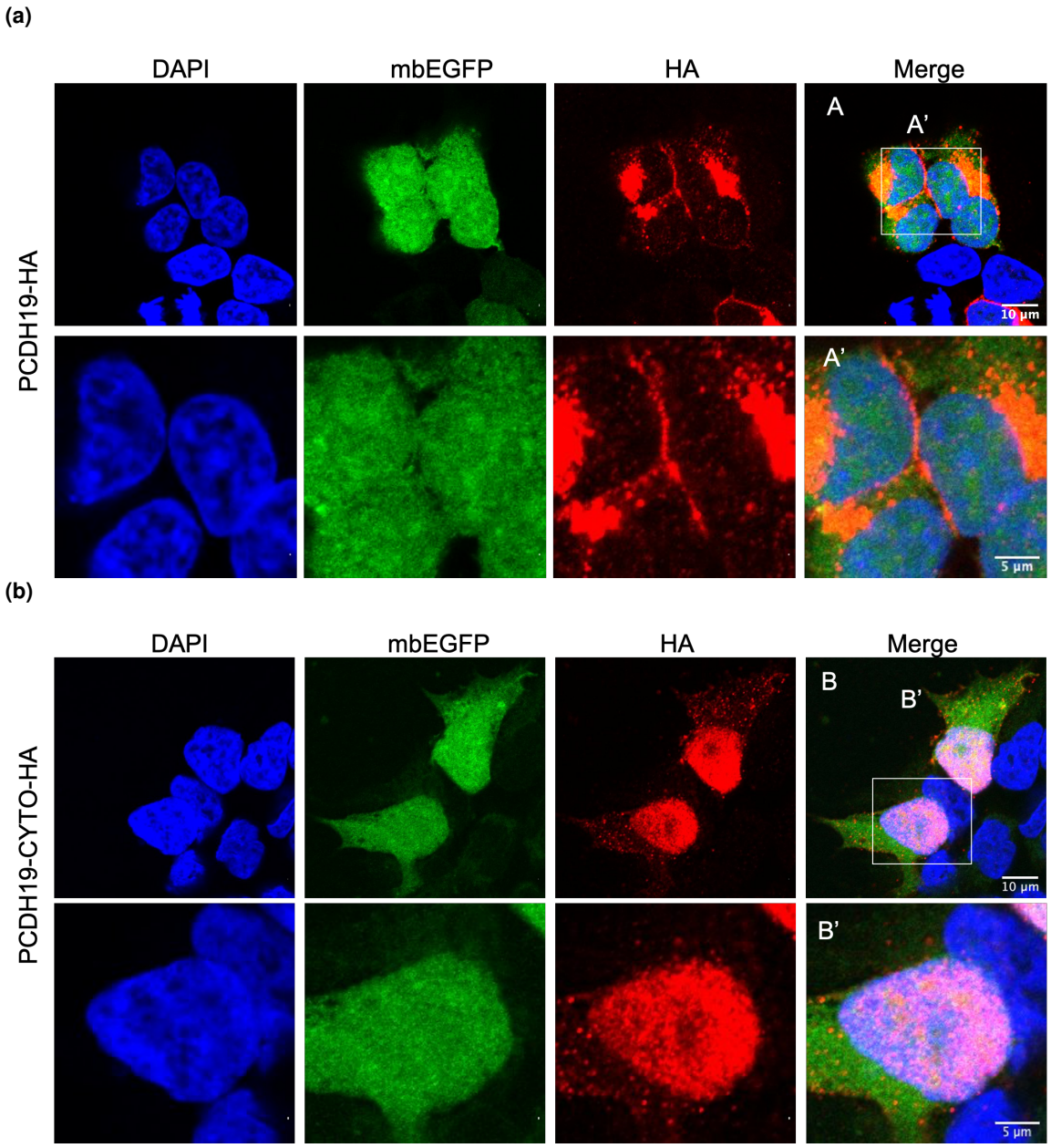


Figure 3.1: PCDH19 cytoplasmic domain localises to the nucleus in HEK293 cells. (a) Full-length PCDH19-HA or **(b)** PCDH19-CYTO-HA transfected HEK293 cells in combination with a mbEGFP reporter for transfection. Anti-HA staining (red) shows full-length protein at the membrane and in the cytoplasm **(a)** whilst PCDH19-CYTO is mainly seen in the nucleus **(b)**. Nuclei are stained with DAPI (blue). Scale bars: A-B, 10 μm; A'-B', 5 μm.

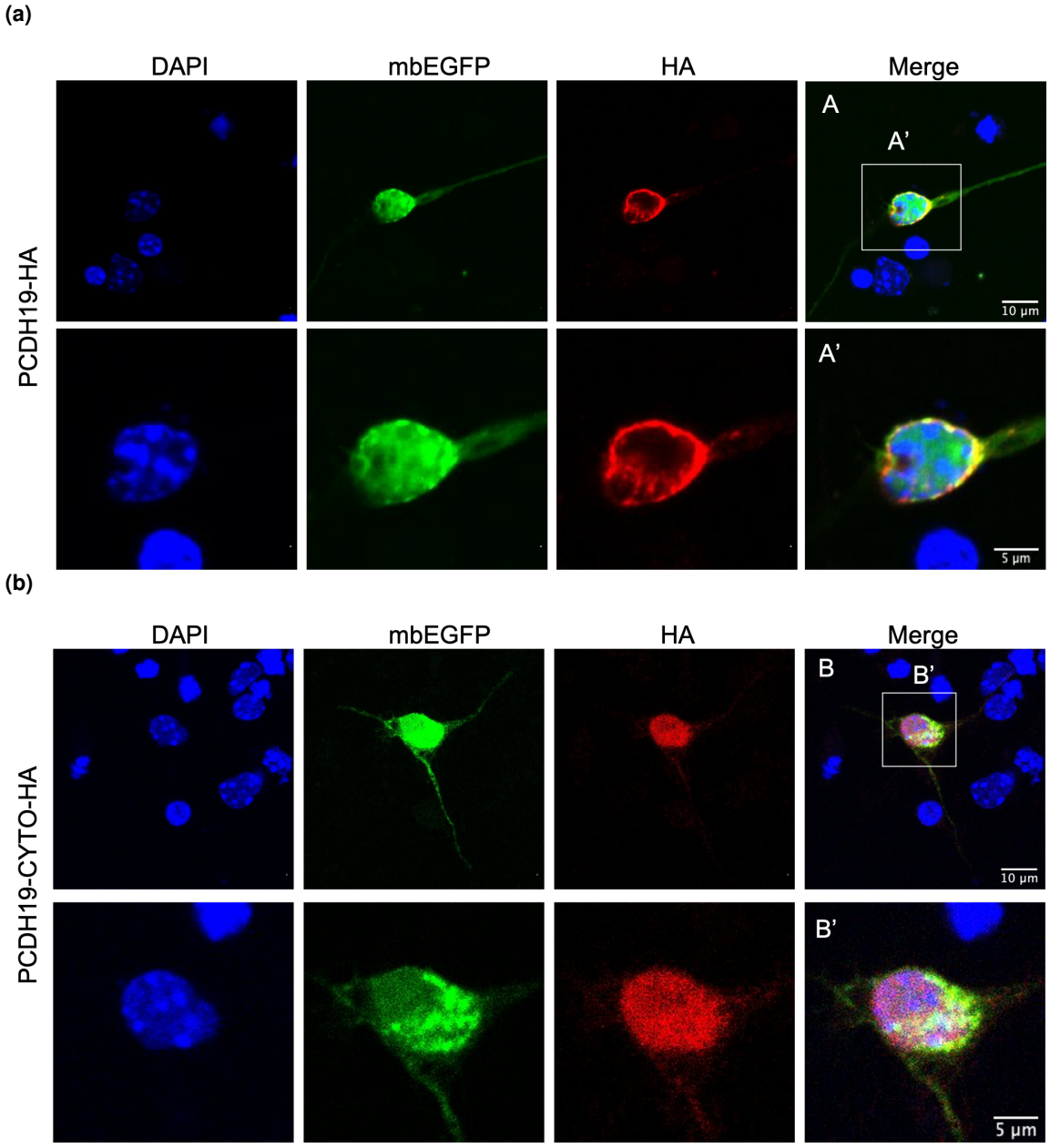


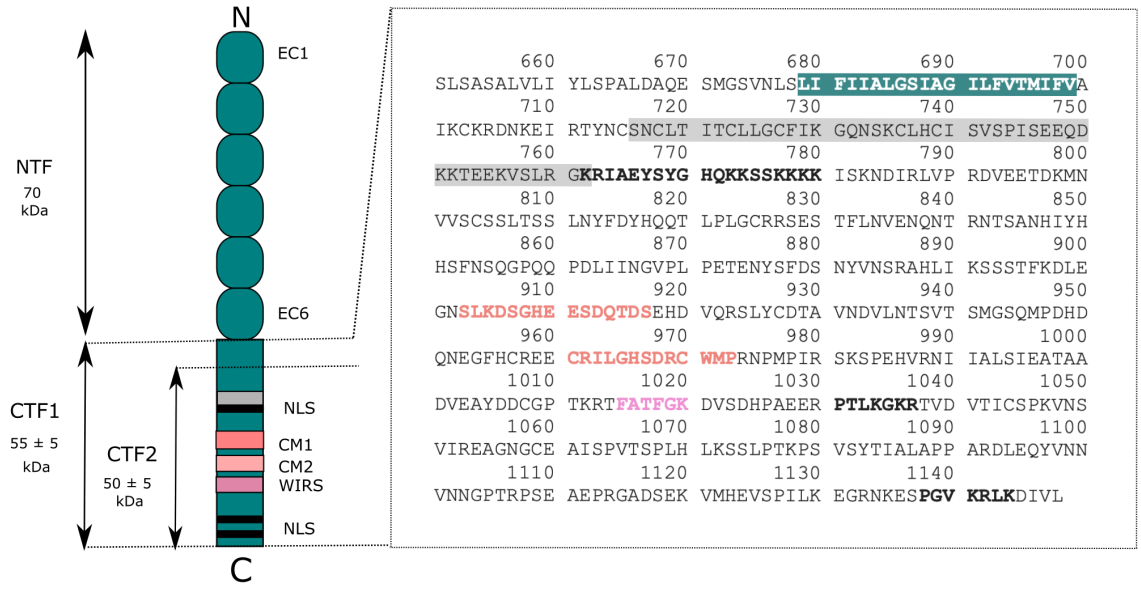
Figure 3.2: PCDH19 cytoplasmic domain localises to the nucleus in mESC-derived neurons. **(a)** Full-length PCDH19-HA or **(b)** PCDH19-CYTO-HA transfected mESC-derived neurons in combination with a mbEGFP reporter for transfection. Anti-HA staining (red) shows full-length protein at the membrane and in the cytoplasm **(a)** whilst PCDH19-CYTO is mainly seen in the nucleus **(b)**. Nuclei are stained with DAPI (blue). Scale bars: A-B, 10 μm; A'-B', 5 μm.

PCDH19 via the HA-tag. Overexpression of the full-lenght PCDH19-HA led to accumulation of the overexpressed protein in the perinuclear region, possibly corresponding to the endoplasmic reticulum. Staining was also observed at the plasma membrane, as seen at the contact site between cells (**Figure 3.1a**). Interestingly, overexpression of the intracellular domain of PCDH19, PCDH19-CYTO, leads to its accumulation in the nucleus, in both HEK293 cells (**Figure 3.1b**) and in mESC-derived neurons (**Figure 3.2b**). This finding validates the predicted nuclear localisation sequences (**Figure 3.3a**) confirming a possible nuclear function for the cytoplasmic domain of PCDH19.

3.2.3 Predicting the size of PCDH19 cleavage fragments

In order to be trafficked into the nucleus, a membrane bound protein needs to be released from its attachment. Commonly, this is achieved by cleavage of the protein on the cytosolic side, as discussed in the introduction. Given basic structural similarities and the knowledge available of N-Cadherin processing, a similar mechanism was hypothesised for PCDH19. Sequence alignment of the two proteins was done in order to compare putative cleavage sites and predict the size of proteolytic fragments of PCDH19 (**Figure 3.3b**). N-Cadherin cleavage by ADAM10 happens in the extracellular domain, very close to the transmembrane domain, at amino-acids R^{714} and I^{715} (Uemura et al., 2006b) (**Figure 3.3b**), whereas the second cut by PS1/γ-secretase happens just outside the transmembrane domain, at the membrane-cytosol interface (ϵ-cleavage), at amino-acids K^{747} and R^{748} (Uemura et al., 2006b). Although, interestingly, the KR residues are conserved in PCDH19, there is no consensus sequence for PS1/γ-secretase cleavage. If processing of PCDH19 happened similarly to N-Cadherin, two fragments would be generated. A first C-terminal fragment PCDH19-CTF1, containing the transmembrane domain of approximately 55 kDa, and a second, smaller C-terminal fragment without the transmembrane domain, PCDH19-CTF2, of approximately 50 kDa. The difference in size is calculated based on the size of the transmembrane domain, which is about 5 kDa. Depending on the inclusion or exclusion of the alternatively spliced exon 2 in CTF1 and CTF2, the fragments could also vary by an additional 5 kDa

(a)



(b)

Figure 3.3: Functional domains and putative cleavage sites of PCDH19. **(a)** Hypothesised proteolytic processing of PCDH19 generates CTF1, which includes the TM, and CTF2, without TM. Amino acid sequence of mouse PCDH19 C' terminal portion. Highlighted in green is the TM domain (679-699); highlighted in grey are amino acids coded by alternatively spliced exon 2 (716-762). Amino acids in bold pink are CM1, CM2 and the WIRS domain. Amino acids in bold black are predicted nuclear localisation sequences (NLS). **(b)** Alignment of PCDH19 and N-Cadherin. Transmembrane domain is highlighted in green and known cleavage sites for N-Cadherin are highlighted in grey.

(**Figure 3.3a**). Although PCDH19 contains a GxxxG motif in its transmembrane domain, which is common to 25% of γ-secretase substrates, cleavage site prediction cannot be determined by sequence analysis alone. This is because PS1/γ-secretase cleavage is not only dependent on primary structure of the substrate protein but also on its secondary structure - for instance, helix destabilisation is one of the mechanisms that triggers processing (Ben-Shem et al., 2007), (Beel and Sanders, 2008). Only a few substrates have known point mutations that abolish processing (ErbB4, Notch1, N-Cadherin) (Beel and Sanders, 2008). Moreover, the GxxxG motif might serve for homodimerisation of substrate, which in turn could affect processing (Beel and Sanders, 2008).

3.2.4 PCDH19 processing generates NTF, CTF1 and CTF2

The A304-468A PCDH19 antibody recognises an antigen present at the end of the C-terminus of PCDH19. For this reason, it is a valuable tool to investigate the production of intracellular PCDH19 fragments which contain the C-terminal end of the protein (**Figure 3.4a**). Previous western blot analysis conducted in the Martinez-Garay lab with this antibody on cell lysates from mouse embryonic fibroblasts (MEFs) showed the appearance of several bands in addition to the expected 120 kDa band corresponding to the full-length protein. In order to characterise these bands, membrane fraction enrichment was used as a method to determine subcellular localisation of the small cytoplasmic fragments of PCDH19. This fractionation method is relatively crude and is used to generate two fractions, a membrane enriched fraction and one containing everything else, which includes both cytoplasm and nucleus (described in methods (**Section 2.9**)). The method was validated by blotting against pan-cadherin and β-actin, seeing enrichment in the membrane and in the ”cytoplasmic” fraction, respectively (**Figure 3.4b**). The pan-cadherin antibody recognises the conserved cytoplasmic cadherin tail, common to classical cadherins.

After validation, the membrane enrichment method was used to determine subcellular localisation of the PCDH19 C-terminal fragments. Together with PCDH19 FL, a C-terminal fragment of about 55 kDa was found to be enriched in the membrane fraction. This band could correspond to

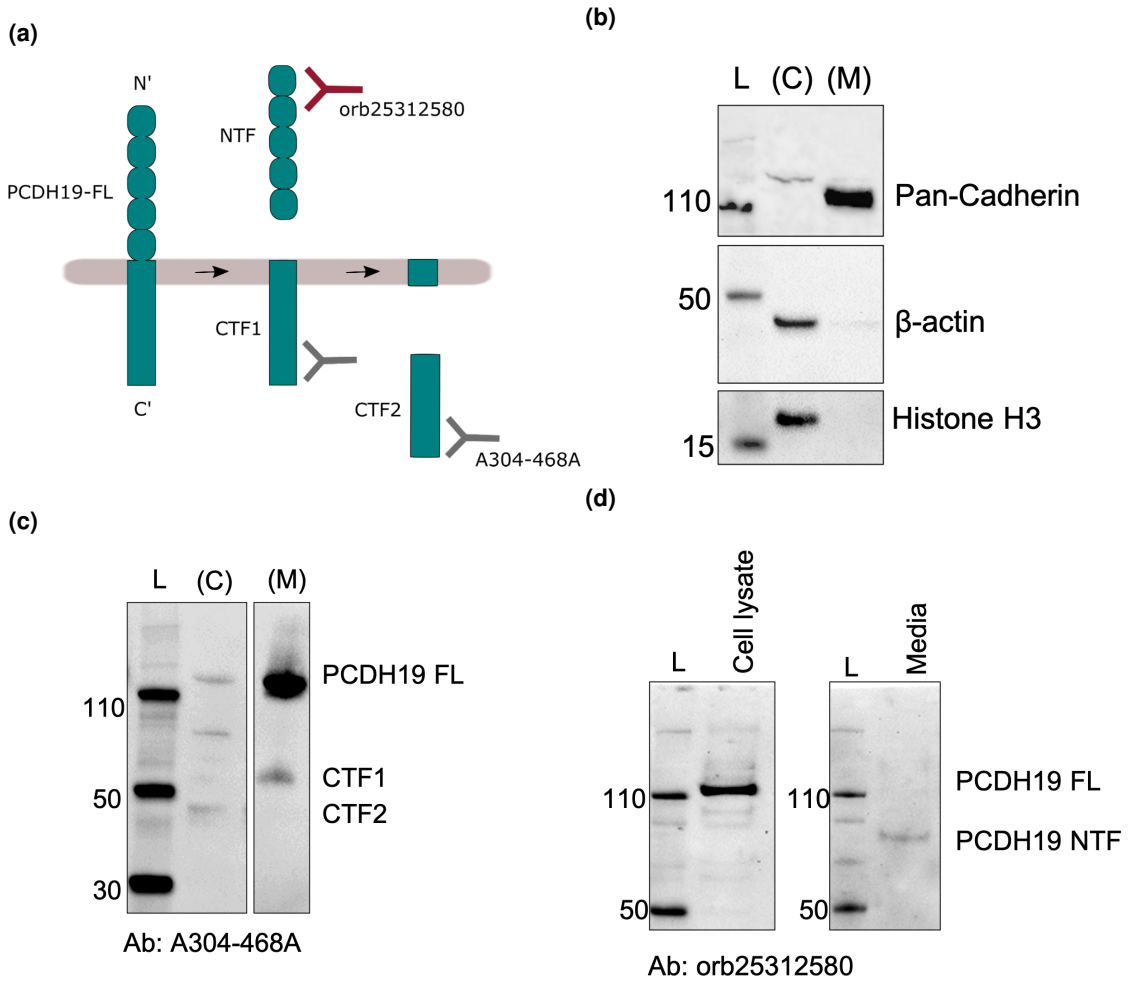


Figure 3.4: Detection of PCDH19-CTF1, CTF2 and NTF in different cellular fractions. **(a)** Schematic of detection of PCDH19-CTFs and NTF with different C-terminal and N-terminal antibodies. **(b)** Validation of membrane fraction enrichment on MEF lysates. Membrane bound cadherins are enriched in the membrane fraction (M) as shown by the pan-cadherin band and cytoplasmic β-actin is enriched in the cytoplasmic fraction (C). The fractionation technique used does not distinguish cytoplasmic and nuclear proteins as seen by presence of Histone H3 in the cytoplasmic fraction. **(c)** PCDH19-CTF1 and PCDH19-CTF2 enrichment in membrane and cytosolic fractions, respectively. **(d)** Detection of PCDH19-NTF in cell media.

PCDH19 CTF1, as CTF1 is predicted to contain the transmembrane domain and therefore should still be anchored at the plasma membrane. A second, smaller band below 50 kDa, which was visibly enriched in the cytoplasmic fraction, would then correspond to PCDH19 CTF2, reflecting the potential loss of membrane attachment (**Figure 3.4c**). The observed sizes of the PCDH19 cytoplasmic fragments are considerably close to the predicted sizes.

If the ectodomain of PCDH19 is being shed, there should be a detectable N-terminal fragment being released to the extracellular space. For cells in culture, this would be the medium, which is easily collectable. Medium of WT MEFs was changed 1 hour prior to collection, and substituted with medium containing no FBS, to avoid western blot saturation with serum proteins. The collected medium was then concentrated by centrifugation and run on a western blot. Using a PCDH19 N-terminal antibody (orb25312580) (**Figure 3.4a**), a fragment of about 70 kDa in size was detected in the cell medium (**Figure 3.4d**). The combination of the molecular weights of the detected PCDH19-NTF and CTF1 fragments, of about 70 kDa and 55 kDa respectively, approximately reconstitutes the molecular weight of the full length protein, predicted to be 126 kDa.

3.2.5 ADAM10 cleaves PCDH19 in mouse embryonic fibroblasts

ADAM10 is the main sheddase for N-Cadherin, as determined using ADAM10 KO mouse embryonic fibroblasts (MEFs) and specific inhibitors (Reiss et al., 2005), (Uemura et al., 2006b). Since N-Cadherin proteolytic processing has been well characterised, this protein was chosen as a positive readout for treatments, to validate the method and to strengthen conclusion on PCDH19 processing. ADAM10 KO and WT MEFs, kindly gifted to us by Prof. De Strooper at KU Leuven in Belgium, were initially used to replicate published data on N-Cadherin processing and to establish a reliable method in the lab. N-Cadherin processing can be induced by changes in intracellular calcium (Marambaud et al., 2003). Ionomycin has often been used to increase intracellular calcium, as it leads to an increase in calcium influx through the plasma membrane and from intracellular calcium stores, thus mimicking a generalised ”activation” and triggering

multiple downstream signalling pathways. Ionomycin has been used also to trigger processing of N-Cadherin.

ADAM10 KO MEFs and ADAM10 WT MEFS were treated with 5 μM ionomycin for 10, 30 or 60 min, in order to stimulate shedding, in the presence or absence of the ADAM10 specific inhibitor GI254023X (GI25X). Lysates were run on western blots and probed against N-Cadherin. As previously observed (Reiss et al., 2005), treatment with ionomycin stimulates production of a 37 kDa C' terminal fragment (CTF). Shedding is rapid and the strongest effect is seen after just 10 min (**Figure 3.5a**). As the CTF is produced, a parallel reduction in N-cadherin full-length (FL) is clearly visible (**Figure 3.5a**). In the presence of the ADAM10 inhibitor, the processing fragment is completely undetectable in WT MEFs. In the case of ADAM10 KO MEFs (**Figure 3.5b**), no processing of N-Cadherin is detected at all in the presence of ionomycin, thus strengthening and validating the previous results that ADAM10 is necessary for ectodomain shedding of N-Cadherin.

The same lysate samples were then tested for PCDH19. Similarly to N-Cadherin, ionomycin treatment seemed to trigger proteolytic processing of PCDH19 in the ADAM10 WT cells, as demonstrated by increased intensity of the CTF1 band, while treatment with GI254023X significantly reduced generation of CTF1 (**Figure 3.6**). The effects of MEF genotype (ADAM19 WT vs ADAM10 KO) and treatment were later confirmed to be significant by a two-way ANOVA analysis of the quantified CTF1s ((**Figure 3.7a**); Genotype: F (1, 32) = 26.74, $P < 0.0001$; Treatment: F (7, 32) = 11.49, $P < 0.0001$; two-way-ANOVA). A Dunnett's multiple comparisons test showed further significant differences in ADAM10 WT MEFs (Control *vs* IM 10', $P < 0.0001$; Control *vs* IM 30', $P = 0.0031$). It is important to note that although there is a significant difference in PCDH19 processing between ADAM10 KO and WT cells, the same fragments, CTF1 and CTF2, can be detected at much lower levels in ADAM10 KO cells, suggesting that PCDH19 can be promiscuously cleaved by other proteases.

ADAM10 KO cells were also blotted next to WT cells to verify the absence of the protein.

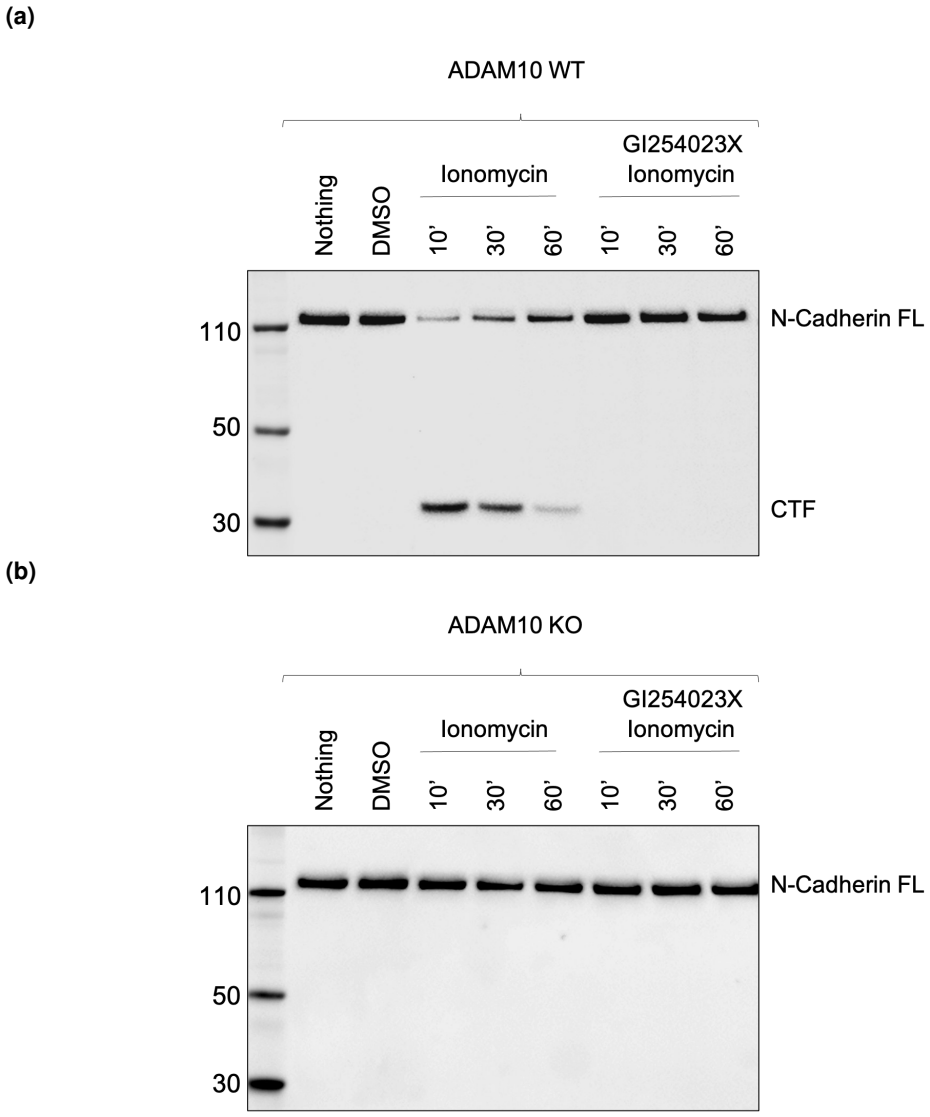


Figure 3.5: Stimulated shedding of N-Cadherin by ADAM10. (a) ADAM10 WT or **(b)** ADAM10 KO mouse embryonic fibroblasts were treated with nothing, DMSO, or 5 μM ionomycin for 10, 30 or 60 min in the presence or absence of GI254023X, an ADAM10 specific inhibitor. Treatment with ionomycin stimulates shedding of N-Cadherin, with the generation of a CTF of 37kDa in the WT but not in the ADAM10 KO MEFs. Shedding is blocked in the presence of the ADAM10 inhibitor.

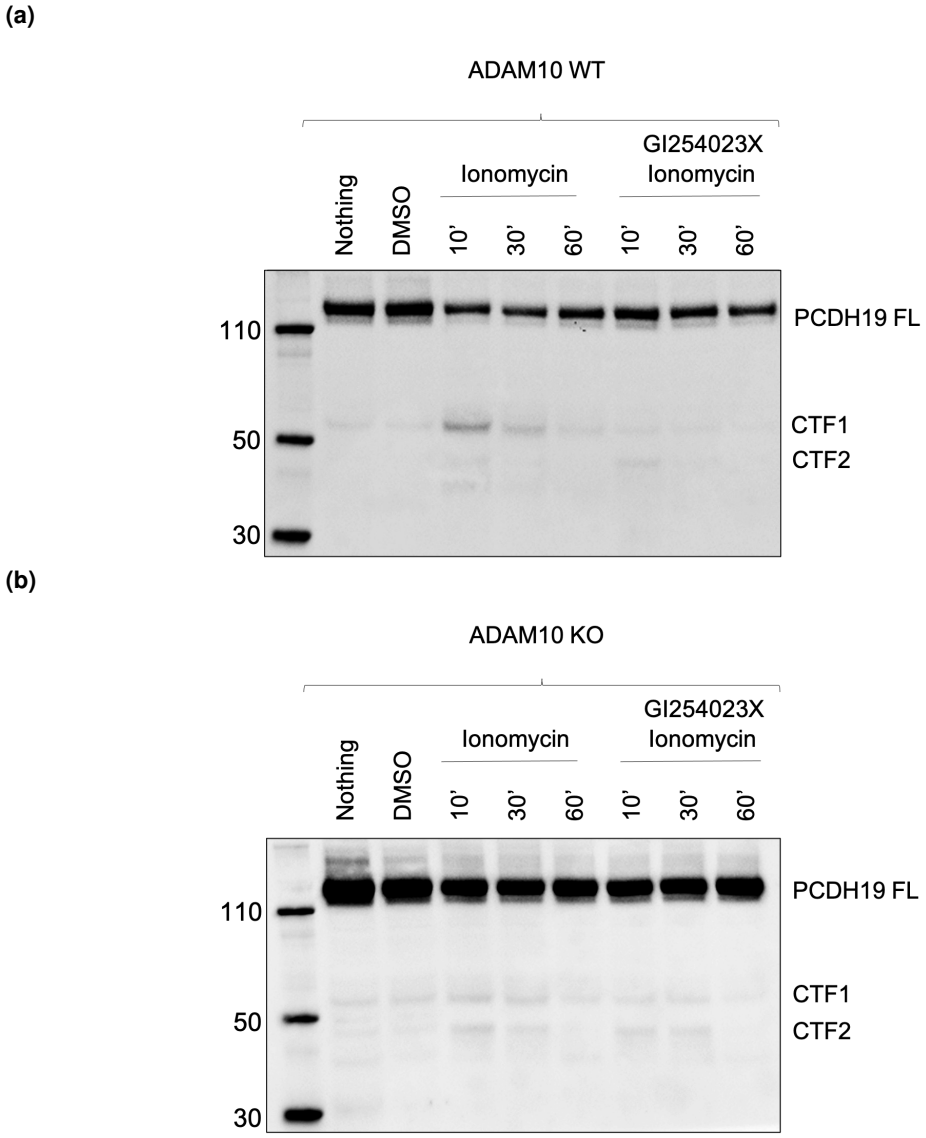
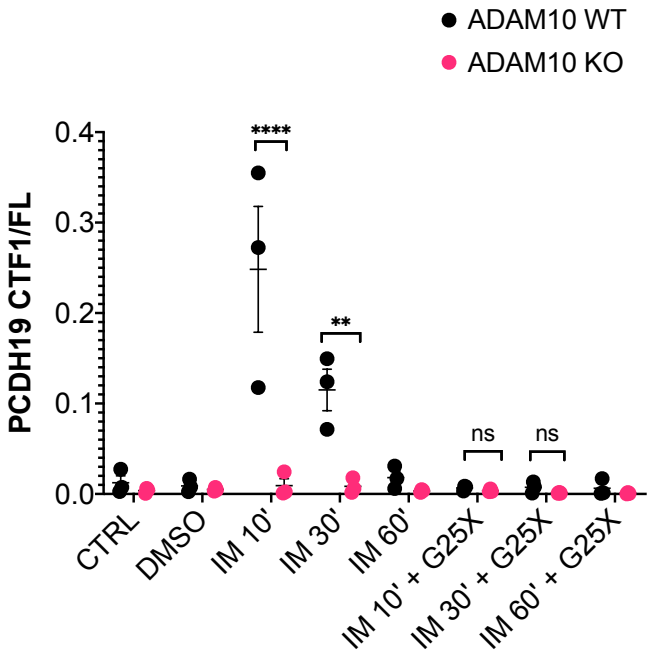


Figure 3.6: Stimulated shedding of PCDH19 by ADAM10. (a) ADAM10 WT or **(b)** ADAM10 KO mouse embryonic fibroblasts were treated with nothing, DMSO, or 5 μM ionomycin for 10, 30 and 60 min, in the presence or absence of GI254023X, an ADAM10 specific inhibitor. Treatment with ionomycin stimulates shedding of PCDH19, with the generation of CTF1 (55 kDa) in the WT cells. The shedding is blocked in the presence of the ADAM10 inhibitor.

(a)



(b)

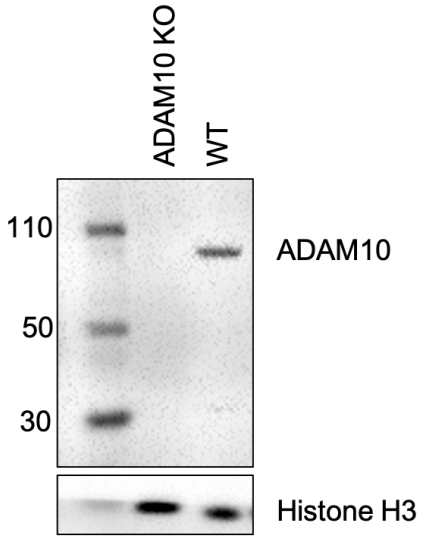


Figure 3.7: PCDH19 CTF1 production in WT MEFs. (a) Densitometric quantification of PCDH19 CTF1/FL in response to ionomycin treatment in ADAM10 WT (black dots) and ADAM10 KO MEFs (pink dots). (****) indicates $P < 0.0001$; (**) indicates $P < 0.01$, (ns) indicates non significant comparison. **(b)** ADAM10 is detected via western blot in WT MEFs but not in ADAM10 KO cells.

ADAM10 could be detected in WT but not KO MEFS, confirming absence of ADAM10 in KO cells (**Figure 3.7b**). Note that the visible band corresponds to ADAM10 precursor, which can be cleaved to generate the mature form (Brummer et al., 2018), as after cell lysis ADAM10 undergoes rapid auto-proteolysis, and in the absence of ADAM10 inhibitors becomes quickly undetectable by western blot (Brummer et al., 2018).

3.2.6 Shedding by ADAM10 in mESC-derived neurons

PCDH19-epilepsy is a neurological disorder. Therefore, it is necessary to transition from mouse embryonic fibroblast to a more pertinent model, so we chose stem-cell-derived neurons, which have been shown to express PCDH19 in previous work carried out in the Martinez-Garay lab. Mouse embryonic stem cells were differentiated into a homogeneous population of cortical-like neurons following the protocol (Bibel et al., 2004) described in (**Section 2.4.1**).

Having determined from the fibroblasts experiments that 10 min was the most effective timing for ionomycin treatment (**Figure 3.6**), this condition was also used to stimulate processing in mESC-derived neurons. As previously described, N-Cadherin extracellular fragment is released by ADAM10, which triggers a subsequent proteolytic event resulting in the generation of a CTF. Initially, we replicated existing evidence that ADAM10 is the neuronal N-Cadherin sheddase. Pre-treatment of neurons with the specific ADAM10 inhibitor GI254023X, as previously done in mouse-embryonic fibroblast, inhibits the generation of N-Cadherin-CTF, elicited by ionomycin (one-way ANOVA: F (2, 6) = 11.02; P = 0.0098. Tukey's multiple comparisons test: CTRL *vs* IM, P = 0.0111; CTRL *vs* IM + G25X, P = 0.7673; IM *vs* IM + G25X, P = 0.0245) (**Figure 3.8a**), and the ADAM10 inhibitor effectively reverts the effect of ionomycin.

Having validated the inhibitor treatments in mESC-derived neurons, the same neuronal lysates were blotted also for PCDH19. Ionomycin led to the accumation of CTF1 and CTF2, but whilst CTF1 is barely detectable, CTF2 appears to be the main PCDH19 cytoplasmic fragment in neu-

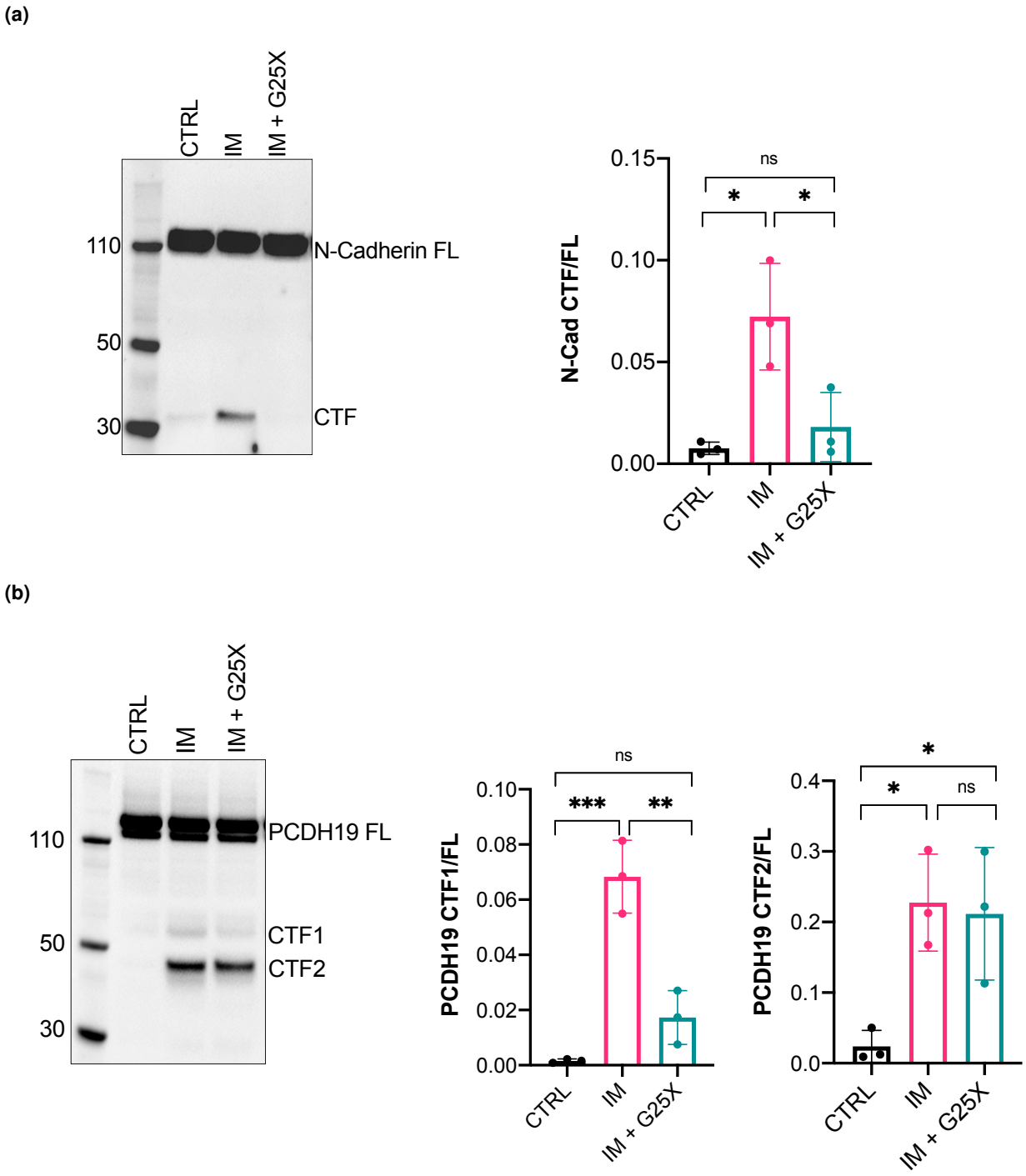


Figure 3.8: Stimulated shedding of N-Cadherin and PCDH19 by ADAM10 in mESC-derived neurons. DIV8 mESC-derived neurons were treated with DMSO (CTRL), or 5 μM ionomycin for 10 minutes (IM), or with 5 μM ionomycin after pre-treatment with 10 μM GI254023X, an ADAM10 specific inhibitor (IM + G25X). Neuronal lysates were blotted against **(a)** N-Cadherin or **(b)** PCDH19, and relative intensity of generated fragments in relation to full-length protein was quantified by densitometric analysis. (***) indicates $P < 0.001$; (**) indicates $P < 0.01$; (*) indicates $P < 0.05$; (ns) indicates non significant comparison.

rons. Nonetheless, after GI254023X, a reduction in the generation of CTF1 was detected (one-way ANOVA: $F(2, 6) = 40.63$; $P = 0.0003$. Tukey's multiple comparisons test: CTRL *vs* IM, $P = 0.0003$; CTRL *vs* IM + G25X, $P = 0.1847$; IM *vs* IM + G25X, $P = 0.0014$) (**Figure 3.8b**) consistent with the findings in MEFs. Interestingly, the treatment with GI254023X had no effect on blocking the generation of PCDH19-CTF2 by ionomycin (one-way ANOVA: F (2, 6) = 8.226; $P = 0.0191$. Tukey's multiple comparisons test: CTRL *vs* IM, $P = 0.0250$; CTRL *vs* IM + G25X, $P = 0.0350$; IM vs IM + G25X, $P = 0.9567$) (**Figure 3.8b**). So whilst ADAM10 can cleave PCDH19, this does not seem to be necessary for the downstream production of PCDH19-CTF2 in mESC-derived neurons.

3.2.7 PS1/γ-secretase cleaves PCDH19

PS1/γ-secretase is known to cleave N-Cadherin and many other substrates at the membrane. In the case of N-Cadherin, PS1 cleavage generates the second fragment, free from membrane attachment, N-Cadherin-CTF2 (Uemura et al., 2006b). PS1/γ-secretase involvement in PCDH19 processing was tested in mESC-derived neurons using a selective inhibitor for the complex, DAPT (**Figure 3.9**).

As demonstrated by the previous results, changes in intracellular calcium in MEFs and mESC-derived neurons, triggered by ionomycin treatment, lead to generation of PCDH19-CTF1 and CTF2, respectively. For this experiment, in order to artificially mimic neuronal activity and increase intracellular calcium, KCl was used. DIV8 mESC-derived neurons were pre-treated with the PS1/γ-secretase inhibitor DAPT for 3 hours, followed by 15 minute treatment with KCl. Whilst triggering processing by KCl led to rapid accumulation of PCDH19-CTF2, pre-treatment of neurons with 10 μM DAPT, led to a significantly reduced generation of PCDH19 CTF2 (one-way ANOVA: F (2, 9) = 11.14, $P = 0.0037$; Tukey's multiple comparisons test: CTRL *vs* KCL, $P = 0.0042$; CTRL *vs* KCL + DAPT, $P = 0.7180$; KCL *vs* KCL + DAPT, $P = 0.0136$) (**Figure 3.9a**).

A parallel experiment was conducted to see the effects of DAPT in the absence of KCl. Neurons

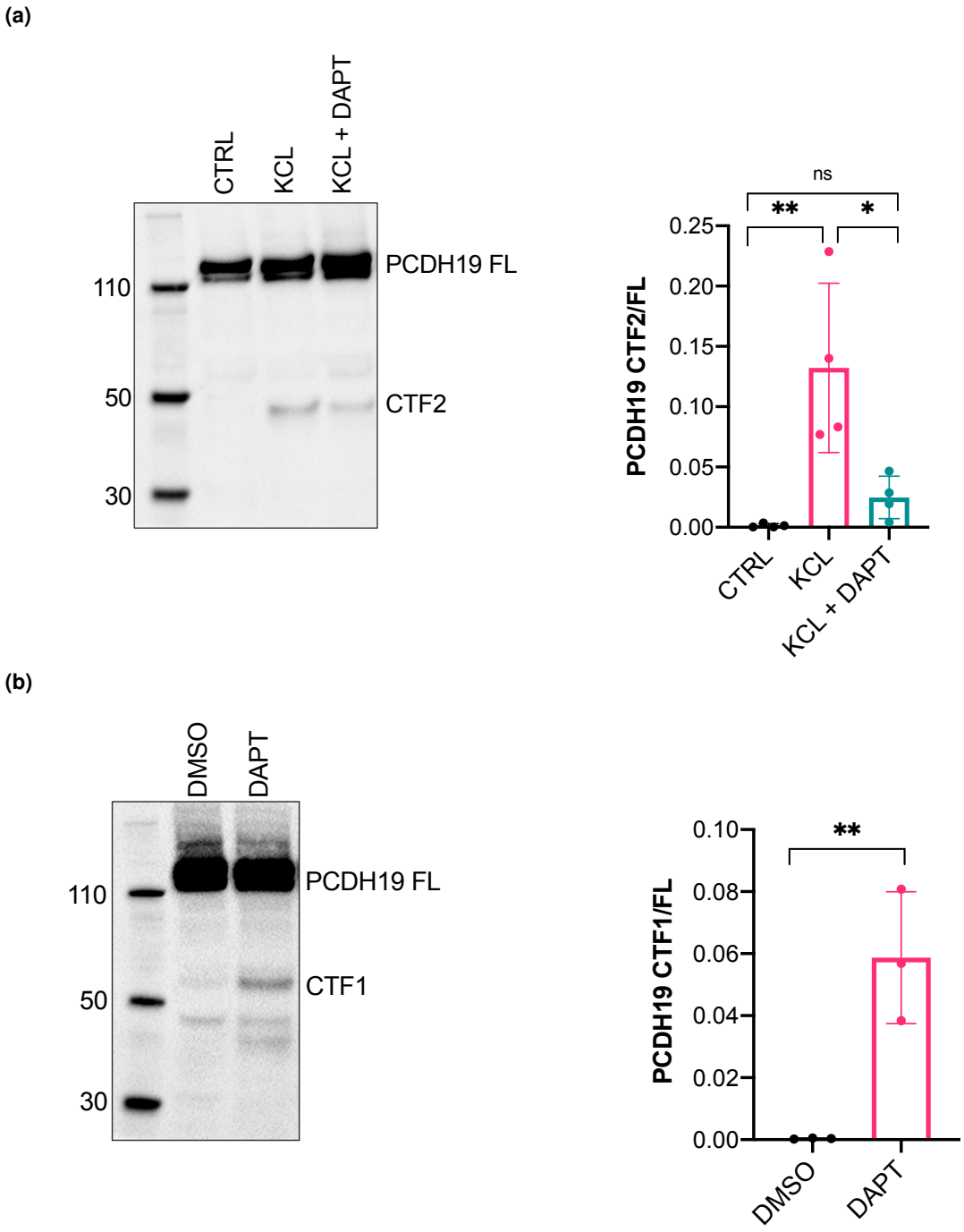


Figure 3.9: PCDH19 can be processed by PS1/γ-secretase. (a) CTF2 intensity is reduced in the presence of DAPT. DIV8 mESC-derived neurons were treated with or without 10 μM DAPT, PS1/γ-secretase inhibitor for 3 hours. 15' before lysis, neurons were activated with 50 mM KCl. KCl triggers production of CTF2 but processing is inhibited in the presence of DAPT. **(b)** DAPT treatment leads to accumulation of CTF1. mESC-derived neurons were treated with 10 μM DAPT for 16 hours. DAPT treatment results in accumulation of PCDH19 CTF1, substrate for PS1/γ-secretase. (**) indicates $P < 0.01$; (*) indicates $P < 0.05$; (ns) indicates non significant comparison.

were treated with DAPT or DMSO for 16 hours, and then blotted against PCDH19. DAPT treatment resulted in the accumulation of PCDH19-CTF1, which is the substrate for PS1/γ-secretase, even in the absence of induced calcium influx (unpaired t-test, DAPT *vs* DMSO, $P = 0.0090$) (**Figure 3.9b**). DAPT treatment for 16 hours also resulted in the accumulation of a third, smaller intracellular fragment suggesting that alternative proteases could be cleaving PCDH19, at a different site, in order to clear membrane-bound fragments. Taken together these data shows that PS1/γ-secretase can cleave PCDH19 CTF1 and generate CTF2.

3.2.8 Activity-dependent processing of PCDH19

Activity-dependent processing of cell-adhesion molecules found at synapses has been shown to play a key role in modulating synaptic plasticity (Conant et al., 2015). Many cell-adhesion molecules, including N-Cadherin, are cleaved at the synapse in response to neuronal activity and this serves as a mechanism to modulate synaptic function. For example, N-Cadherin cleavage at glutamatergic synapses affects size and maturation of dendritic spines (Malinverno et al., 2010).

As PCDH19 has been shown to partially localise with synaptic markers such as Homer and Synapsin1/2 (Hayashi et al., 2017), it was important to determine if levels of full-length PCDH19 could also be altered by activity. For this experiment more mature neurons were used, at DIV11. At this stage, mESC-derived neurons in culture have formed functioning synapses and are electrically active (Bibel et al., 2007). N-Methyl-D-aspartic acid (NMDA), the specific ligand of the NMDA receptor, was used to trigger neuronal activity in the mESC-derived neuronal cultures. 30 minute treatment with NMDA led to increased generation of PCDH19 CTF2, whilst pre-treatment with MK-801, a specific antagonist of the NMDA receptor, in combination with NMDA blocked generation of CTF2 (one-way ANOVA: F (2, 6) = 12.83, $P = 0.0068$; Tukey's multiple comparisons test: CTRL *vs* NMDA, $P = 0.0108$; CTRL *vs* NMDA + MK-801, $P = 0.9994$; NMDA *vs* NMDA + MK801, $P = 0.0112$) (**Figure 3.10**). Interestingly, when this experiment was carried out on less mature DIV8 mESC-derived neurons there was no detectable CTF2 triggered by treatment with NMDA (data

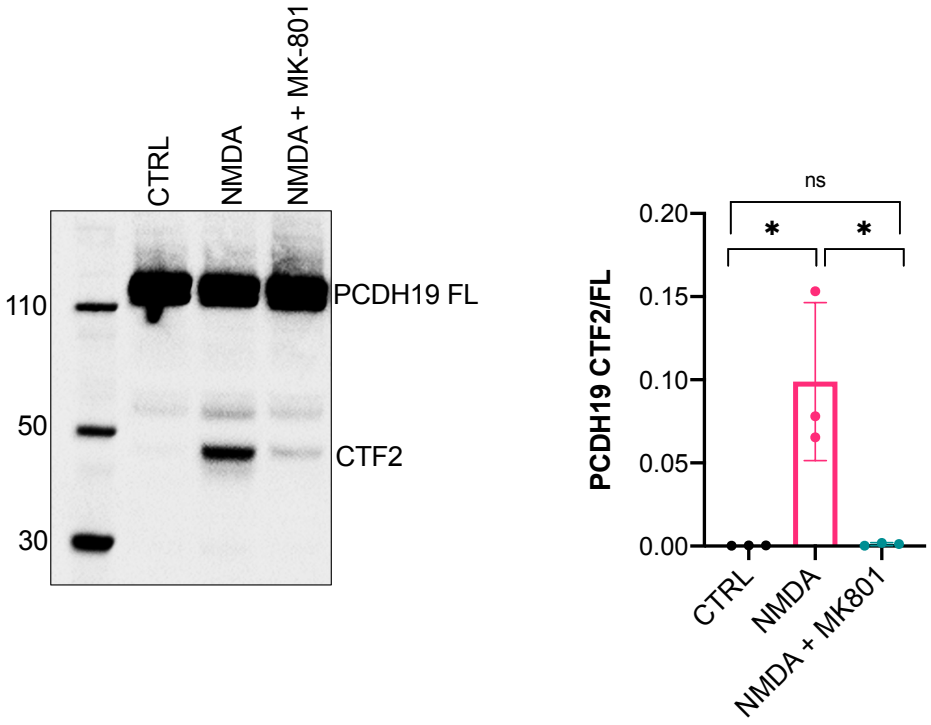


Figure 3.10: PCDH19 processing is triggered by activity of the NMDA Receptor. DIV11 mESC-derived neurons treated with 50 μM of NMDA or 50 μM + 1 μM MK-801, 30 min pre-lysis. FL: full-length; CTF2: C-terminal fragment 2.

not shown). Whilst GRIN1 is expressed at this timepoint, at DIV8 neurons are still not electrically active (Bibel et al., 2004), therefore although the machinery for PCDH19 processing is present at DIV8, processing via activation of NMDA receptor requires electrically mature neurons.

3.3 Discussion

This chapter discusses the proteolytic processing of PCDH19. Using two different cell-culture systems, MEFs and mESC-derived neurons, it was determined that, *in vitro*, PCDH19 can be processed into a soluble, extracellular N-terminal fragment (PCDH19-NTF) and at least two intracellular C-terminal fragments (PCDH19-CTF1 and PCDH19-CTF2). Slightly different than originally predicted, the sizes of the fragments are 55 kDa and about 45 kDa for PCDH19 CTF1 and CTF2, respectively, with 10 kDa difference between them, meaning cleavage sites of either protease might be slightly apart from the membrane interface.

Using inhibitors, an ADAM10 KO cell line and N-Cadherin as a positive control it was determined that PCDH19 ectodomain shedding can be mediated by ADAM10 in MEFs and in mESC-derived neurons *in vitro*, but interestingly, in neurons, ADAM10 inhibition has no effect on downstream generation of CTF2, suggesting the possible production of CTF2 directly from the full-length protein or the involvement of other sheddases. ADAM10 is expressed and active in mESC-derived cortical-like neurons, as seen by cleavage of N-Cadherin, and seems to be able to cleave PCDH19 to generate PCDH19-CTF1, but with no effect on CTF2. Therefore ADAM10 does not seem to be the main neuronal sheddase for PCDH19 and it remains necessary to investigate the involvement of other possible sheddases. The most efficient approach would involve using broad-spectrum inhibitors for whole families of proteases such as the matrix metalloproteases or other ADAMs. Initial investigations probing BACE-1 and MMPs in the cleavage of PCDH19 seem to imply that these proteases are not involved (data not shown), but they constitute very preliminary experiments so far. Processing *in vivo* can also be cell-type specific, for instance p75^{NTR} is differentially cleaved in hippocampal neurons and in cerebellar neurons, resulting in

the prevalence of different cytoplasmic fragments in each neuronal cell-type (Vicario et al., 2015).

Treatment with DAPT, a PS1/γ-secretase inhibitor, led to a reduction in the production of the cytoplasmic fragment CTF2 in mESC-derived neurons. In parallel, it was seen that DAPT treatment, in the absence of stimulation also led to the accumulation of CTF1, the substrate for CTF2 production, as downstream processing is inhibited. Combined, these data show that PS1/γ-secretase can indeed process PCDH19. It is also possible that, in some instances, PS1/γ-secretase could be cleaving PCDH19 without prior ectodomain shedding, generating CTF2 directly from full-length PCDH19. A similar mechanism has been reported for the neurotrophin receptor p75. In response to nerve growth factor binding, p75-ICD is directly released via PS1/γ-secretase and translocated to the nucleus, independently from the shedding of p75-ECD via metalloproteases (Frade, 2005). It is also possible that another protease, such as calpain, could be generating PCDH19-CTF2 in parallel, as the band does not entirely disappear in the presence of the γ-secretase inhibitor. N-Cadherin, for instance, in addition to being cleaved by γ-secretase can also be cleaved by calpain in the cytoplasmic domain (Jang et al., 2009).

Although some processing of PCDH19 is detectable in untreated cells at baseline, stimulation of cells with substances that increase intracellular calcium (such as ionomycin, KCl or NMDA) greatly increased the production of the cytoplasmic fragments. Interestingly, ionomycin treatment in MEFs leads to rapid accumulation of CTF1, whereas in neurons, it is CTF2 which is the most visible. The mechanism underlying this observation remains to be elucidated, but it is interesting to speculate that CTF2 might have a more relevant role in neurons as opposed to MEFs and therefore it is not immediately lost to proteasome degradation. Alternatively, this discrepancy could also be due to different expression levels of PS1/γ-secretase in MEFs and neurons, whereby CTF1 in neurons is immediately processed into CTF2 because of localised presence of PS1/γ-secretase.

This chapter also investigated the effect of neuronal activity on proteolytic processing of PCDH19. It was determined that PCDH19 undergoes activity-dependent processing in mature

mESC-derived neurons following treatment with NMDA. The effect of NMDA could be blocked by simultaneous use of the NMDA antagonist MK-801. NMDA receptor activation can trigger processing of N-Cadherin, EphA4, (Inoue et al., 2009), Notch (Alberi et al., 2011) and other synaptic proteins regulating the development and function of excitatory synapses. Interestingly, a study found that *Pcdh19* mRNA is downregulated in response to neuronal activity. Benito *et al.* triggered activation of neurons *in vitro* by expression of constitutively-active CREB (cAMP Response Element-Binding Protein) and constitutively active cFOS in hippocampal-derived neurons and detected downregulation of *Pcdh19* by RNAseq, amongst many other genes (Benito et al., 2011). So whilst neuronal activity stimulates processing of PCDH19, which alters PCDH19-mediated adhesion, it can at the same time downregulate transcription of *Pcdh19*.

Adding to the complexity, combinatorial expression of α- and γ-pcdhs has been shown to affect their susceptibility to processing, meaning that heteromers of α- and γ-pcdhs are processed differently to homomer formations (Bonn et al., 2007).Therefore, depending on their cell-specific expression, processing of different pcdhs is altered leading to differential downstream signalling. As mentioned, PCDH19 is know to form a complex with N-Cadherin. The formation of the PCDH19/N-Cadherin complex in neurons could therefore be affecting the susceptibility to processing of PCDH19, as well as N-Cadherin, but this has yet to be investigated.

This chapter also confirmed that, when overexpressed alone, the cytoplasmic domain of PCDH19 can be found in the nucleus of both HEK293 cells and mESC-derived neurons but it remains to be confirmed that, *in vivo*, the PCDH19 cytoplasmic fragment released after processing, CTF2, can translocate to the nucleus. Another experiment to investigate subcellular localisation of PCDH19 could involve mutational analysis done by altering nuclear localisation signals and seeing if distribution of protein changes. This has been done successfully for other ICDs, such as for NICD (Huenniger et al., 2010), NRG1-ICD (Fazzari et al., 2014) and ErbB4-ICD (Williams et al., 2004). A further experiment that could resolve this would involve fusion of PCDH19 to a Gal4 transactivation domain that can induce luciferase reporter activity. The construct could

be transfected together with a reporter plasmid and cells treated with ionomycin in order to induce processing. Reporter activity would indicate cleavage and nuclear translocation, as further discussed in (**Section 7.6**).

Many cadherins are processed and the released fragments have biological activity but for PCDH19-CTF1/2 it remains to be elucidated if these fragments retain biological activity or if they are subjected to proteosomal degradation. The fragments are most visible after 10 min of ionomycin treatment and almost undetectable one hour after treatment meaning that the signalling function of the fragment is short lived.

3.3.1 Conclusion

This chapter demonstrates the previously unknown processing of PCDH19. It was determined that PCDH19 can be processed generating smaller intracellular fragments. Although the neuronal role of ADAM10 in PCDH19 processing remains to be investigated, it has been determined that PS1/γ-secretase can generate PCDH19-CTF2 and that, in ESC-derived neurons *in vitro*, processing is triggered by activation of the NMDA receptor. These observations highlight a novel role for PCDH19 in mediating signalling in neurons. The following chapters aim to resolve if PCDH19-CTF2 can have a signalling function: (**Chapter 5**) describes the generation of PCDH19-ICD overexpressing and PCDH19-KO ESCs as an *in vitro* model to investigate the role of the intracellular domain of PCDH19 and (**Chapter 6**) describes the results of an RNA-sequencing analysis carried out with those PCDH19-ICD overexpressing and PCDH19-KO cortical-like progenitors and neurons.

Chapter 4

PCDH19 interactors

4.1 Introduction

Although currently not much is known about PCDH19 interacting proteins, increasing evidence demonstrates a role for PCDH19 in diverse cellular processes, ranging from cell-cell interactions and reorganisation of the cytoskeleton to nuclear signalling, as reviewed in (Gerosa et al., 2018).

So far, PCDH19 has been shown to be involved in cell-adhesion in three ways: by binding other PCDH19 molecules on neighbouring cells *in trans* via EC1-4 (Cooper et al., 2016); by forming a *cis* complex with N-Cadherin, which has stronger adhesion properties than PCDH19 alone (Emond et al., 2011), (Biswas et al., 2010); and by *cis* interaction with other δ-Pcdhs, PCDH10 and PCDH17 (Pederick et al., 2018).

As mentioned in the introduction (**Figure 1.2**), PCDH19 has a WAVE regulatory complex (WRC) interacting receptor sequence (WIRS) in its cytoplasmic domain, via which it can interact with the WRC. The WAVE regulatory complex consists of 5 proteins: CYFIP1/2, NAP1, ABI1/2/3, HSPC300 and WAVE1/2 /3, also known as WASF1/2/3. It has been demonstrated that CYFIP and ABI form a composite surface that is recognised by the WIRS sequence, generating a

binding interface which exists only when the WRC is assembled and functioning. Membrane proteins such as PCDH19, which contain a WIRS sequence, can recruit the WRC to the plasma membrane in order to modulate cytoskeletal rearrangements (Chen et al., 2014). In fact, the interaction between PCDH19 and different WRC components has been experimentally validated, in particular CYFIP2 and NAP1 have been shown to interact with PCDH19 (Tai et al., 2010).

PCDH19 can also bind DOCK7, a guanine nucleotide exchange factor (GEF) which regulates Rac activity (Emond et al., 2021) and seems to be important for interkinetic nuclear migration during corticogenesis (Yang et al., 2012). In the same study it was determined that PCDH19 interacts with NEDD1, a centrosomal protein (Emond et al., 2021). In the hippocampus, PCDH19 binds the GABA-A receptor α1, and is capable of modulating GABAergic transmission by altering levels of receptor found at the surface (Bassani et al., 2018). It is speculated that this could be mediated by endocytosis and that PCDH19 could be involved in vesicle trafficking. Finally, PCDH19 has been shown to have a nuclear role, via its interaction with the paraspeckle protein NONO, which is capable of modulating gene expression (Pham et al., 2017).

The observations summarized above are derived from independent studies in different organisms (mouse, rat, chicken, zebrafish) and *in vitro* models (HEK293 cells, primary cultures). No comprehensive neuronal PCDH19 interactome has been published so far, and there is also no study comparing interactors during development and in adulthood. Although conserved functions are highly probable, PCDH19 will be playing different and specific roles during cortical development and in adulthood. Uncovering new pathways in which PCDH19 is involved can also advance our current understanding of *PCDH19*-epilepsy.

4.1.1 Aims

The aim of this chapter is to investigate novel interactors of PCDH19 *in vivo*, in adult and in embryonic brain, by immunoprecipitation combined with mass spectrometry analysis, in order to

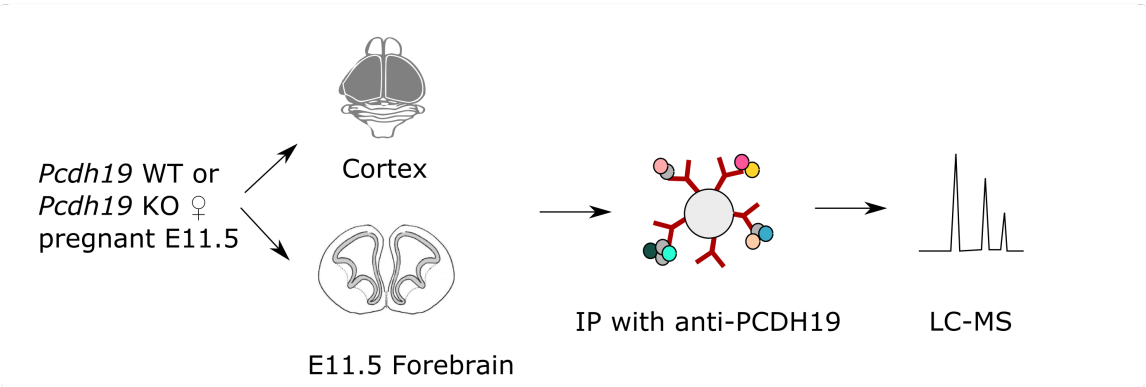
identify novel processes or pathways in which PCDH19 might be involved. In particular, relating to our previous findings on PCDH19 processing and nuclear translocation of overexpressed PCDH19 cytoplasmic domain, this chapter aims to uncover potential links to a nuclear function of PCDH19. Because of the expression pattern of PCDH19, adult cortical tissue and forebrain at the onset of neurogenesis at E11.5, when PCDH19 is highly expressed, were chosen as starting material.

4.2 Results

4.2.1 Experimental setup

In order to investigate PCDH19 interacting proteins in an unbiased manner, we performed liquid-chromatography mass-spectrometry (LC-MS) analysis on brain lysate samples after immunoprecipitation with the PCDH19 antibody. We compared cortex samples from pregnant animals and forebrain samples from their E11.5 litters from 3 *Pcdh19* KO and 3 WT animals and their respective E11.5 litters (**Figure 4.1a**). The *Pcdh19* KO litters originated from KO x KO matings, whilst the WT litters originated from WT x WT matings, to ensure that all embryos in the litter had the same genotype and the samples could be pooled to obtain enough starting material. Before LC-MS, all samples were checked via western blot to confirm PCDH19 expression and successful immunoprecipitation (**Figure 4.1b**). As expected, all WT samples showed a band around 120 kDa, corresponding to full-length PCDH19, that was absent in all KO samples. Also as expected, the PCDH19 band appeared enriched after immunoprecipitation when compared to the band from the input, which contains the total protein lysate. Despite high expression levels of PCDH19 in neuroprogenitor cells, immunoprecipitation was more successful for the adult cortex samples compared to the E11.5 forebrain samples, possibly due to the total amount of starting tissue. It is important to note that in the adult cortex samples there are two specific bands over 120 kDa corresponding to full-length PCDH19, whilst in the E11 tissue, there is only one. The two bands in the adult cortex probably represent glycosilated/unglycosilated forms or alternatively spliced isoforms or a combination of both.

(a)



(b)

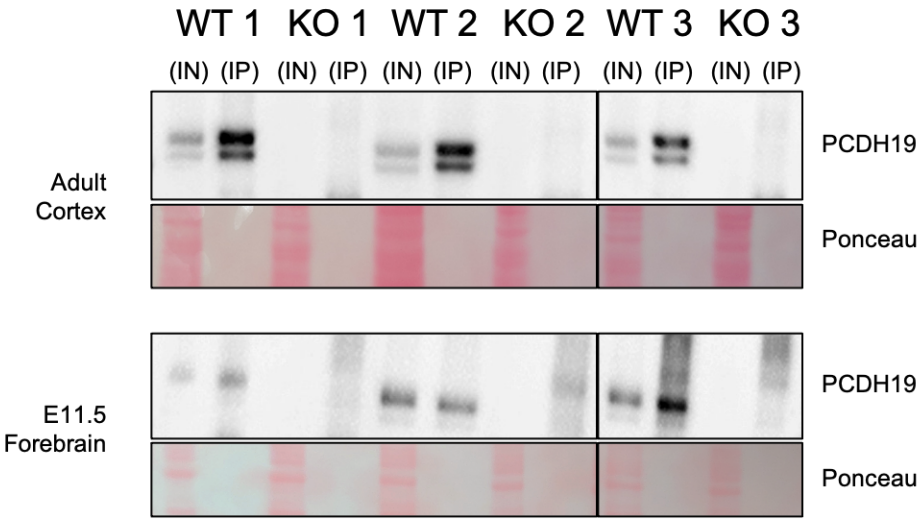


Figure 4.1: Immunoprecipitation of PCDH19 for LC-MS analysis. **(a)** Experimental design. **(b)** Western blot anti-PCDH19 to confirm immunoprecipitation of PCDH19 in WT samples and absence of PCDH19 in the KO samples. Ponceau stain is used as a loading control.

4.2.2 PCDH19 peptides

First of all, the presence of peptides corresponding to PCDH19 was checked in all samples. Consistent with the immunoprecipitation, PCDH19 was one of the most abundant peptides in the WT samples. However, a few PCDH19 peptides were unexpectedly detected in the KO adult cortex samples, although at much lower levels than in the WT samples (**Table 4.1**). An analysis of the distribution of the PCDH19 peptides along the PCDH19 protein (**Figure 4.2**) showed that although in cortex sample KO 1 peptides mapped along the whole protein, all PCDH19 peptides detected in cortex samples KO 2 and 3 mapped exclusively to the cytoplasmic domain of the protein (**Figure 4.2**). Dissections and immunoprecipitation were carried out on different days for KO and WT animals, therefore samples could not have been mixed or contaminated during those procedures. It was later determined that the first KO sample had been run on a gel together with the WT 1, 2 and 3 samples and could thus have been cross-contaminated in the process. Therefore, the KO 1 sample was eliminated from all downstream analysis. KO samples 2 and 3 were instead retained for the analysis based of the following rationale: as mentioned previously, (**Figure 1.4a**) (**Section 2.1.2**), in the *Pcdh19* KO mouse model used in this thesis, exons 1, 2 and 3 have been replaced by a β-gal/NeoR fusion cassette. This means that the last three exons of *Pcdh19* (exons 4, 5 and 6) are still present. Hence there is a possibility that in the *Pcdh19* KO mouse there might be residual expression of an alternative transcript encompassing exons 4, 5 and 6. Expression levels of this potential truncated form of PCDH19, would be very low, as this was undetectable by western blot using the C-terminal antibody (the uncropped version of the western blot is shown in the appendix (**Figure 8.2**)). Unfortunately, these results indicated that the KO samples are not true negative controls; in fact, if there is residual expression, the resulting protein would be entirely cytoplasmic and therefore might be enriching the KO samples for non-membrane bound interacting partners of PCDH19. Nonetheless, for the purpose of this work, downstream analysis was carried out as it might still provide interesting and useful information.

Table 4.1: PCDH19 peptides identified by LC-MS at 1% FDR

Adult Cortex	Unique	Total	Coverage (%)	E11.5 Forebrain	Unique	Total	Coverage (%)
WT 1	16	21	19.83	WT 1	24	32	29.17
WT 2	43	147	47.86	WT 2	17	24	20.42
WT 3	29	53	34.41	WT 3	4	4	4.56
KO 1	14	25	20	KO 1	0	0	0
KO 2	5	6	5.33	KO 2	0	0	0
KO 3	3	4	3.23	KO 3	0	0	0

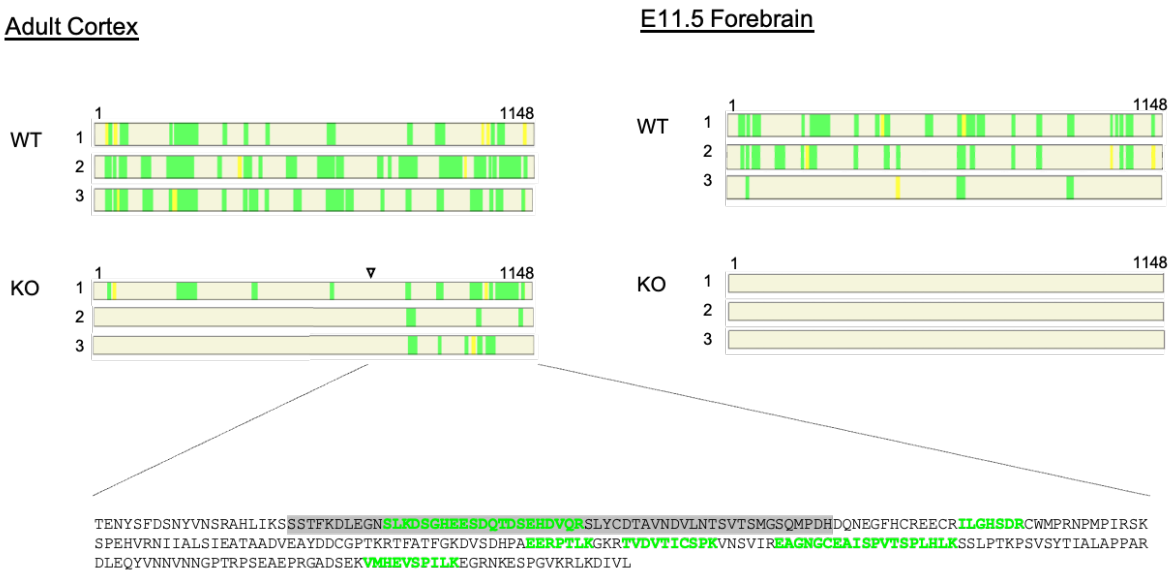


Figure 4.2: Distribution of detected peptides along the PCDH19 protein. Peptides mapping onto mouse PCDH19 identified by LC/MS. Arrowhead indicates end of exon 3. All peptides identified in the adult cortex KO 2,3 samples (green bold sequences) map onto the C-terminal region of the protein – shown amino acid sequence encoded by exon 4, 5 and 6 with alternated highlight. Green 1% FDR, Yellow 5% FDR.

4.2.3 PCDH19 interactors

LC-MS identified 3298 total proteins in the adult cortex samples and a total of 1767 proteins in the E11.5 forebrain samples. These lists were strictly filtered, by eliminating any protein that appeared in any of the KO samples of the same dataset. After filtering, 632 proteins appeared only in the WT samples for adult cortex . Of these 632, only 3 proteins appeared in 3/3 samples. For E11.5 forebrain, 361 proteins were only found in the WT samples after filtering. Of these 361 proteins, only 4 appeared in 3/3 litters. Interesting, only 29 proteins are detected at least once in both datasets. Those proteins are listed in the appendix (**Appendix 8.2.1**). Protein expression is very different in the two tissues analysed, but common interactors could give a clue of the core processes in which PCDH19 is involved.

Before proceeding with the analysis, the lists were scanned for known PCDH19 interactors, in order to validate our results. N-Cadherin (CDH2) was detected in the E11.5 dataset. ABI1 and ABI2, and WASF1, 2 and 3 were also detected in the E11.5 sample, as well as DOCK7. PCDH17 was instead detected in the cortex samples (**Table 4.2**).

Using the R package ”UniprotR” (v.2.0.3), and the ”GetSubcellular location” function, known subcellular distribution of identified proteins, was retrieved from the UniProt database (Soudy et al., 2020). Almost half of the proteins were classified as undefined, meaning that distribution is promiscuous or has yet to be experimentally determined (**Figure 4.3**). The vast majority of defined proteins are located in ”cytoplasm” or ”membrane”. Interestingly, there is a substantial proportion of nuclear proteins in the cortex samples that is much smaller in the E11.5 samples.

Table 4.2: Known PCDH19 interactors identified by LC-MS

Gene ID	Protein Name	Protein Function	Dataset	Reference
Cdh2	N-Cadherin	adhesion molecule	E	(Emond et al., 2011), (Biswas et al., 2010)
Pcdh17	Protocadherin-17	adhesion molecule	C	(Pederick et al., 2018)
Abl1-2	Abl interactor 1-2	WRC component	E	(Chen et al., 2014)
Wasf1-3	WASP family member 1-3	WRC component	E	(Chen et al., 2014), (Emond et al., 2021)
Dock7	Dedicator of cytokinesis protein 7	Rac GEF	E	(Emond et al., 2021)

(E) E11.5 dataset; (C) cortex dataset.

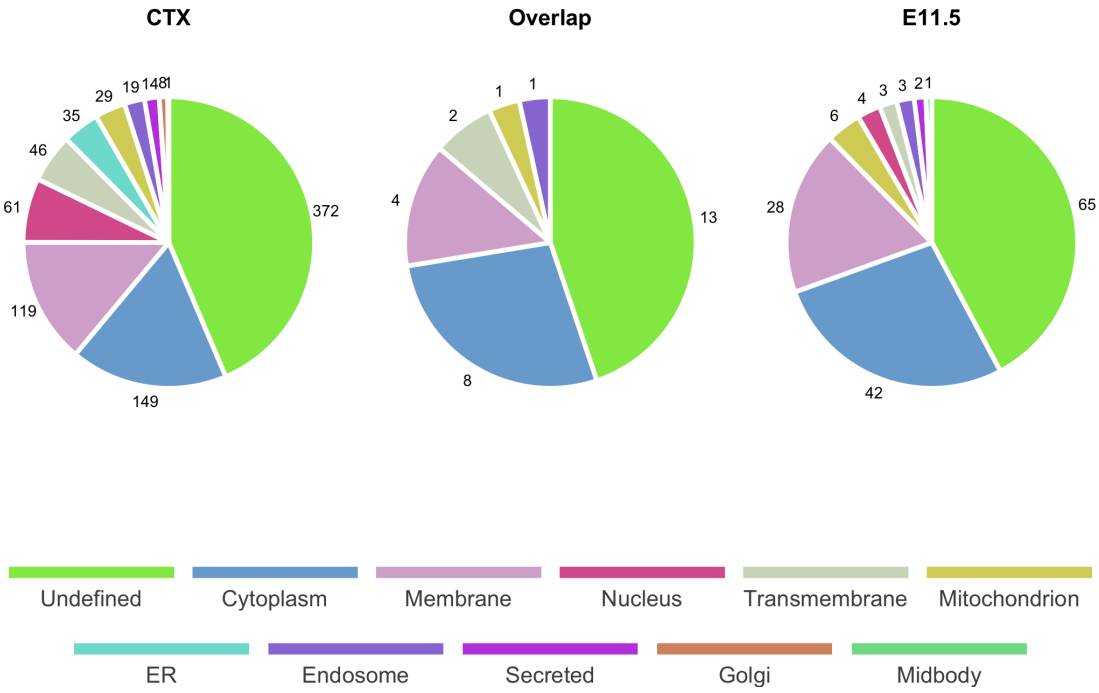


Figure 4.3: Subcellular distribution of detected proteins. Number of proteins detected in each subcellular compartment for adult cortex (CTX), E11.5 and the overlap between the two datasets.

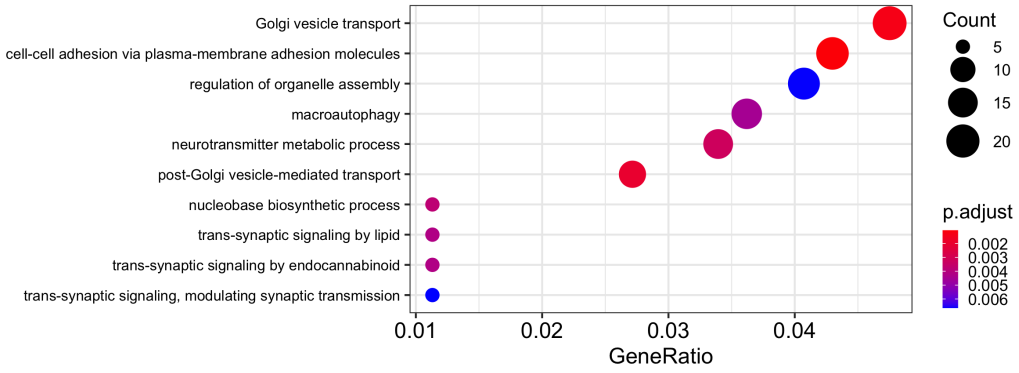
4.2.4 Gene ontology analysis

Gene ontology analysis was carried out on R (v.4.3) using the package "Cluster Profiler" (v.3.16.1) (Yu et al., 2012). Analysis of the two datasets (CTX, E11.5) was performed for Biological Process (BP), Molecular Function (MF), and Cellular Component , with a p-value cut off < 0.05. Results were visualised in dotplots and cnetplots.

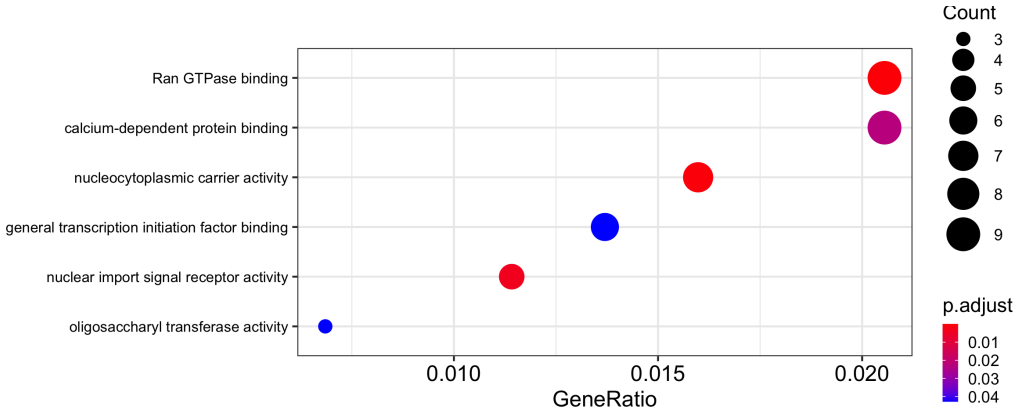
In the adult cortex dataset the analysis revealed several gene ontology terms which reflected previously known functions of PCDH19. In particular, for biological processes (GO:BP), the term "*cell-cell adhesion via plasma membrane cell adhesion molecules"* highlights the known function of PCDH19 at the cell membrane in mediating contacts with neighbouring cells. Interestingly, the terms *"Golgi vesicle transport"*, *"neurotransmitter metabolic process"* and "*trans-synaptic signalling, modulating synaptic transmission*" were also significant, suggesting a role for PCDH19 at the synapse, possibly in mediating or participating in vesicle transport (**Figure 4.4a**). For molecular function (GO:MF) *"calcium dependent protein binding"* was found amongst the significant terms, in line with known function of protocadherins. Interestingly, the terms "*RAN GTPase binding*", "*nucleocytoplasmic carrier activity*" and "*nuclear import signal receptor activity*" were also significant (**Figure 4.4b**). Those terms suggest the possibility of nuclear transport of PCDH19 via interaction with different transporters. Some of these nuclear transporters are shown in (**Figure 4.5a**). "*General transcription initiation factor binding*" is also enriched, suggesting a potential nuclear function of PCDH19. Finally, for cellular compartment (GO:CC), some of the most significant terms related to the synapse, including *"integral component of synaptic membrane"* and *"GABA-ergic synapse"*. Several terms were also associated with vesicle trafficking, in particular *"synaptic vesicle'* and "*transport vesicle membrane*" (**Figure 4.4c**). Network depictions of the interacting proteins found in the analysis indicating their relationship to selected GO terms for biological process, molecular function and cellular compartment are shown in (**Figure 4.5**).

In the E11.5 dataset, most of the biological process (GO:BP) terms relate to actin dynamics,

(a) BP



(b) MF



(c) CC

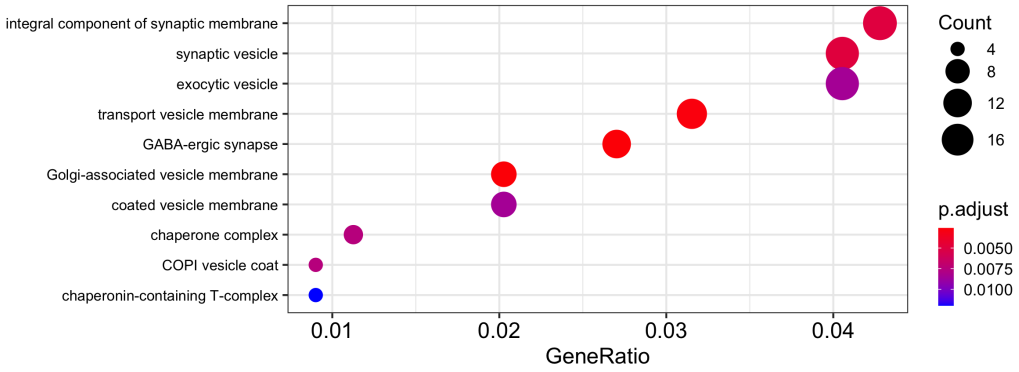


Figure 4.4: Dotplot of GO enrichment analysis for adult cortex. (a) Biological Process (BP). **(b)** Molecular Function (MF). **(c)** Cellular Compartment (CC). ”Count” indicates the number of genes in the dataset associated with a specific GO term and is represented by the size of the dot. ”Gene ratio” is the ratio between the number of genes associated with a specific GO term present in the dataset, over the total number of genes in the dataset. ”p.adjust” value is indicated by the colour of the dots with bright red dots corresponding to most significantly enriched GO terms.

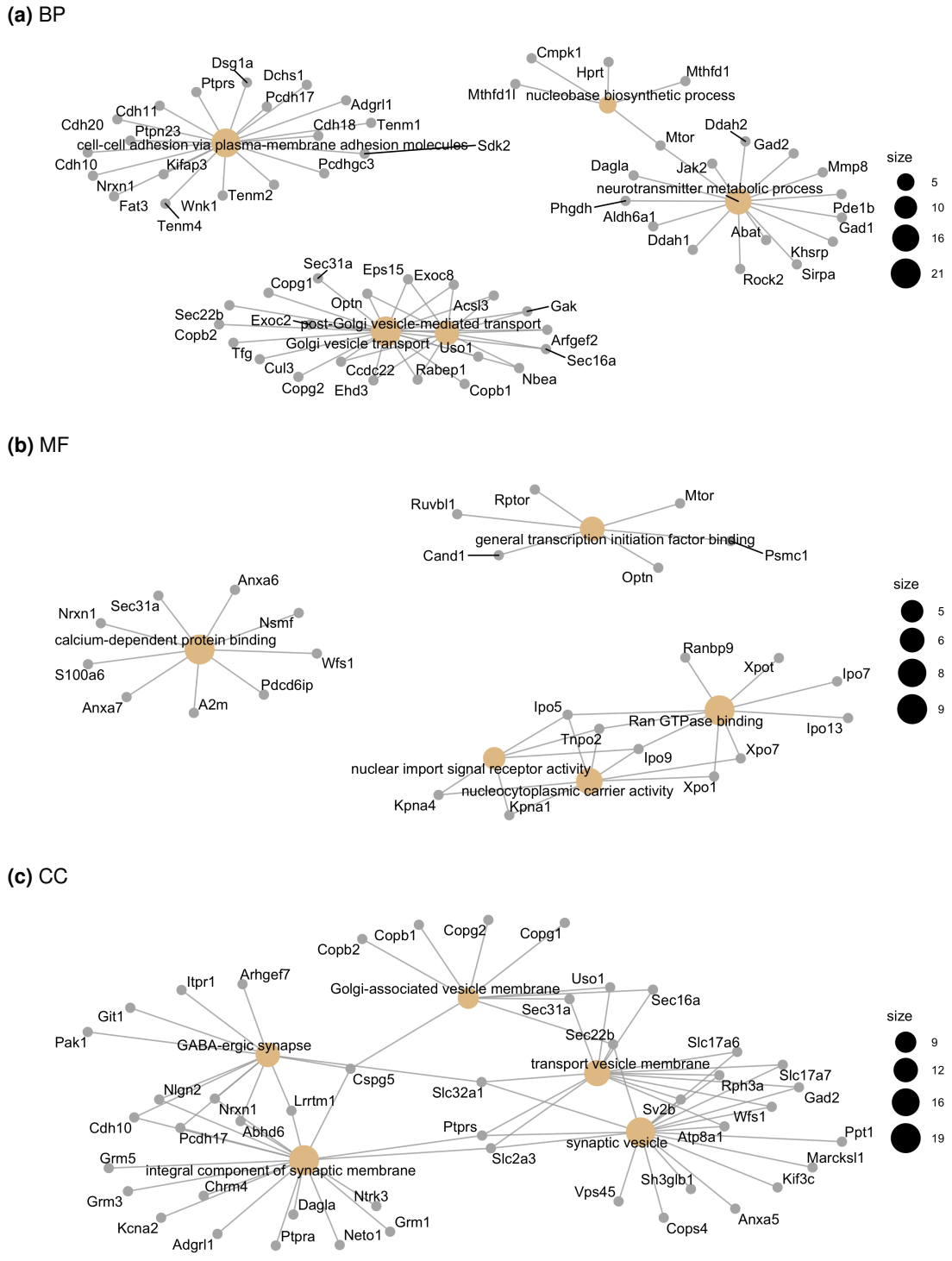


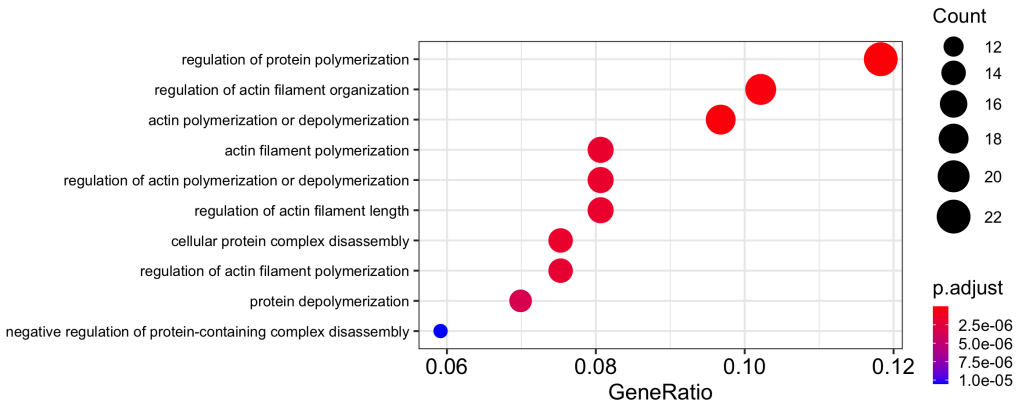
Figure 4.5: Cnetplot of GO enrichment analysis for adult cortex . (a) Biological Process. **(b)** Molecular Function. **(c)** Cellular Compartment. Size of the dot corresponds to number of genes associated with the specific GO term.

such as *"actin polymerization and depolymerization"*, *"regulation of actin filament organization"* and *"regulation of actin filament length"* (**Figure 4.6a**), in line with previous reports of PCDH19 interacting with the WRC and modulating actin dynamics (Chen et al., 2014), (Emond et al., 2021). The significant molecular function terms (GO:MF) include *"microtubule binding"* and *"actin binding"*, as well as *"SH3 domain binding"* and *"PDZ domain binding"* (**Figure 4.6b**), highlighting a role for PCDH19 in cytoplasmic signalling and signal transduction. Amongst the interactors associated with "*PDZ domain binding*" we find, for example, DLG1, a multidomain scaffolding protein, necessary for recruiting and anchoring receptors and ion channels to the plasma membrane. The significant cellular compartment terms (GO:CC) include again terms relating to the actin cytoskeleton, such as *"actin-based cell-projection"*, "*actin filament*" and also several terms related to the microtubules. (**Figure 4.6c**). Interestingly, *"nuclear pore*" was also identified in this dataset, driven by the presence of several components of the nuclear pore (NUP155, NUP50, NUP153, NUP133), which are important for nucleocytoplasmatic trafficking and can interact with importins (Ogawa et al., 2012). Again, network depictions of several of the interacting proteins found in the analysis, indicating their relationship to different GO terms for biological process, molecular function and cellular compartment are shown in (**Figure 4.7**).

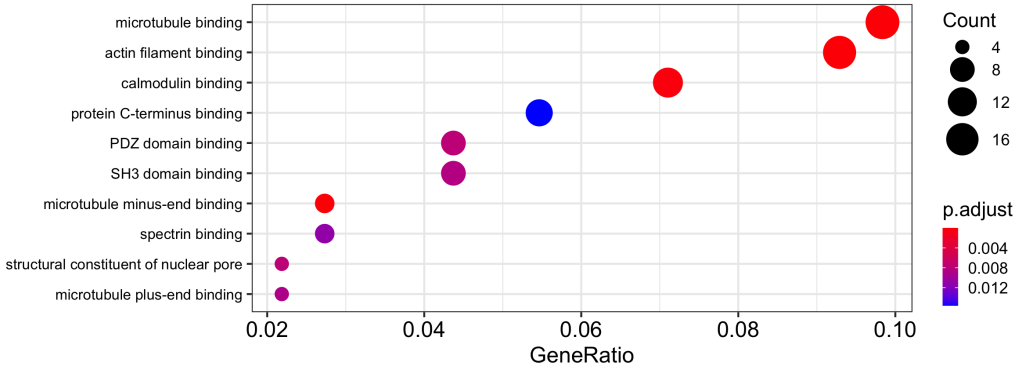
4.2.5 Interesting potential novel interactors of PCDH19

Several members of the cadherin superfamily were identified in this analysis. In addition to CDH2 and PCDH17 that have been described previously as PCDH19 interactors (Pederick et al., 2018) (Emond et al., 2011), PCDH1, PCDH7, PCDH9, PCDH10, PCDHα12, PCDHβ1,2,9, PCDHγa3,12, PCDHγb1,6, PCDHγc3,4, CDH10, CDH11, CDH18 and CDH20 were also identified. This is interesting because interactions between members of the same or a different cadherin subfamily can alter adhesion properties of the protein alone (Pederick et al., 2018) (Bisogni et al., 2018). Another RacGEFs was detected: DOCK3, belonging to the same protein family as the recently described interactor of PCDH19, DOCK7 (Emond et al., 2021). Both DOCK7 and DOCK3 are guanine nucleotide exchange factor (GEFs), which activate Rac and promote restructuring

(a) BP



(b) MF



(c) CC

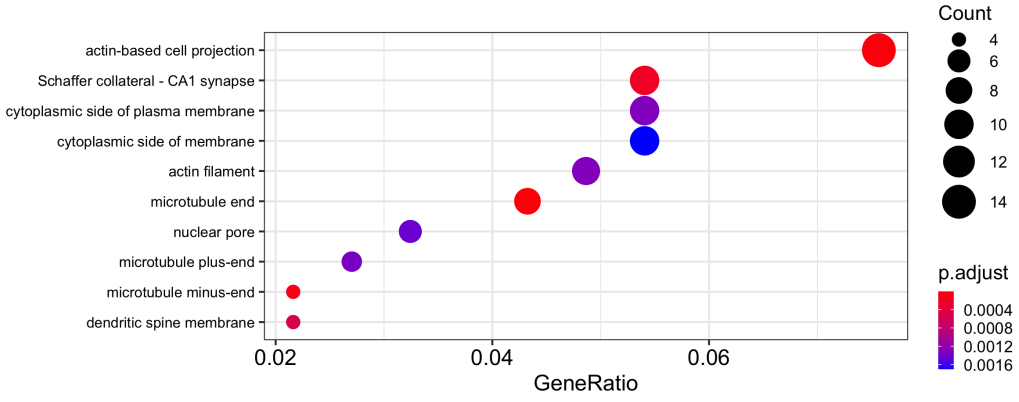
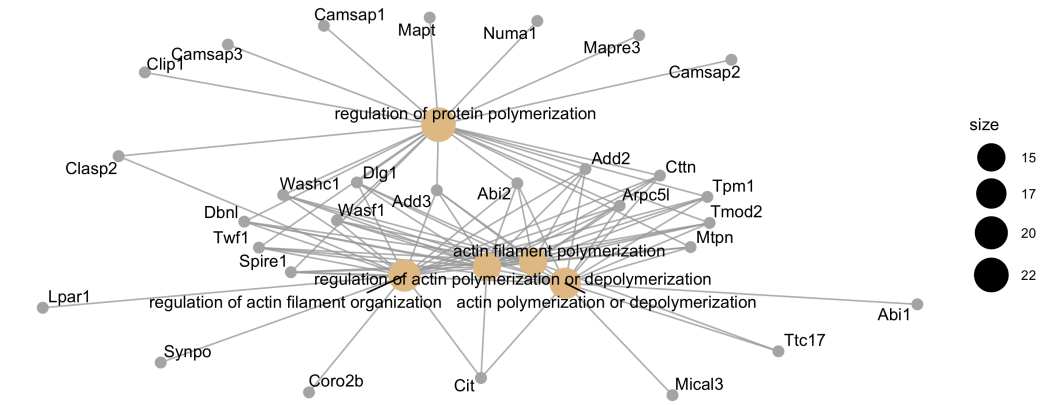
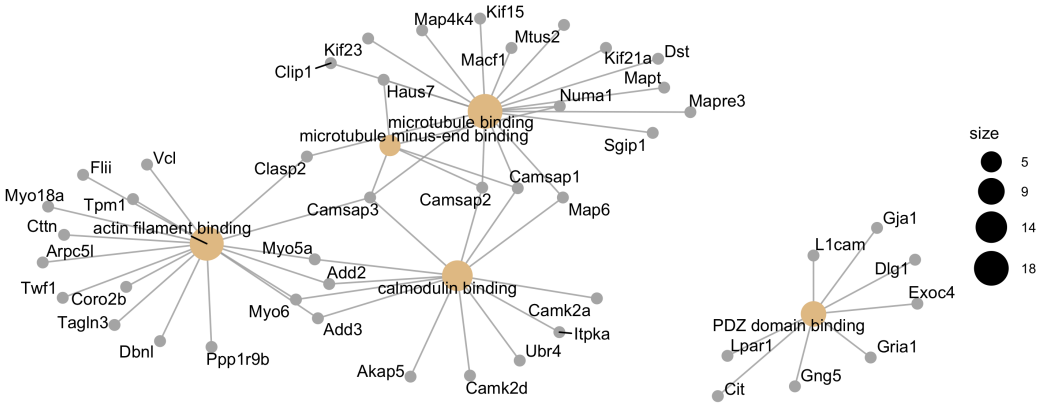


Figure 4.6: Dotplot of GO enrichment analysis E11.5 forebrain. **(a)** Biological Process. **(b)** Molecular Function. **(c)** Cellular Compartment. ”Count” indicates the number of genes in the dataset associated with a specific GO term and is represented by the size of the dot. ”Gene ratio” is the ratio between the number of genes associated with a specific GO term present in the dataset, over the total number of genes in the dataset. ”p.adjust” value is indicated by the colour of the dots with bright red dots corresponding to most significantly enriched GO terms.

(a) BP



(b) MF



(c) CC

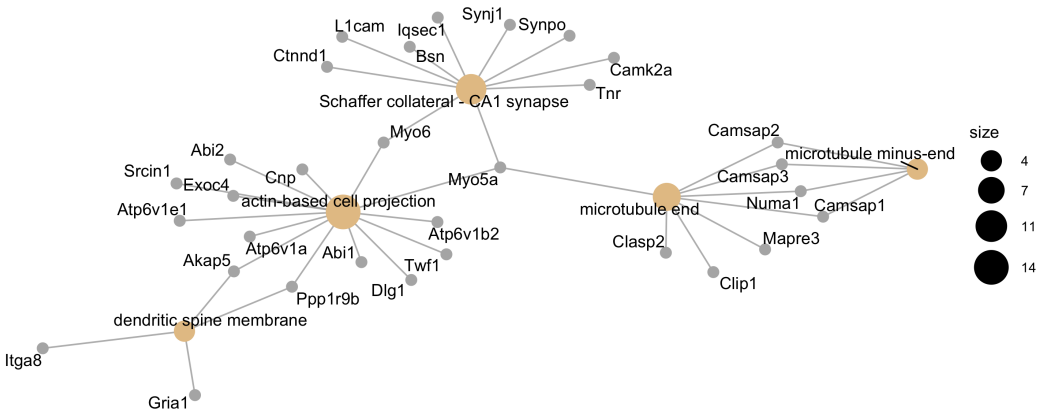


Figure 4.7: Cnetplot of GO enrichment analysis E11.5 forebrain. **(a)** Biological Process. **(b)** Molecular Function. **(c)** Cellular Compartment. Size of the dot corresponds to number of genes associated with the specific GO term.

of the acting cytoskeleton. DOCK3 is also known as presenilin-binding protein, as it interacts with PSN1/2 and can also bind the glutamate receptor subunit NR2D (Bai et al., 2013). Remarkably, several nuclear transport proteins, including importins and exportins were identified in the MS analysis and are listed in (**Table 4.3**). They include importins KPNA1, KPNA4, KPNB1, IPO5, IPO7, IPO9 and IPO13, and exportins XPO1, XPO5 and XPO7. Importins transport cargo proteins to the nucleus by recognising and binding specific sequences known as nuclear localisation sequences (NLS), whereas exportins mediate nuclear export instead. As described in (**Section 3.2.1**), PCDH19 contains at least one NLS signal in its cytoplasmic domain, which would explain some of these interactions and support a role of PCDH19-ICD in the nucleus.

Table 4.3: Nuclear transporters identified by LC-MS

Gene ID	Protein Name	Protein Function	Dataset
Ipo5	Importin 5	Nuclear import transporter	C
Ipo7	Importin 7	Nuclear import transporter	C
Ipo9	Importin 9	Nuclear import transporter	C
Ipo13	Importin 13	Nuclear import transporter	C
Kpna1	Karyopherin subunit alpha 1	Nuclear import transporter	C
Kpna4	Karyopherin subunit alpha 4	Nuclear import transporter	C
Kpnb1	Karyopherin subunit beta 1	Nuclear import transporter	C
Xpo1	Exportin 1	Nuclear export transporter	C
Xpo5	Exportin 5	Nuclear export transporter	C
Xpo7	Exportin 7	Nuclear export transporter	C

(C) cortex dataset.

4.2.6 KPNA1 interacts with PCDH19-CYTO

For the purpose of this thesis, which mainly focusses on the proteolytic processing of PCDH19 and its potential involvement in membrane to nucleus signalling, the identification of several importins was a particularly relevant finding. Hence, it was decided to validate the PCDH19 interaction with one of the identified proteins *in vitro*. Karyopherin alpha 1 (KPNA1), also known as importin subunit α5, was selected because of the commercial availability of a MYC- tagged construct that would simplify pull-down and blotting of this protein of interest. Validation of PCDH19 interaction with KPNA1 was done by overexpression of the KPNA1-MYC construct in combination with either PCDH19-HA or PCDH19-CYTO-HA in HEK293T cells. Using immunoprecipitation with an anti-MYC antibody, it was confirmed that PCDH19 can interact with KPNA1, and whilst this interaction is barely detectable for full length PCDH19, interaction with PCDH19-CYTO seems to be much stronger. These results suggest that, when cleaved, PCDH19 interacts more readily with KPNA1 (**Figure 4.8**).

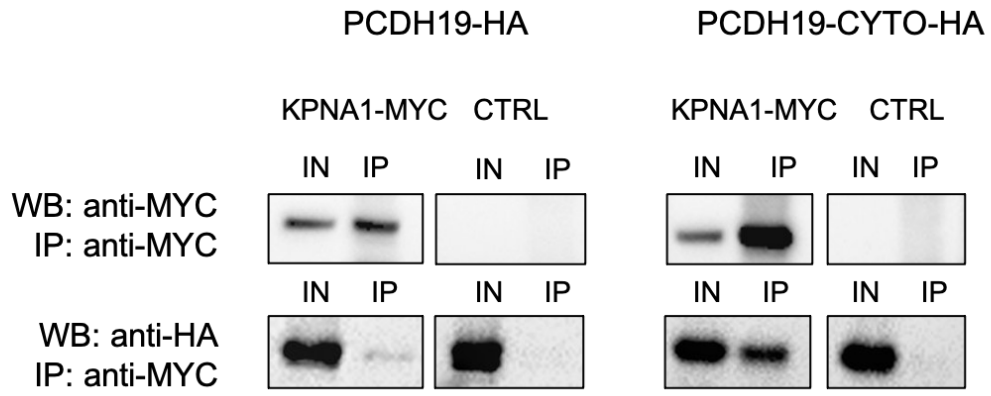


Figure 4.8: PCDH19-CYTO-HA interacts with KPNA1-MYC. HEK293T cells were transfected with PCDH19-HA or PCDH19-CYTO-HA with or without MYC-tagged KPNA1, karyopherin alpha 1. Immunoprecipitation was performed with anti-MYC antibody and INPUT/IP samples were blotted with anti-MYC and anti-HA. PCDH19-CYTO-HA shows a much stronger interaction with KPNA1 than the full lenght protein, suggesting that PCDH19 needs to be cleaved in order to interact with the importin.

4.3 Discussion

This chapter focussed on determining novel interactors of PCDH19 in mouse cortical tissue, derived from either embryos or adult animals. The overlap between the two obtained datasets was relatively small, suggesting that PCDH19 might be playing different roles during development and in adulthood. Given the highly diverse cellular context and protein composition of progenitors and neurons, it is not entirely surprising that the PCDH19 interactomes from adult and E11.5 tissue would not have much in common. Importantly, several of the previously described interactors of PCDH19 were also identified during this experiment, thus adding validity to the method and increasing confidence in novel interactions. Amongst these, N-Cadherin, the components of the WAVE regulatory complex, ABI1-2 and WASF1-3 were detected. Moreover PCDH17 and DOCK7 were also detected. The results of the Gene Ontology analysis carried out on the two datasets and results highlighted a main, previously described role of PCDH19 in cytoskeletal reorganisation. This result was driven by the presence of several members of the WRC and also various actin binding proteins. Of note, for the adult cortex dataset, terms relating to nucleocytoplasmic trafficking were enriched.

Interestingly, several nuclear transporters were identified in this preliminary screen including both nuclear import and nuclear export proteins. This observation has not been reported previously but strengthens the initial hypothesis, brought forward in the previous chapter (**Chapter 3**), that PCDH19, or fragments of PCDH19, can be shuttled to and from the nucleus. NLSs present in the cytoplasmic domain of PCDH19, could be recognised by importin-α subunits. The interaction between protocadherins and importins has not been reported before, but there are different studies that show interaction between cytoplasmic fragments of membrane proteins and importins. For instance, Notch-ICD is bound by importin-α3 (also known as karyopherin-α3) to be trafficked to the nucleus, and this interaction is necessary for downstream effects on cell proliferation (Sachan et al., 2013). Several studies have shown localisation of protocadherin fragments in the nucleus (Haas et al., 2005), (Hambsch et al., 2005) (Magg et al., 2005). In particular, for protocadherin

Fat1 it was shown that the cytoplasmic domain localised to the nucleus, but when a predicted NLS was removed, Fat1 cytoplasmic domain was instead found in the cytoplasm (Magg et al., 2005).

Although the data are very preliminary and will need to be confirmed in other ways, one of the detected importins was further investigated. The interaction between PCDH19 and KPNA1 was replicated *in vitro*, using overexpressed, tagged proteins. Interestingly, the interaction between PCDH19 and KPNA1 is stronger with the PCDH19 cytoplasmic domain. This suggests that this interaction could happen more readily after PCDH19 processing. It is important to note that a PCDH19 C-terminal antibody was used for immunoprecipitation for the mass spectrometry analysis, therefore both full-length and cytoplasmic fragments, with their respective interactors, would have been pulled-down.

Unfortunately, the whole analysis has to be carefully interpreted because of the unexpected and yet unexplained detection of PCDH19 peptides in the *Pcdh19* KO cortex samples. Although no transcript has been described to date that encodes only for the cytoplasmic domain of PCDH19, in the *Pcdh19* KO transgenic mouse there could be some residual expression due to the presence of exons 4, 5 and 6 in the genome. Contamination of samples or residual expression in KO samples means that the cortical samples did not have a true negative control. Nonetheless, filtering was done using the KO samples as a list of non-specific interactions. Alternatively, the ”crapome” database could have been utilised as another feasible filtering system, but it only contains human and *S.Cerevisiae* proteins (Mellacheruvu et al., 2013). Overall, the filtering method applied for this analysis was quite stringent. Another analysis, applying a less stringent filtering, which allows very few (1-2) peptides to be detected in the KO samples, could have also been implemented and potentially given interesting results. In fact, the paraspeckle NONO, which is a PCDH19 interactor (Pham et al., 2017), would have been also part of the list of interactors, if the analysis would have been less stringent.

The method chosen in this study to identify PCDH19 interactors was immunoprecipitation. A

clear advantage is that it was done *in vivo* and on the endogenous protein, using a reliable and specific antibody. This was possible because of the availability of KO animals in the laboratory to be used as controls, although as discussed above, they seem to express very low levels of the cytoplasmic domain, which was only discovered once the analysis had been carried out. A main disadvantage of using this pull-down method is that it will only detect strong or stable interactions, that will survive the pull-down. Alternatively, the PCDH19 interactome could have been investigated by proximity-labelling by biotinylation using BioID (Roux et al., 2018). This method would make detection of interactions easier but has other drawbacks such as having to use an engineered and overexpressed protein fusion of BioID, which could interfere with protein function and/or interactions. BioID could also potentially create false positives by labelling proteins which are in very close proximity but not actually interacting. The strong advantage of BioID labelling, though, is that it can pick up very transient interactions between proteins, which comprise the majority of interactions necessary for intracellular signalling.

4.3.1 Conclusion

This chapter comprised a preliminary investigation into the interactome of PCDH19 and led to the uncovering of previously undescribed interactions between PCDH19 and proteins of the nuclear import pathway, suggesting a role for PCDH19 in the nucleus. In particular, the interaction between PCDH19 and KPNA1 was replicated *in vitro*. This interaction appears to be stronger when involving just the cytoplasmic domain of PCDH19, suggesting that PCDH19 might need to be processed in order to interact with importins and be trafficked into the nucleus.

Chapter 5

Generation of an *in vitro* model to study the function of PCDH19 cytoplasmic domain

5.1 Introduction

Embryonic stem cells can be easily manipulated *in vitro* in order to induce modifications at the genomic level and generate cell lines that stably express a construct of interest. For this reason, ESCs are a good model to investigate disease processes and biological function of proteins. Remarkably, protocols have been generated to differentiate ESCs *in vitro* in many different cell-types in a reproducible manner. The greatest advantage of these protocols is the generation of a homogeneous population of cells, crucial for any bulk downstream analysis, such as RNA sequencing. Genetic engineering can be achieved by different methods, of which zinc finger nucleases and CRISPR/Cas9 are the best-established. ZFNs are engineered restriction enzymes that contain a series of DNA-binding zinc finger domains fused to a FokI endonuclease. ZFNs create a double strand break at the target DNA sequence (Perez-Pinera et al., 2012)

(**Figure 5.1a**). CRISPR-Cas9 technology, emerged more recently, is even more adaptable than ZFNs (Ran et al., 2013). It allows precise genome editing, efficiently and cost-effectively, and is adaptable to a myriad of different DNA sequences, with the only limiting factor being the presence of a 3 bp PAM sequence, upstream of which Cas9 will induce a double strand break (DBS) (Ran et al., 2013) (**Figure 5.1b**). For both methods, precise genome editing can be achieved when supplementing the cells with a repair template which initiates the cell endogenous repair mechanism of homology directed repair (HDR). If a repair template is not provided, the cells will instead resort to patching up the DNA damage via the non-homologous end joining (NHEJ) pathway, which is an error prone mechanism and can often result in insertion or deletions (indels) and consequently frameshits and premature stop codons.

Although different efforts have been undertaken to study behaviour of PCDH19 KO cells *in vitro*, reviewed in the introduction (**Section 1.2.4**), these cells have in most cases been primary cells derived from the *Pcdh19* KO mouse (Mincheva-Tasheva et al., 2021). PCDH19-epilepsy patient-derived induced-pluripotent stem cells (Homan et al., 2018) have also been derived and studied but, so far, there are no reports on the generation and investigation of mouse ESC-derived neurons which are PCDH19-KO or express mutated/truncated forms of PCDH19.

As previously discussed (**Chapter 3**), PCDH19 is subject to proteolytic processing, which results in the release of its intracellular domain (ICD). Once released, PCDH19 ICD can interact with importins (**Chapter 4**) and be shuttled into the nucleus, but it is yet to be determined if proteolytic processing is indeed a membrane-to-nucleus signalling mechanism.

In order to assess if PCDH19 proteolytic processing and release of the ICD is a mechanism capable of inducing changes in gene expression in neurons, a mESC line overexpressing the ICD of PCDH19 was generated. This *in vitro* model was produced in order to assess the effect of overexpression of the cytoplasmic domain of PCDH19 on gene expression during neuronal formation and maturation and compare WT, KO and ICD overexpression transcriptional profiles of

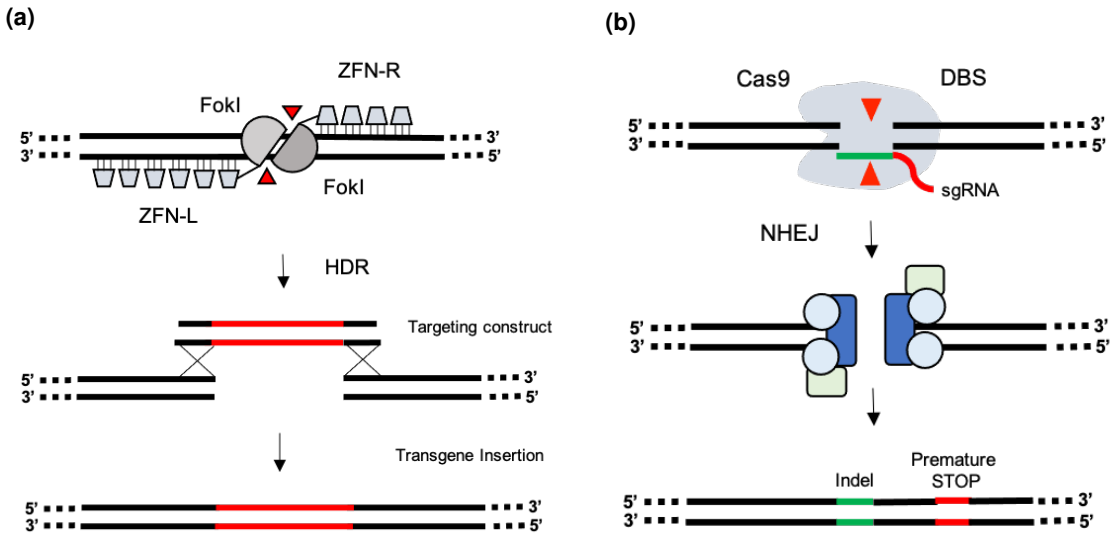


Figure 5.1: Geneting engineering of ESCs. (a) Zinc finger nuclease engineering. ZFNs work as dimers, each consisting of a series of zinc finger DNA-binding modules fused to the type IIS restriction endonuclease FokI. In the case of *Rosa26* locus targeting, the left ZFN (ZFN-L) consists of 6 zinc finger domains, whilst the right ZFN (ZFN-R) is formed by 4. Each zinc finger domain binds to a specific 3bp sequence on the target DNA. FokI is unspecific but will cleave the DNA where directed by the ZFNs, inducing a double strand break (red arrowheads). Genome editing is achieved by providing a repair template, with homology to the disrupted site, in order to allow for homology directed repair (HDR) to take place. Adapted from (Perez-Pinera et al., 2012). **(b)** Cas9 activity creates a double strand break (DBS) at the target site. Endogenous DNA-repair mechanism of Non-Homologous End Joining (NHEJ) occurs to repair the damage but, being an error prone mechanism, it often results in insertions or deletions (indels) which lead to frameshifts and/or premature stop codons. Adapted from (Ran et al., 2013).

mESC-derived neurons by RNA sequencing analysis.

5.1.1 Aims

The aim of this chapter is to set up an *in vitro* model to investigate the downstream effects of PCDH19 proteolytic processing on gene expression in mouse ESC-derived cortical neurons. This chapter outlines the generation of the targeting vectors and of the mouse embryonic stem cells overexpressing the ICD of PCDH19 (PCDH19-CYTO). It also outlines the design and generation of gRNAs against *Pcdh19* for the creation of isogenic PCDH19-KO mouse embryonic stem cells. Finally, it includes the validation of these cells in terms of PCDH19 protein expression profile and their ability to differentiate into cortical-like neurons following an established protocol.

5.2 Results

5.2.1 Generation of *Rosa26*-PCDH19-ICD-HA targeting vector

To assess the function of the intracellular domain (ICD) of PCDH19, a mouse embryonic stem cell line, referred to as PCDH19-CYTO, constitutively expressing the cytoplasmic domain of PCDH19 was generated. The Gt(ROSA)26Sor locus, or *Rosa26* locus was chosen as the targeting locus for insertion of the transgene, as it is a well-characterised, constitutively and ubiquitously expressed locus on chromosome 6 of the mouse genome (Perez-Pinera et al., 2012).

A targeting vector for the *Rosa26* locus was generated that contained the ICD of PCDH19, including the last portion of exon 1, and exons 2, 3, 4, 5 and 6, and corresponding to the whole intracellular domain with a predicted protein size of 51 kDa. PCDH19 ICD was inserted in between the 5' and 3' *Rosa26* homology arms. Expression of PCDH19 ICD is blocked by the presence of a neomycin resistance cassette which is flanked by two loxP sites. Expression of PCDH19 ICD therefore can only happen after Cre recombination. The neomycin resistance cassette allows for antibiotic selection of targeted clones (**Figure 5.2**).

The *Rosa26*-PCDH19-ICD-HA targeting vector plasmid was generated starting from the pZDRosa-loxP-NeoR-loxP-IRES-EGFP plasmid (**Table 2.8**). The pZDRosa-loxP-NeoR-loxP-IRES-EGFP plasmid contains the 5' and 3' targeting arms for the Rosa26 locus and a floxed neomycin resistance cassette followed by an IRES-EGFP. The IRES-EGFP fragment was not necessary and was excised by double restriction digestion with *BsrGI*-HF and *AscI*, leaving a 7322 bp backbone (**Figure 5.3a**). Details are provided in the methods (**Section 2.5.1**).

In parallel, the PCDH19-CYTO-HA fragment was amplified from the pre-existing plasmid pCIG-PCDH19-CYTO-HA (**Table 2.2**) using primers *Rosa26*-CYTO-HA-F2 and *Rosa26*-CYTO-HA-R (**Table 2.5**). The resulting PCR product of 1385bp (**Figure 5.3b**) was cloned between the 3' and 5' arms of the *Rosa26* targeting vector by InFusion reaction. Clones were screened for the correct

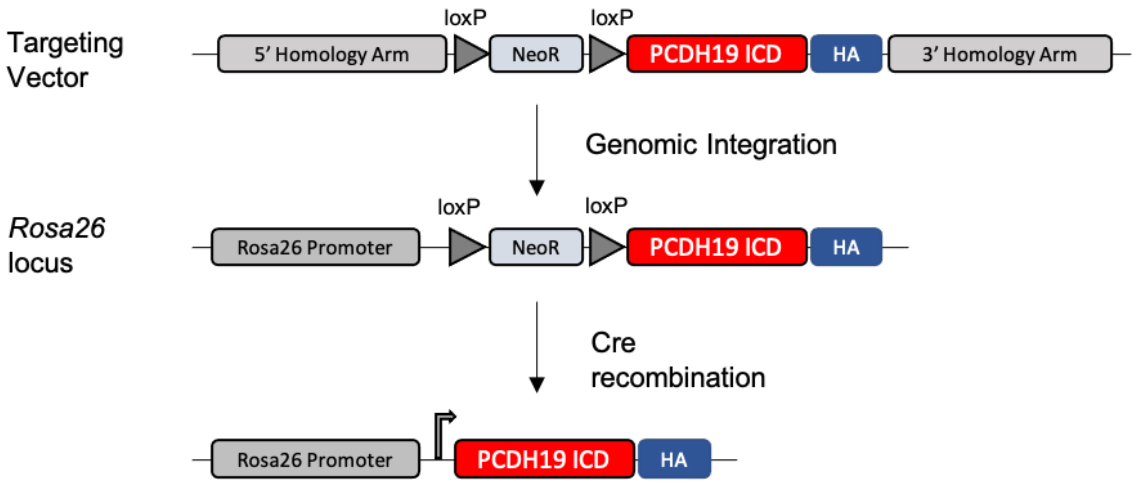


Figure 5.2: Zinc-finger nuclease targeting of the *Rosa26* locus. The provided repair template consists of 5' and 3' homology arms to the *Rosa26* locus, a Neomycin resistance cassette (NeoR) flanked by loxP sites, and the whole intracellular domain of PCDH19 with a fused HA-tag at the C-terminus. The NeoR allows for selection of positively recombined clones but prevents expression of the transgene. Only after Cre recombination, PCDH19-ICD-HA expression is constitutively turned on, under the control of the *Rosa26* promoter. ZFN-L/R, zinc finger nuclease left/right; HDR, homology directed repair; ICD, intracellular domain.

insert by restriction analysis with *HindIII*. Clones containing PCDH19-CYTO-HA were predicted to generate three fragments of 4.7 kb, 2.5 kb and 1.4 kb whilst negative clones produced 2 fragments (**Figure 5.3c**). Clones 1, 9 and 10, which displayed the expected restriction digest pattern, were sent for sequencing, which confirmed insertion of PCDH19-CYTO-HA.

5.2.2 Targeting mESCs for generation of PCDH19-CYTO mESC line

The PCDH19-CYTO mouse embryonic stem-cell line was created by zinc finger nuclease (ZFN) targeting of the *Rosa26* locus, using previously engineered plasmids (**Table 2.7**) (Perez-Pinera et al., 2012). In brief, stem cells were nucleofected with the zinc finger nuclease plasmids and the *Rosa26* targeting construct containing PCDH19-ICD-HA, as described in methods (**Section 2.5.2**). Positive clones were selected through antibiotic selection (**Section 2.5.3**). Clones that displayed antibiotic resistance, indicating the insertion of the targeting construct, were genotyped with primers inside and outside the 5' and 3' targeting arms (**Table 2.9**), in order to assess that insertion of the transgene occurred in the correct locus. Clones that showed bands of the predicted sizes were further nucleofected with the pCIG-CRE plasmid, expressing Cre recombinase, to excise the neomycin selection cassette that prevents expression of the transgene (**Figure 5.2**). This second nucleofection step had to be followed by another round of colony picking, hence generating subclones from the originally targeted clones. After excision of the NeoR cassette, PCDH19-ICD-HA is constitutively expressed under the control of the *Rosa26* promoter (**Figure 5.2**). PCDH19-ICD-HA expression was tested by western blot on protein lysates from several targeted clones using both anti-HA and anti-PCDH19 antibodies. Both antibodies detected a band of the appropriate size just below the 50 kDa mark, consistent with the predicted size of the protein, 51 kDa (**Figure 5.4**).

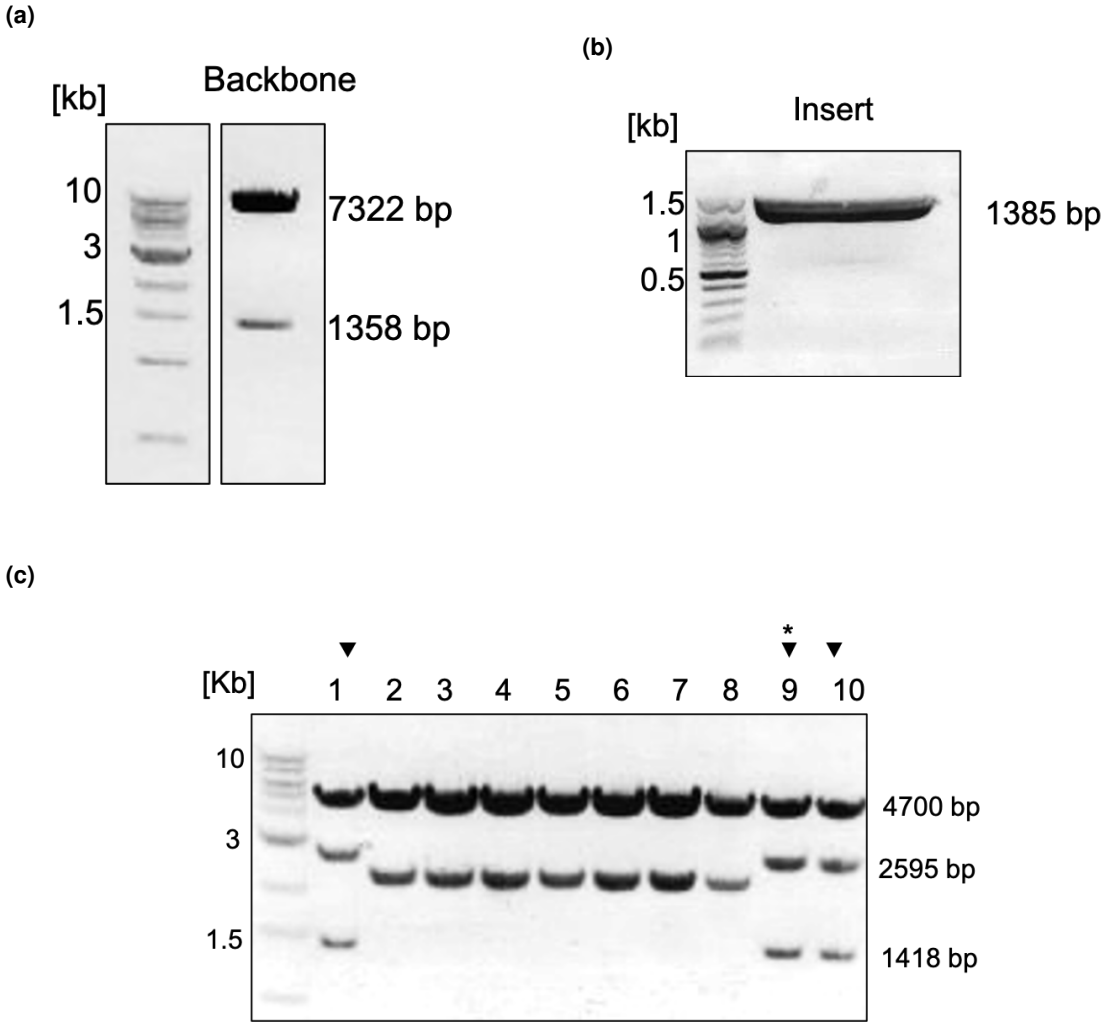


Figure 5.3: Generation of the *Rosa26*-floxedNeo-PCDH19-ICD-HA targeting vector. **(a)** *BsrGI*-HF and *AscI* restriction digestion of starting plasmid pZDRosa-floxedNeo-IRES-EGFP for excision of the IRES-EGFP fragment. **(b)** PCR amplification of PCDH19-ICD from pCIG-PCDH19-ICD. **(c)** *HindIII* restriction digestion analysis of floxedNeo-PCDH19-ICD-HA vector. Correct clones contain three fragments of 4.7 kb, 2.5 kb and 1.4 kb (black arrowheads). The selected clone used for targeting is indicated by an asterisk.

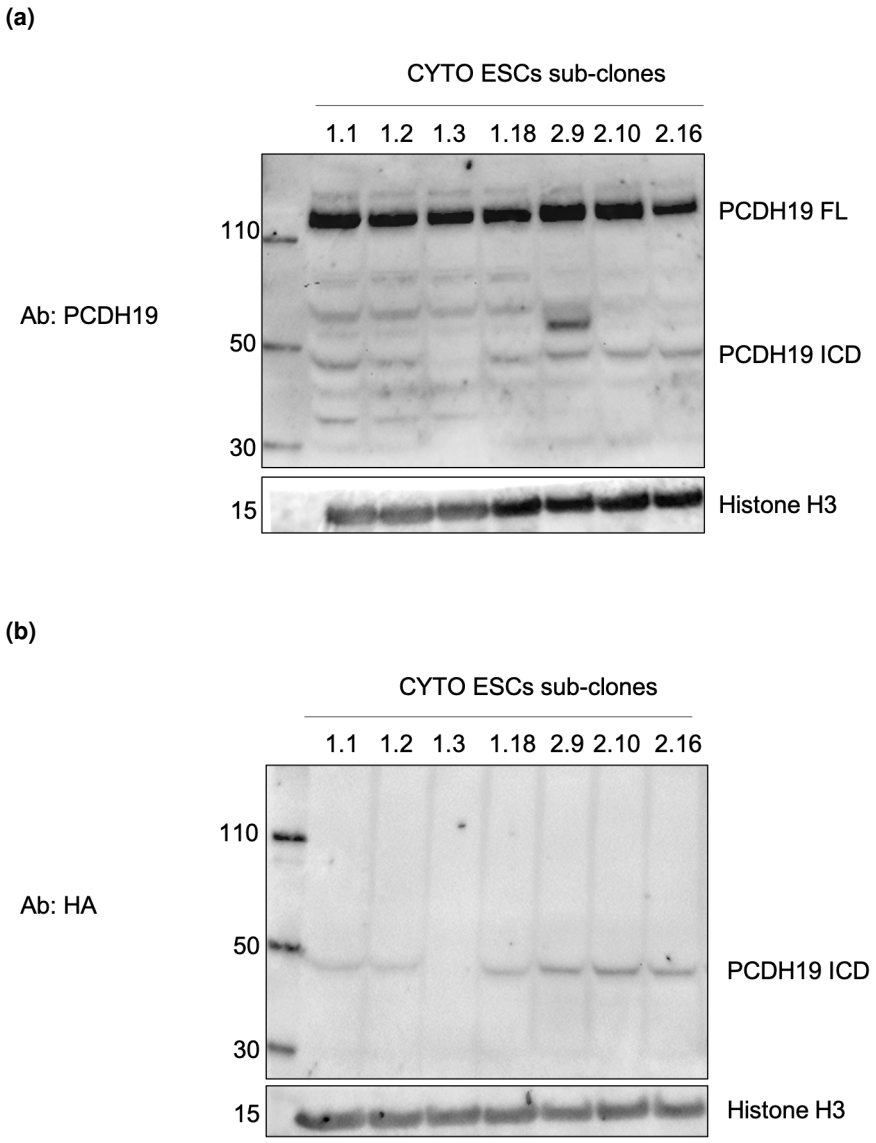


Figure 5.4: PCDH19-ICD-HA expression in targeted ESC subclones. (a) Western blot against PCDH19 on protein lysates from targeted subclones after reverse selection shows that targeted subclones express the full-length PCDH19 (over 110 kDa) and the cytoplasmic domain (below 50 kDa). **(b)** PCDH19-ICD-HA expression is further validated in the same ESC subclones by reactivity to the anti-HA antibody. Subclones 1.1, 1.2, 1.3 and 1.18 are derived from clone 13, whilst subclones 2.9, 2.10 and 2.16 are derived from clone 40.

5.2.3 Design and cloning of the gRNAs against *Pcdh19*

Although *Pcdh19* KO ES cells were present in the Martinez-Garay lab, these had been derived from the *Pcdh19* KO mouse and therefore have a different genetic background than the E14 cells. In order to have an isogenic cell line to the PCDH19-CYTO cell line, a PCDH19-KO E14 ESC line was generated by CRISPR-Cas9 genome editing. Two gRNAs were designed to target the beginning of *Pcdh19* exon 1 using the online CRISPR design tool CRISPOR (crispor.tefor.net, by Maximilian Haeussler and Jean-Paul Concordet (Concordet and Haeussler, 2018), a website that predicts off-target effects and helps design effective guide RNAs for different Cas9 proteins. gRNA1 5'-CGGGACGGTGATCGCTAACG (TGG)-3' and gRNA2 5'-CGCATTACGGCTCTCGACGG (AGG)-3' were chosen (**Table 2.6**). In parenthesis is the Protospacer Adjacent Motif (PAM) 5'-NGG-3' (with N being any nucleotide base), required for Cas9 nuclease cleavage which needs to be downstream of the target sequence (**Figure 5.5b**). Both chosen gRNAs have only 5 predicted off-target effects based on four mismatches – with only one off-target being in a protein-coding region (*Acin1* for gRNA1 and *Stat5a* for gRNA2) and with the other off-target effects being either intergenic or intronic. Target specificity based on the MIT Specificity score was 99 for both guides and was a preferred criterion to out-of- frame and indel scores (prediction of the successful indel) to minimise off-target effects.

The gRNAs were individually cloned into the Cas9-mCherry fusion plasmid (pU6-(*BbsI*)-CBh-Cas9-T2A-mCherry) created by Ralph Kuhn (Chu et al., 2015) (**Table 2.8**), as explained in methods (**Section 2.5.1**). This plasmid contains the Cas9 nuclease from *Streptococcus Pyogenes*, flanked by nuclear localisation sequences, linked through a T2A sequence (self-cleaving peptide) to a mCherry fluorescent reporter protein (**Figure 5.5a**). Expression of the Cas9 nuclease will therefore be matched by expression of mCherry, providing a readout for successful nucleofection. The plasmid also includes a *BbSI* restriction site for insertion of customised gRNAs.

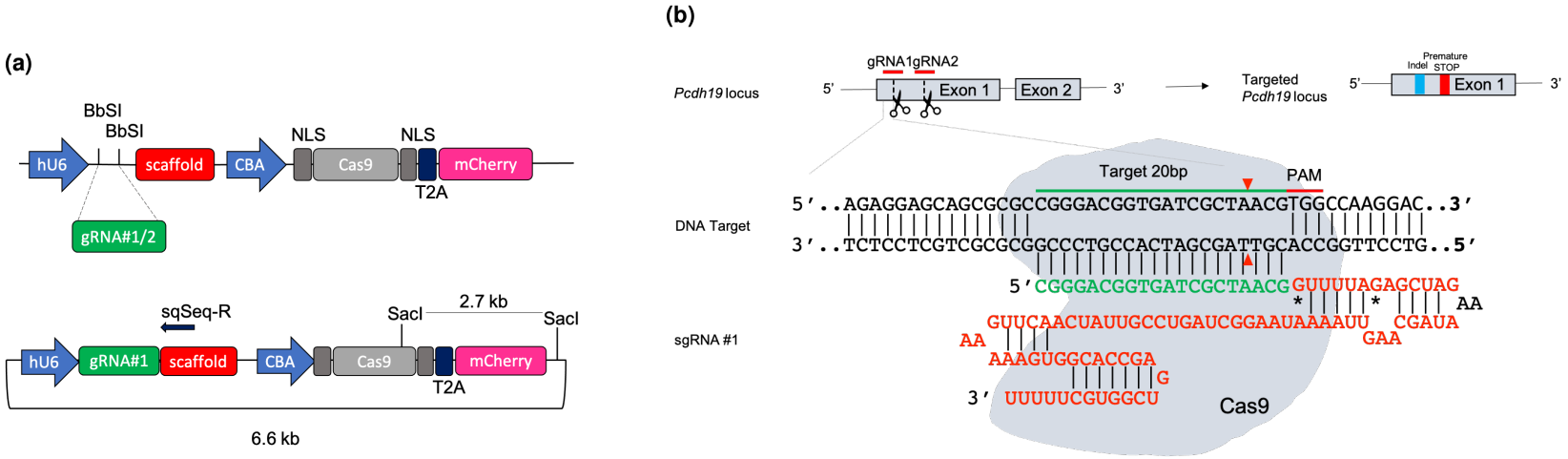
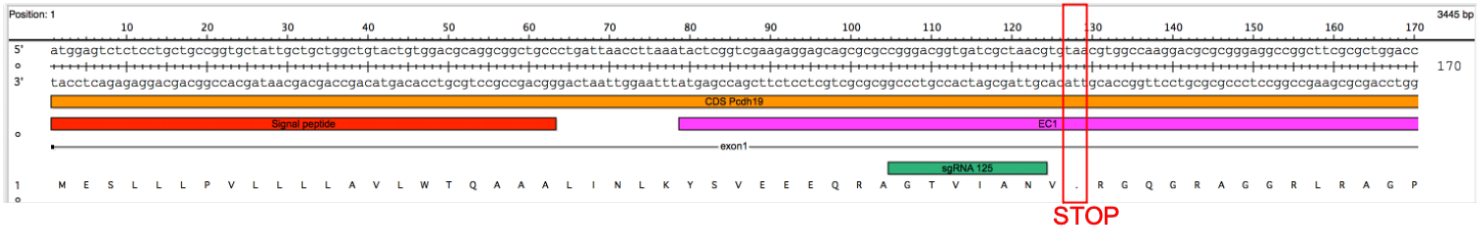


Figure 5.5: CRISPR/Cas9 targeting of the *Pcdh19* locus. (a) Outline of the Cas9-mCherry plasmid. **(b)** Schematic for RNA-guided Cas9 nuclease activity with gRNA1 used as example. The gRNA (green) binds a homologous sequence on the target DNA. In red is the gRNA scaffolding. The red arrowheads indicate the Cas9 cleavage site, 3 bp upstream of the Protospacer Adjacent Motif (PAM) 5'- NGG -3' (with N being any nucleotide base), the grey shape is the Cas9 nuclease enzyme, asterisks indicate mismatched base. Adapted from (Ran et al., 2013).

5.2.4 Generation of the PCDH19-KO mESCs

To generate the PCDH19-KO mESCs, the same mouse embryonic stem cells were utilised, E14s. This was done to ensure that the PCDH19-CYTO and -KO cell lines had the same genetic background, in order to minimise differences not due to expression levels of PCDH19. Cells were nucleofected with the pU6-gRNA1-CBh-Cas9-T2A-mCherry and pU6-gRNA2- CBh-Cas9-T2A-mCherry plasmids (**Table 2.8**), as previously described (**Section 2.5.2**). In order to select cells that were successfully nucleofected and facilitate the screening process, mCherry+ cells were selected via fluorescence-activated cell-sorting (FACS). Cells were individually plated as single-cells in a 96-well plate and allowed to recover for several days until they had grown into colonies (**Section 2.5.4**). Colonies were amplified for extraction of both genomic DNA and protein. Subsequently, clones were screened by PCR, with primers located either side of the predicted Cas9 cut site (**Table 2.10**). In parallel, protein lysates collected from the ESC clones were screened by western blot against PCDH19. As a consequence of CRISPR-Cas9 targeting, several clones lost expression of the PCDH19 full-length protein (**Figure 5.7**). Successfully targeted clones, lacking PCDH19, show instead a faint band at the height of the PCDH19-full length protein. However, this band is also present in control KO mESC, which are derived from the *Pcdh19* KO mouse model. As previously described (**Figure 1.4a**), the *Pcdh19* KO mouse completely lacks exons 1-3, lacking the full-length protein; therefore that band can safely be assumed to be unspecific. Finally, in order to confirm successful targeting, the genomic DNA around the cut site was amplified and sequenced. PCDH19-KO clones 2 and 18 had a premature STOP codon just after the gRNA1 sequence (**Figure 5.6**). Clone 2 was used for RNA sequencing (**Chapter 6**). Changes in gene expression of the predicted off-target effects of gRNA1 and gRNA2, described in (**Subsection 5.2.3**) were checked by RNA sequencing (**Chapter 6**). It was confirmed that expression of both *Acin1* and *Stat5a* was unchanged between *Pcdh19* KO and WT cells, for both progenitors and neurons.

KO clone #2



KO clone #18

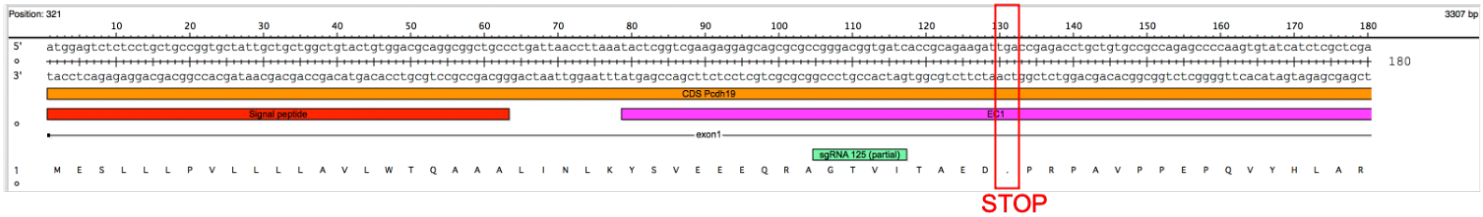


Figure 5.6: *Pcdh19* KO sequencing. Sequence of PCDH19-KO clones 2 and 18 which show premature stop codon within exon 1 of *Pcdh19*, indicated by a dot within the amino acid sequence. The sequence recognised by gRNA1 (gRNA 125) is shown in green and is just before the premature stop codon.

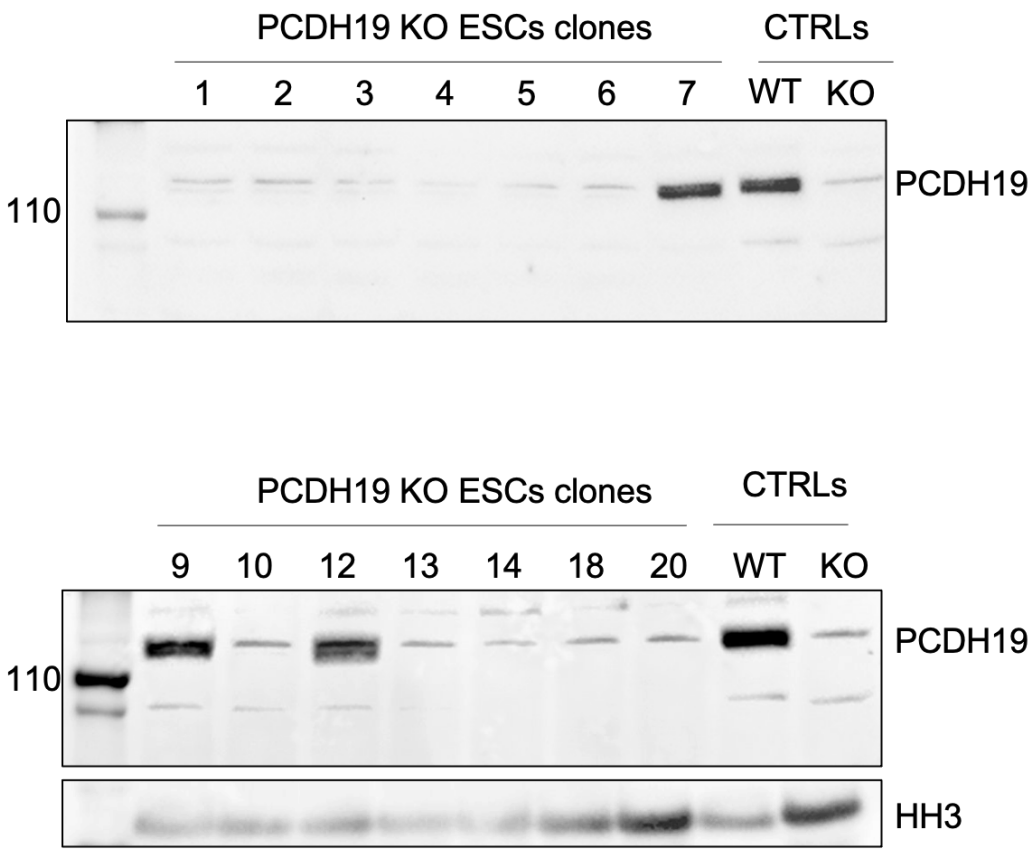


Figure 5.7: Lack of PCDH19 in PCDH19-KO clones. Western blot with anti-PCDH19 antibody on lysates from the PCDH19-KO ESC clones. KO control sample is *Pcdh19* KO ESCs derived from the *Pcdh19* KO mouse.

5.2.5 Karyotyping of PCDH19-CYTO and PCDH19-KO mESCs

Genome editing can lead to DNA damage at unwanted locations, even resulting in chromosomal truncations (Cullot et al., 2019). It is therefore necessary to ensure the clones have the right number of chromosomes before using them in downstream applications like RNAseq. ESC clones were karyotyped to ensure that at least 70% of cells per clone had the correct number of chromosomes (40 chromosomes), as previously described (**Section 2.5.6**). Only clones that had at least 70% of the cells with the right number of chromosomes were used for further experiments. PCDH19-CYTO cells were karyotyped twice, before and after removal of the selection cassette.

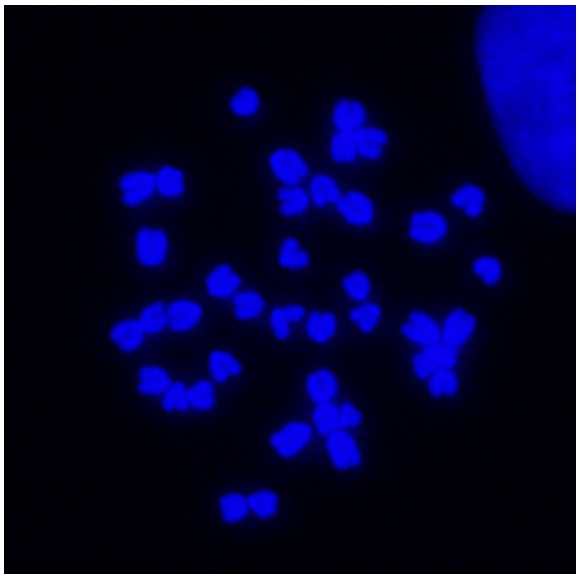


Figure 5.8: Representative karyotype. Chromosomes are visualised under the fluorescence microscope via DAPI stain in order to count individual chromosomes.

5.2.6 Neuronal differentiation of PCDH19-CYTO and PCDH19-KO mESCs

Once PCDH19-CYTO and PCDH19-KO embryonic stem cells were obtained, it was necessary to assess if these cells were capable of differentiating into cortical-like neurons. Based on genotyping and protein expression, several clones from both cell lines were selected for neuronal differentiation (CYTO clones 13.18, 40.10 and 40.16 and PCDH19 KO clones 2, 18 and 20). The previously described protocol (**Section 2.4.1**) (Bibel et al., 2004) was used to induce neuronal differentiation (Bibel et al., 2004).

PCDH19-KO clones 2, 18 and 20 were all able to differentiate, although giving rise to a noticeable number of non-neuronal cells. For the PCDH19-CYTO clones, there was a very high variability, with clones 40.16 and 40.10 unable to produce any neurons at all, whilst clone 13.18 could differentiate but giving rise to a high proportion of non-neuronal cells.

E14 cells, unlike other embryonic stem cells, are “feeder-independent”, which means that they can be grown directly on gelatin-coated dishes, without the need for an underlying feeder layer. Feeder cells are growth-arrested mouse embryonic fibroblasts (MEFs) that secrete growth factors that promote proliferation of ESCs. In some instances, growing ESCs on feeder cells can improve ESC quality, although the precise mechanism of how this occurs is still unknown. Given the non-homogeneity of the initial neuronal cultures obtained, it was decided to improve the quality of the ESCs by passaging them on MEFs before differentiation. Homogeneity of the neuronal cultures was very important for the successful outcome of the experiment, since the cultures were to be used for RNA sequencing. Indeed, passaging on MEFs drastically improved the morphological quality of the cells, both CYTO and KO, and reduced the overall number of non-neuronal cells during differentiation. After passaging on MEFs, CYTO and KO neurons appeared morphologically similar to WT neurons and gave rise to a negligible amount of non-neuronal cells.

Previous work in the Martinez-Garay laboratory, conducted by Dr. Jessica Griffiths, had shown a significant difference in the number of *Pcdh19* KO neuronal progenitor cells produced during neuronal differentiation, compared to WT. As discussed, these *Pcdh19* KO ESCs were derived from the *Pcdh19* KO mouse model. It was therefore important to monitor absolute number of cells obtained at CA8, day of dissociation of the aggregates. Interestingly, in the E14 genetic background, both KO and CYTO cells, showed proliferation rates similar to those of WT (one-way ANOVA: F (2, 12) = 0.3320, P = 0.7239) (**Figure 5.9**).

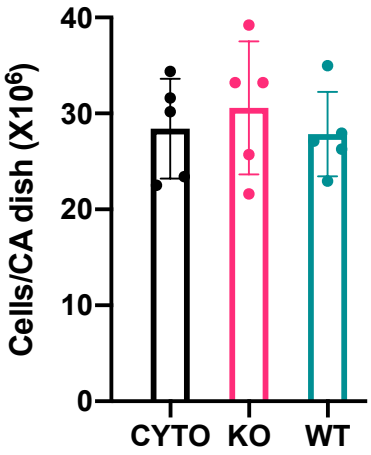


Figure 5.9: PCDH19-KO and PCDH19-CYTO ESC-derived cellular aggregates gave rise to equal numbers of progenitor cells compared to WT. Absolute number of cells counted on the day of dissociation of the cellular aggregates (CA8). On the day of aggregation (CA0), 4 million cells are used per dish. Over the course of 8 days – cell numbers increased 6-7 fold, giving rise to approximately 25 million progenitors/dish. No differences in the numbers of progenitor cells of different genotypes were observed.

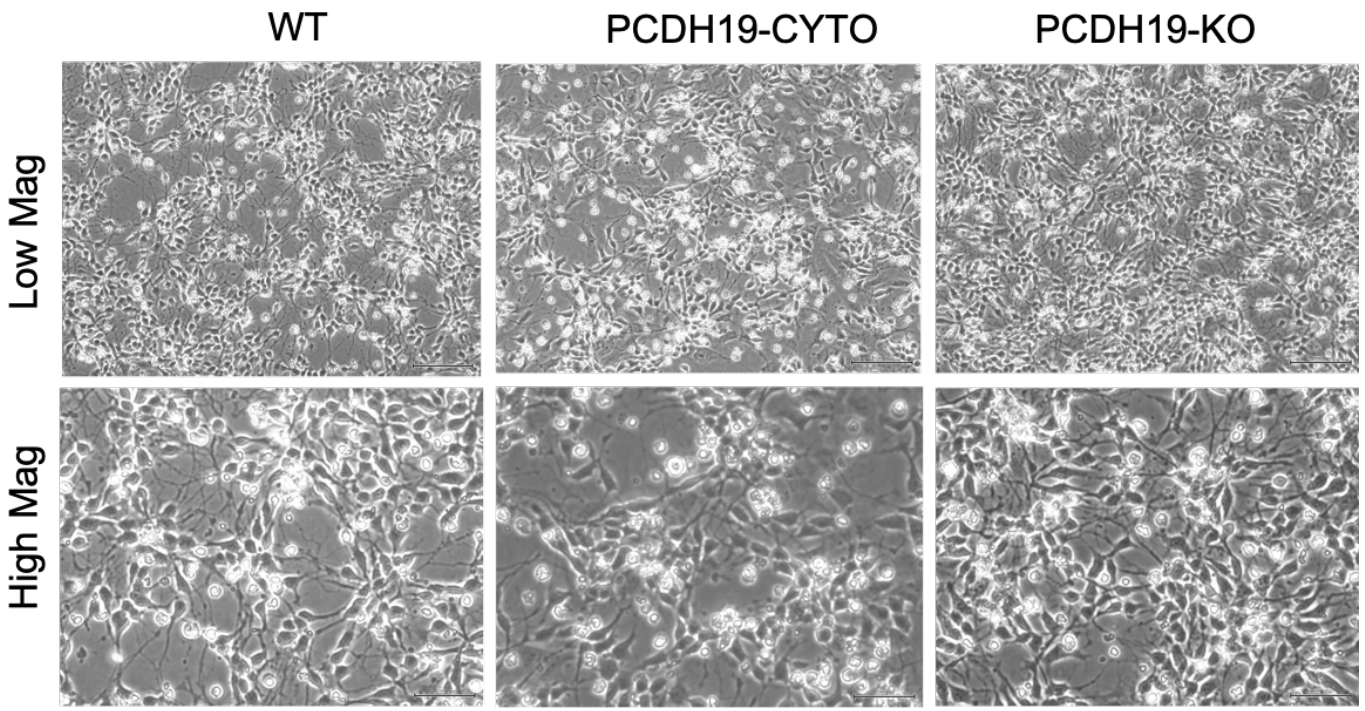


Figure 5.10: PCDH19-KO and PCDH19-CYTO ESCs can differentiate into cortical-like neurons. Post-passaging on MEFS, PCDH19-CYTO and PCDH19-KO mESC quality is improved and cell can differentiate into cortical-like neurons, as WT.

5.2.7 PCDH19 expression in PCDH19-CYTO and PCDH19-KO neurons

It was also necessary to assess whether PCDH19-ICD expression was maintained during neuronal differentiation and to verify that in the PCDH19-KO neurons, expression of the protein was completely gone. For this purpose, protein lysates were extracted from neurons at DIV8 and DIV12 and blotted with the PCDH19 C-terminal antibody. DIV8 and DIV12 were selected because these timepoints were used, later on, for the RNA sequencing analysis. Both WT and CYTO cells express the full-length protein (above 110 kDa), whilst in PCDH19 KO neurons the band is absent. In PCDH19-CYTO neurons an extra band, just under 50 kDa is visible, corresponding to the intracellular domain (**Figure 5.11**), in agreement with what was previously observed in the ESCs. PCDH19-CYTO neuronal lysates were also blotted against HA, to reveal the cytoplasmic domain, confirming that the fragment seen below 50 kDa corresponds to overexpression of the transgene and is not just the result of processing of the endogenous protein (data not shown).

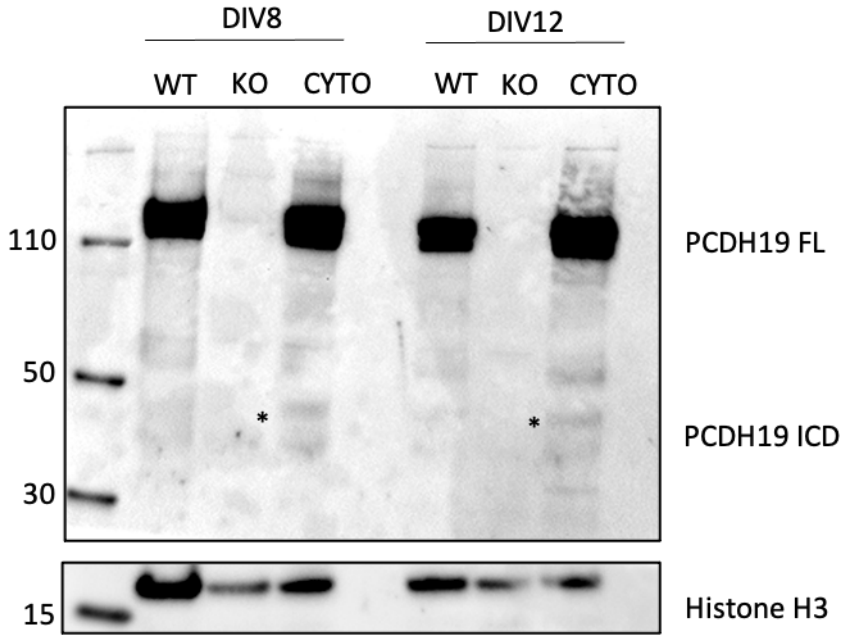


Figure 5.11: PCDH19 expression in WT, KO and CYTO neurons. Western blot against PCDH19 shows expression profile of PCDH19 at DIV8 and DIV12 (timepoints chosen for RNA sequencing). KO neurons show complete lack of the protein, whilst CYTO neurons maintain expression of the endogenous full-length protein (over 110 kDa) and also show the overexpression of the cytoplasmic domain (below 50 kDa).

5.3 Discussion

This chapter outlines the generation of an *in vitro* system to study the role of PCDH19 in cortical neurons and, in particular, to assess if the cytoplasmic domain, which is released by proteolytic processing, has a role in regulating gene expression. This chapter also outlined the design and generation of gRNAs against *Pcdh19* for the creation of isogenic PCDH19 KO mouse embryonic stem cells, to be used for RNA sequencing analysis. *Pcdh19* KO mouse embryonic stem cells derived from the *Pcdh19* KO mice were already present in the Martinez-Garay lab, but being a different genetic background, they were unsuitable for RNA-sequencing analysis, as they would have introduced unnecessary variability to the experiment. Generation of both PCDH19-CYTO and PCDH19-KO cells was successful, as demonstrated by PCDH19 expression profile in these cells.

As this was the first attempt at overexpressing the cytoplasmic domain of PCDH19, it was necessary to assess if the created cells were in fact capable of producing cortical-like neurons, following the established protocol (Bibel et al., 2004). This is especially critical when considering that overexpression of the cytoplasmic domain of PCDH19 could result in a dominant-negative effect by sequestering endogenous binding partners in the cytoplasm. A dominant-negative effect has been, in fact, previously reported for the overexpression of the cytoplasmic domains of N-Cadherin and E-cadherin (Nieman et al., 1999), (Togashi et al., 2002). In some instances, in order to have the dominant-negative effect, N-Cadherin cytoplasmic domain had to be targeted to the membrane. This was achieved by addition of a myristoylation signal (Nieman et al., 1999). In other cases, membrane targeting of N-Cadherin was not necessary in order to generate a dominant negative effect. Expression of dominant negative cadherins can have a dramatic impact on cell morphology and/or on cell proliferation, driven by loss of adhesion. For example dominant-negative N-Cadherin expressed by *in utero* electroporation in cortical progenitors can disrupt radial migration (Franco et al., 2011), (Gil-Sanz et al., 2013). Another example is provided by NF-Pcdh lacking the extracellular domain, which leads to ectoderm disruption in *Xenopous*

embryos, due to sequestration of binding partners, such as TAF1 (Heggem and Bradley, 2003). NF-Pcdh dominant negative expression also leads to defects in axon elongation of retinal ganglion cells (Piper et al., 2008).

PCDH19-CYTO cells were able to differentiate into cortical-like neurons and could proliferate at the same rate as WT cells. The PCDH19-KO ESCs also proliferated at the same rate as WT cells, differently from what was previously shown with the *Pcdh19* KO ESCs (Dr. Jessica Griffiths, unpublished), confirming the importance of the genetic background. Unlike PCDH19-CYTO mESC-derived neurons, PCDH19-KO mESC-derived neurons had previously been generated in the Martinez-Garay lab, following the same differentiation protocol (Dr. Jessica Griffiths, unpublished), hence it was expected that the CRISPR/Cas9 generated PCDH19-KO ESCs could also be differentiated into cortical-like neurons. Moreover, both PCDH19 KO mouse models and male carriers of *PCDH19* mutations show no defects.

It is important to note that both PCDH19-CYTO and PCDH19-KO ESCs were passaged on MEFs in order to improve the quality of the neuronal differentiation, decrease numbers of non-neuronal cells and consequently increase homogeneity of the neuronal cultures. MEFs provide support for ESCs by secreting growth factors that are important for maintenance of pluripotency and by providing physical attachment via cell-cell and cell-matrix contacts. For instance MEFs produce and secrete different types of extracellular matrix components, such as laminin, collagens an fibronectins (Hongisto et al., 2012). MEFs also secrete transforming growth factor beta (TGFß), activin A (Eiselleova et al., 2008) and Leukemia Inhibitory Factor (LIF). In fact, LIF is essential for ESC culture and is also supplemented in the ESC-medium as it is crucial for maintaining pluripotency of mESCs, via the JAK/STAT pathway (Williams et al., 1988). Alternatively to culturing on feeder cells, quality of ESC clones could have been improved by testing culturing of ESCs in different types of medium, such as, for example, 2I+LIF medium, which includes inhibitors of GSK-3 and MAPK/ERK which inhibit differentiation and promote self-renewal (Ying and Smith, 2017).

It is possible that culturing ESC on MEFs could have had long-lasting effects on the transcriptome of the ESCs and, retrospectively, WT ESCs, although feeder-independent and of good quality, should have been cultured on MEFs for an equivalent number of passages, in parallel, as they were used at controls for RNA sequencing (**Chapter 6**).

5.3.1 Conclusion

This chapter outlines the successful generation and validation of an *in vitro* model to study the function of the cytoplasmic domain of PCDH19. This *in vitro* model consists of E14 mESCs overexpressing the cytoplasmic domain from the *Rosa26* locus and isogenic PCDH19 KO mESCs. These cells were able to generate cortical-like neurons, which were used, as will be discussed in the next chapter, for RNA sequencing analysis at different time-points during neuronal differentiation.

Chapter 6

Transcriptome analysis of PCDH19-CYTO and PCDH19-KO cells

6.1 Introduction

Cadherins have important signalling functions that are mediated by their proteolytically processed intracellular domains. For instance, N-Cad/CTF2, the intracellular fragment of N-Cadherin that is generated via γ-secretase cleavage, is a transcriptional repressor of CBP/CREB mediated transcription. N-Cad/CTF2 can physically bind CBP in the cytoplasm to promote its degradation via the ubiquitin pathway, hence reducing its levels in the nucleus and consequently decreasing CBP/CREB mediated transcription (Marambaud et al., 2003). CREB, cyclic AMP response element binding protein, regulates expression of genes containing cyclic AMP response elements in their promoters, such as the immediate early-gene *c-Fos*. Therefore, in neurons, CREB is involved in mediating synaptic plasticity and activity-driven γ-secretase cleavage of N-Cadherin keeps levels of these genes in check.

E-Cad/CTF2, the cytoplasmic fragment of E-Cadherin, which is also generated by γ-secretase cleavage, has a nuclear function in combination with the cadherin-binding protein p120-catenin, which enhances its nuclear translocation. In the nucleus, E-Cad/CTF2 can regulate p120/Kaiso transcription (Ferber et al., 2008). Moreover, γ-PCDH ICDs can translocate to the nucleus and induce gene transcription at the γ-*Pcdh* locus, in an auto-regulatory manner (Hambsch et al., 2005).

Finally, although a proteolytic processing of PCDH19 has not been reported before, PCDH19 is also known to have a nuclear function. In fact, PCDH19 can interact with the paraspeckle protein NONO in order to modulate ER-α-dependent gene expression. Whilst this study was carried out in heterologous cells and using a candidate approach, it confirmed via overexpression of full-length PCDH19 that PCDH19 can modulate expression of the ER-α-dependent genes *AKR1C3*, *APOD*, *ENC1* and *OXTR*. Interestingly, the effect was lost when using the mutant form of PCDH19 Asn557Lys, with a single amino acid change in the extracellular domain (Pham et al., 2017), suggesting that a functional full-length PCDH19 capable of adhesion is necessary to induce these transcriptional effects.

The forementioned genes, *AKR1C3*, *APOD*, *ENC1* and *OXTR*, are also dysregulated in fibroblasts derived from *PCDH19*-epilepsy patients, which also have abnormal neurosteroid levels. It remains unclear if this phenotype is causative or consequence of the disease, as neurosteroids are potent modulators of neuronal activity and altered levels can trigger epileptic episodes. In particular, patients have reduced levels of allopregnanolone, a potent positive allosteric modulator of GABA-A receptors (Tan et al., 2015), and corresponding decreased levels of the neurosteroid metabolising enzyme *AKR1C3*, which is responsible for the conversion of progesterone into allopregnanolone. Many of the dysregulated genes are regulated by chorionic gonadotropin (Cg), progesterone and estrogen through their respective receptors: progesterone receptor (PGR) and estrogen receptor alpha (ESR1) (Trivisano et al., 2017).

Transcriptome analysis has been performed, via microarray, on PCDH19 KO cells derived from the *Pcdh19* KO mouse (Homan et al., 2018). For this experiment, progenitors were extracted at E14.5, grown into neurospheres and differentiated into cortical neurons *in vitro*. *Pcdh19* KO, WT and mixed cultures were compared. Differentially expressed genes were linked to cell-fate decision, cell adhesion, neuronal maturation and differentiation. Interestingly, they were also linked to the estrogen signalling pathway but it was not specified if this pathway was upregulated or downregulated (Homan et al., 2018).

In this thesis it has been demonstrated that PCDH19 can be proteolytically processed with the resulting generation of at least 2 intracellular C-terminal fragments, CTF1 and CTF2 (**Chapter 3**). It has also been determined that CTF2, which is not attached to the membrane and can therefore readily translocate to other cellular compartments, is the most prominent form in mouse ESC-derived neurons, suggesting a neuronal specific function for this fragment of PCDH19. Moreover, the mass spectrometry analysis in (**Chapter 4**), revealed several nuclear import transport proteins as potential interactors of PCDH19 and one of these, KPNA1, was validated through *in vitro* experiments. These results suggest that PCDH19 can be transported into the nucleus, via interactions with importins. In (**Chapter 5**) an *in vitro* model was established to investigate the role of the PCDH19-CYTO fragment and to determine if it could be involved in modulating gene expression. RNA sequencing analysis was the method chosen for this purpose, to obtain a full picture of the transcriptome of PCDH19 KO and PCDH19 CYTO cells in an unbiased manner, and to allow identification of differentially expressed transcripts between the cell lines at different time points during neuronal differentiation.

6.1.1 Aims

This chapter aims to investigate the transcriptional profile of neuronal progenitor cells and neurons derived from the PCDH19-CYTO and PCDH19-KO mouse ESCs generated in (**Chapter 5**), by conduction of a whole transcriptome RNA-sequencing analysis. The aim is to uncover potential

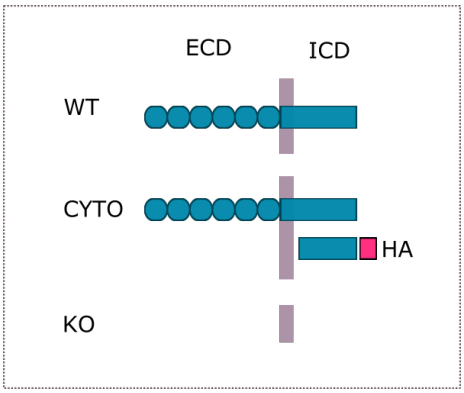
transcriptional changes arising from the overexpression of the cytoplasmic domain of PCDH19 or the complete loss of the protein, in order to begin to understand the physiological consequences of PCDH19 processing in a neuronal system.

6.2 Results

6.2.1 Experimental set up and quality control

PCDH19-CYTO and PCDH19-KO ESCs (**Figure 6.1a**) derived in the previous chapter (**Chapter 5**) were differentiated alongside WT ESCs into cortical-like neurons following a previously described protocol (**Section 2.4**) (Bibel et al., 2004). RNA was extracted at 3 timepoints, CA8, DIV8 and DIV12, during three independent differentiations from each of the three cell lines (**Figure 6.1b**). Each individual RNA sample was derived from the pooling of 3 wells (750' 000 cells/well), in order to reduce potential artefacts due to individual well variability (such as small differences in cell density). CA8 samples were collected at the end of the aggregation phase without dissociation for plating of progenitors, so as to minimise stress to the cells that could result in transcriptomic changes. Because the samples were obtained from independent ESC differentiations, this experimental design compares biological, rather than technical replicates. This design was chosen because, despite the increase in variability, it reduces the overall effect of possible differentiation-dependent biases on the results and increases the chances of identifying cell line specific phenotypes. RNA extraction was performed in 3 batches of 9 samples, each including the WT, KO and CYTO samples from one timepoint. This was done to reduce extraction time, preserve high RNA quality and to limit variability across groups. All samples were checked for RNA integrity and the RIN score was calculated with a Tapestation. All samples scored 8.0 and above, in line with the Illumina recommendations for sequencing (**Figure 6.2**). Sample concentration was determined with the QUBIT and samples were sequenced by the Genomics Hub of the School of Biosciences, as described (**Section 2.13**). The raw FASTQ files were quality checked via the FASTQC programme. Sequence quality was checked via Mean Quality Scores (Phred Scores) and all sequences had an average score above 30 (= more than 99.9 base call accuracy). Sequences were also checked for CG content (49-51%) and for duplication levels across library (50%). All samples passed quality controls and were therefore used in downstream analysis. Sequences were trimmed and aligned to reference genome and counts were calculated as described (**Section 2.13**).

(a)



(b)

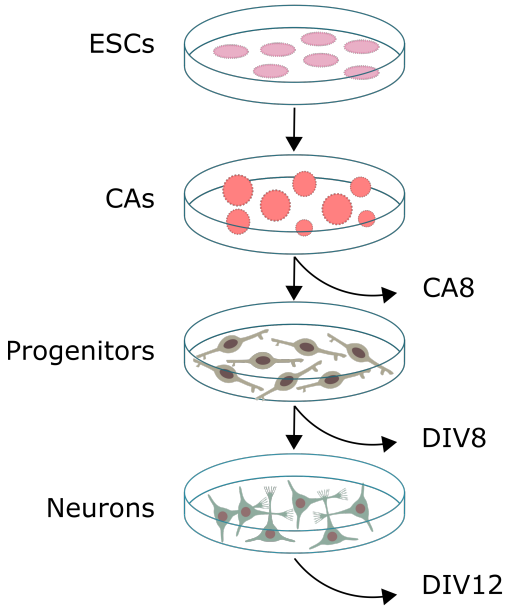
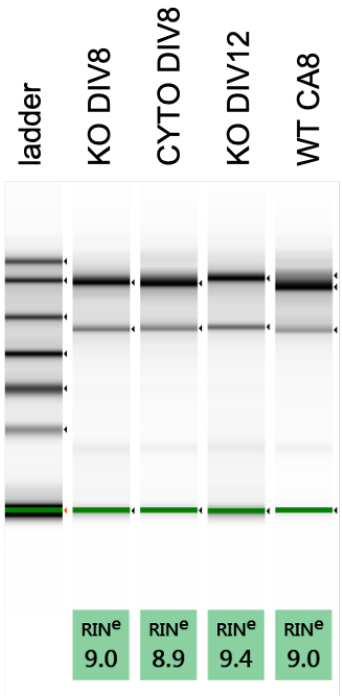


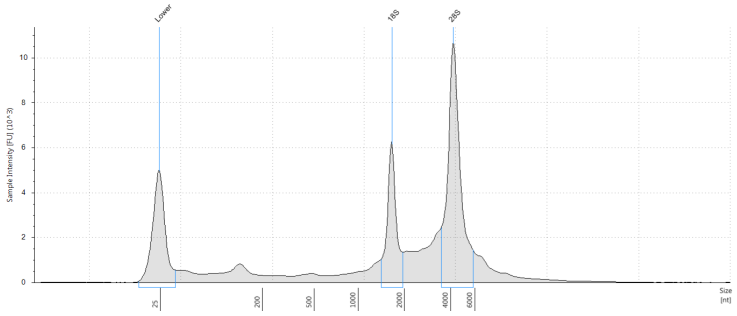
Figure 6.1: RNA-seq experimental design. (a) Schematic of PCDH19 protein expression in WT, CYTO and KO mESCs used for RNA sequencing. **(b)** Extraction time points for RNA sequencing: samples were collected for WT, CYTO and KO cells during neuronal differentiation at CA8 (progenitors), DIV8 (immature neurons) and DIV12 (mature neurons).

(a)



(b)

D1: KO DIV12



Sample Table

Well	RINe	28S/18S (Area)	Conc. [ng/µl]	Sample Description	Alert	Observations
D1	9.4	2.4	127	KO DIV12		

Peak Table

Size [nt]	Calibrated Conc. [ng/µl]	Assigned Conc. [ng/µl]	Peak Molarity [nmol/l]	% Integrated Area	Peak Comment	Observations
25	40.0	40.0	4710	-		Lower Marker
1673	20.3	-	35.6	29.49		18S
4227	48.5	-	33.7	70.51		28S

Figure 6.2: RNA samples passed RNA Integrity Number (RIN) quality control. (a) Tapestation gel example with 4 representative RNA samples and respective RINs. **(b)** KO DIV12 sample as example. RIN score is calculated from the 28S/18S RNA peaks.

6.2.2 Principal Component Analysis and clustering

The obtained RNAseq data were initially subjected to a Principal Component Analysis (PCA) on the whole transcriptome of all samples. PCA on all samples showed that Principal Component 1 (PC1) accounted for 92% of the variance and PC2 for 2%, highlighting major transcriptional differences between progenitors and neurons, regardless of cell line (WT, KO or CYTO) (**Figure 6.3**). Because of the magnitude of this difference, PCA was then repeated separately for progenitors and neurons in order to visualise differences due to genotype as opposed to differentiation state. At CA8, PC1 accounted for 47% of the variance and PC2 accounted for 23% of the variance (**Figure 6.3.b**). For neurons (DIV8 + DIV12) PC1 accounted for 31% of the variance and PC2 accounted for 23% (**Figure 6.3.c**). Cell-line differences are not pronounced based on whole transcriptome changes and samples only cluster minimally showing that differentiation procedure does increase variability. High variability between samples of the same cell-line and time-point could be the result of biological variability as each sample was derived from a completely separate neuronal differentiation procedure. This was done to reduce false positives that could have arisen from the analysis of technical replicates. Because of the big difference between progenitor and neuronal samples overall, hierarchical clustering of samples was done separately for progenitors and neurons. Although samples belonging to the same cell line and from the same time-point do not cluster perfectly together, for progenitors, there is a clear separation between WT and CYTO samples. Also CYTO and KO samples cluster together, except for a single KO sample. In the neuronal samples, CYTO and WT again separate from each other, both at DIV8 and DIV12, whilst the separation between KO and WT is less evident. In particular, at each time point there is one KO sample which clusters with the WT samples, and all three (CA8, DIV8 and DIV12) belong to the same KO differentiation (KO differentiation number 1) . For neurons, there also are two WT samples, one DIV8 and one DIV12 that cluster closer to CYTO and KO. In this case they also belong to the same differentiation (WT differentiation number 1) (**Figure 6.4**). These observations strengthen the need for biological over technical replicates.

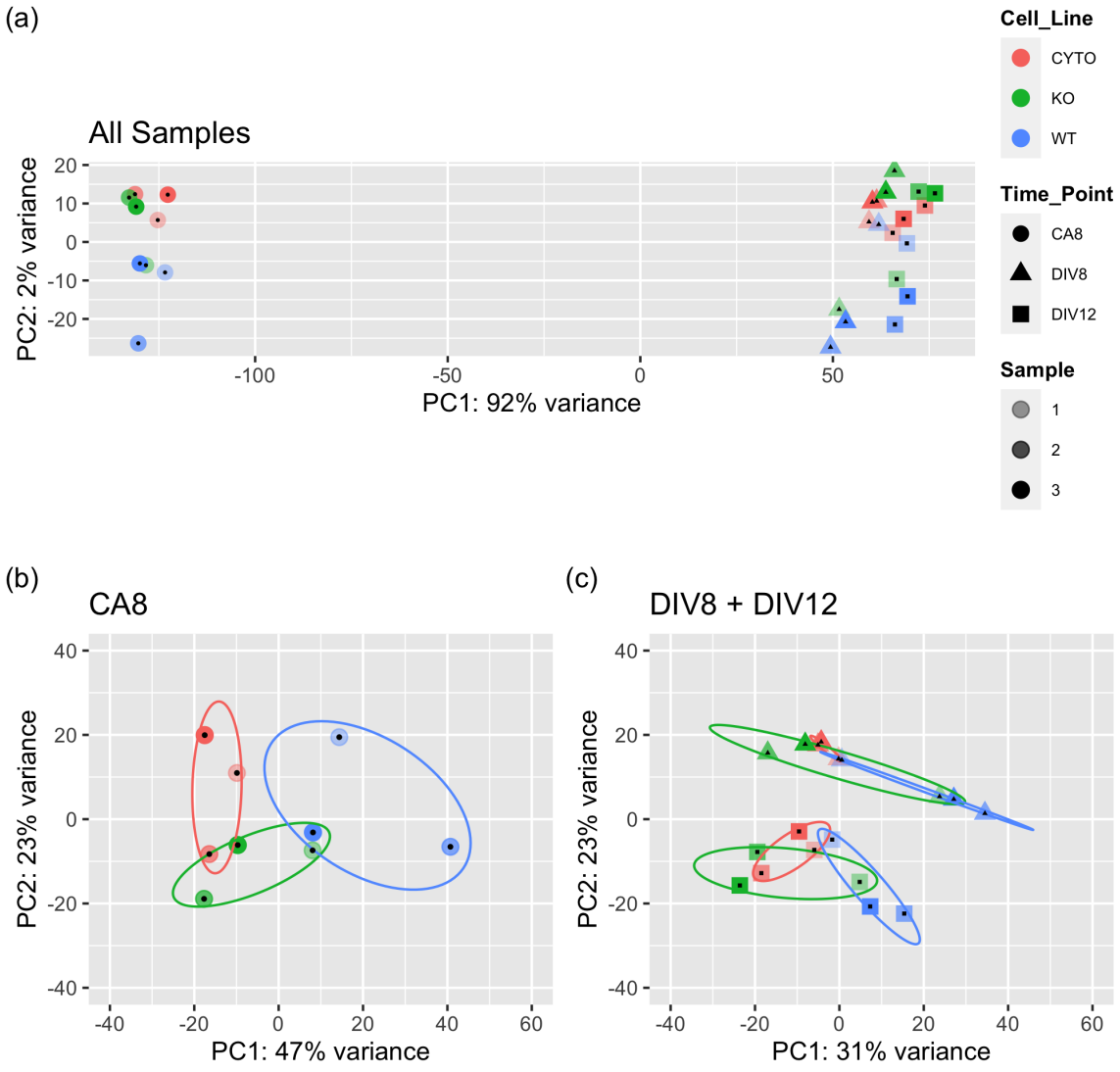
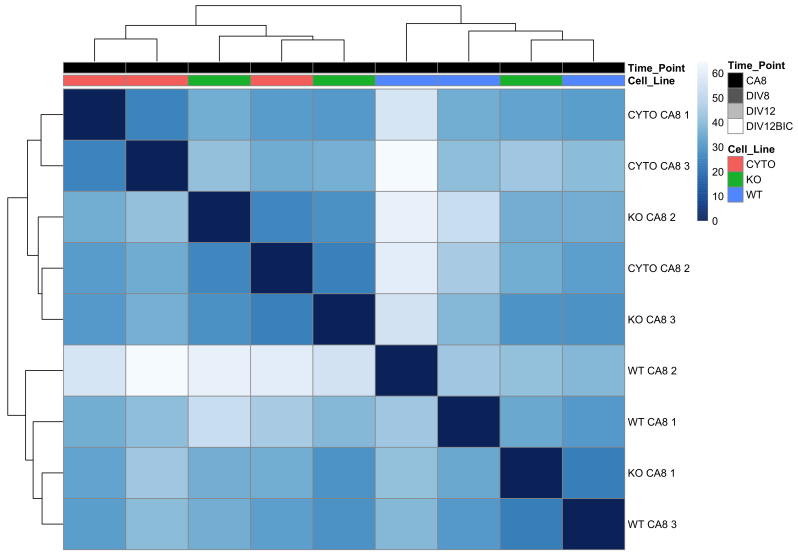


Figure 6.3: Principal Component Analysis (PCA). (a) PCA on all samples. **(b)** PCA on progenitors (CA8). **(c)** PCA on neuronal samples (DIV8 and DIV12). Shapes indicate time points (CA8, DIV8, DIV12), colours indicate cell-lines (CYTO, KO, WT) and alpha scale indicates samples belonging to the same differentiation (1, 2, 3).

(a)



(b)

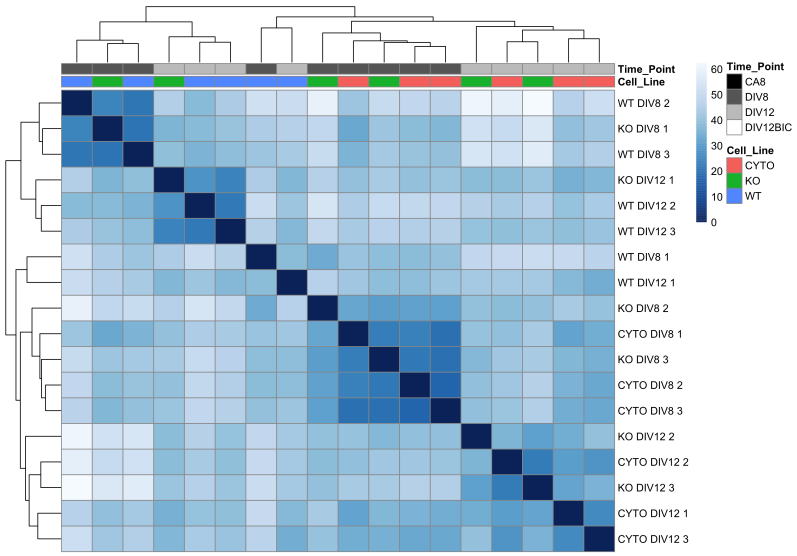
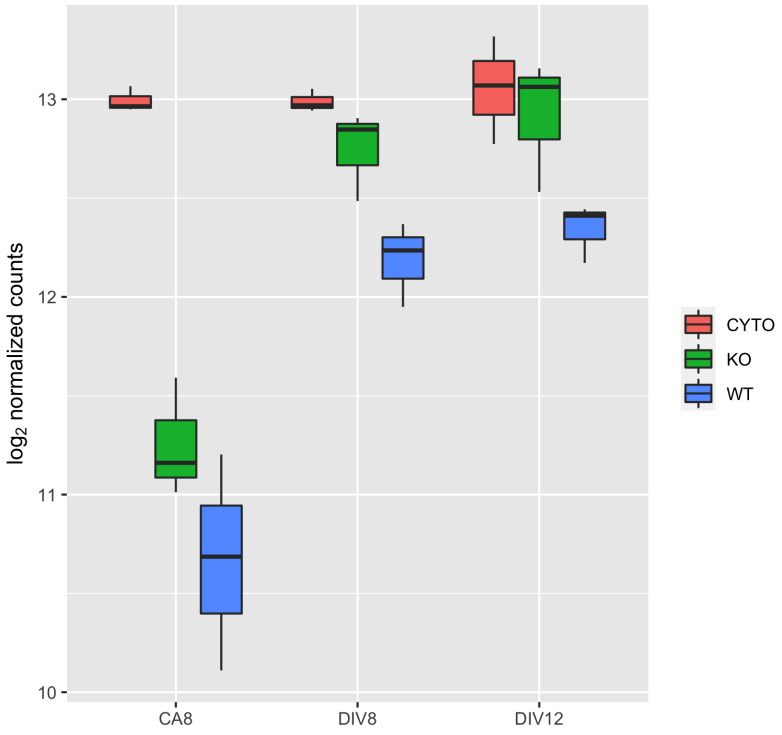


Figure 6.4: Correlation matrix showing hierarchical clustering of samples based on expression of all genes. (a) Clustering of progenitor samples (CA8). **(b)** Clustering of neuronal samples (DIV8 + DIV12). Blue scale represents distance between samples, with darker blue corresponding to smaller distances.

6.2.3 *Pcdh19* expression

As a further way to verify the validity of the KO and CYTO ESC lines generated in (**Chapter 5**), RNAseq data were checked for expression of *Pcdh19*. *Pcdh19* expression was analysed by comparing normalised counts for all time-points and cell lines. As expected, in CYTO cells *Pcdh19* mRNA is significantly upregulated (CYTO *vs* WT: CA8, p.adj = 9.22×10^{-28}; DIV8, p.adj = 0.00213; DIV12, p.adj = 0.00626) (**Figure 6.5b**). *Pcdh19* overexpression is especially evident at the progenitor stage, CA8, as endogenous *Pcdh19* is not strongly expressed at this stage in WT cells (**Figure 6.5a**). At DIV8 and DIV12 expression of endogenous *Pcdh19* in WT neurons is increasing, therefore the difference between CYTO and WT is less pronounced. As RNAseq was paired-end, it was possible to conduct a splicing analysis, which was done via DEXSeq2 (Anders et al., 2012) (**Section 2.13**). Analysis of *Pcdh19* splicing shows differential exon usage of *Pcdh19* in CYTO *vs* WT (**Figure 6.6a**) but not in KO *vs* WT (**Figure 6.6b**), due to the overexpression of the cytoplasmic domain, encoded by exons 2-6 (**Figure 6.6c**). This is evident at all 3 time-points, confirming expression of transgenic PCDH19-CYTO-HA. When comparing normalised counts between CYTO and WT for each exon bin, designated by an E, there is a significant upregulation in usage of E013-10 and E008-7 which correspond to exons 2, 3, 4, 5 and 6 – the whole of the intracellular domain, and significant downregulation in E017-14, corresponding to exon 1. E001-6 correspond to the 3' UTR of PCDH19 whilst E009 corresponds to 3 bp belonging to a less common *Pcdh19* isoform (**Figure 6.6a**) (**Figure 6.6c**) that were not included in the targeting vector. *Pcdh19* is "differentially spliced" in CYTO compared to WT at CA8, DIV8 and DIV12, confirming expression of PCDH19-ICD across all 3 time-points (data shown for CA8). Importantly, and despite the lack of protein in ES cells and differentiated neurons (**Figure 5.7**), *Pcdh19* mRNA is present in KO cells and its levels follow a developmental trajectory similar to WT *Pcdh19*, increasing from CA8 to DIV12 (KO *vs* WT: CA8, p.adj= 0.0936; DIV8, p.adj= 0.153; DIV12, p.adj= 0.0861) (**Figure 6.5b**). Therefore, it seems that *Pcdh19* KO mRNA is escaping non-sense mediated decay.

(a)

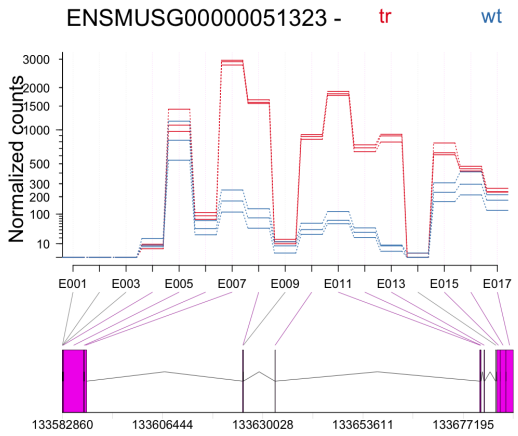


(b)

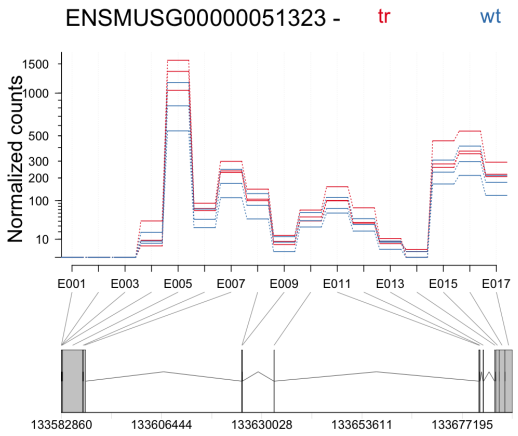
Pcdh19	CA8		DIV8		DIV12	
	$\log_2$ FC	p.adj	$\log_2$ FC	p.adj	$\log_2$ FC	p.adj
CYTOvsWT	2.26	9.22×10^{-28}	0.795	0.00213	0.724	0.00626
KOvsWT	0.542	0.0936	0.563	0.153	0.595	0.0861
CYTOvsKO	1.72	1.99×10^{-15}	0.232	1	0.129	0.838

Figure 6.5: *Pcdh19* expression in CYTO, KO and WT cells. (a) Normalised counts for *Pcdh19* for CYTO, KO and WT cells at CA8, DIV8 and DIV12. **(b)** *Pcdh19* $\log_2$ fold change and p.adjusted values for CYTO *vs* WT, KO *vs* WT and CYTO *vs* KO at CA8, DIV8 and DIV12.

(a)



(b)



(c)

ID	Start	End	Width	Exon	CA8 CYTO *vs* WT CA8	
					$\log_2$ FC	p.adj
E001	133582860	133582874	15	3' UTR	NA	NA
E002	133582875	133582878	4	3' UTR	NA	NA
E003	133582879	133582879	1	3' UTR	NA	NA
E004	133582880	133583104	225	3' UTR	-3.62	2.42×10^{-7}
E005	133583105	133587901	4797	3' UTR	-2.49	3.32×10^{-83}
E006	133587902	133588011	110	3' UTR	-2	1.16×10^{-12}
E007	133588012	133588577	566	6	1.07	1.23×10^{-97}
E008	133625269	133625438	170	5	1.29	4.25×10^{-103}
E009	133625439	133625441	3	5	-2.04	0.36
E010	133632895	133632953	59	4	1.38	6.85×10^{-65}
E011	133681081	133681319	239	3	1.53	1.18×10^{-137}
E012	133681320	133681408	89	3	1.45	1.15×10^{-53}
E013	133682092	133682232	141	2	4.31	2.56×10^{-149}
E014	133684817	133684991	175	1	-2.97	NA
E015	133685118	133685893	776	1	-1.43	1.69×10^{-22}
E016	133685894	133687264	1371	1	-2.41	3.6×10^{-171}
E017	133687265	133688987	1723	1	-2.43	1.28×10^{-93}

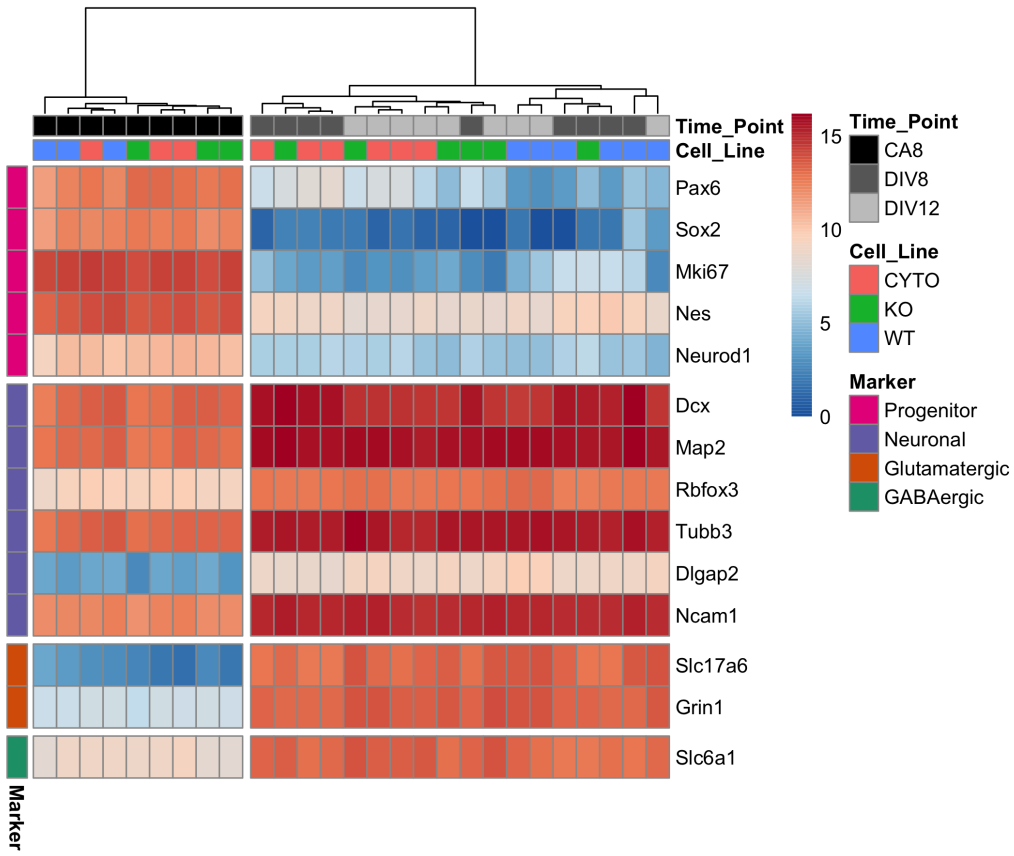
Figure 6.6: *Pcdh19* exon usage. **(a)** CYTO *vs* WT at CA8. **(b)** KO *vs* WT at CA8. Normalised counts plotted along genomic coordinates of *Pcdh19*. Purple exons are differentially expressed, grey exons are not. The red line represents counts for the transgenic lines (CYTO **(a)**, KO **(b)**), the blue line represents counts for WT. **(c)** Genomic coordinates of exon bins, their size in base pairs and corresponding *Pcdh19* exons. p.adjusted values and $\log_2$ fold change are reported for CYTO *vs* WT at CA8. NA means comparison was not applicable due to very low number of counts/ no counts.

6.2.4 Expression of key marker genes

Analysis of expression of well-established marker genes through the differentiation time-course can give a comprehensive overview of the cells and be very useful to determine if cells of the different genotypes behaved as expected during the differentiation procedure. Markers of neuronal progenitor cells such as *Sox2*, *Ki-67*, *Pax6*, *Neurod1* and *Nestin* were upregulated at CA8 compared to DIV8 and DIV12 across CYTO, KO and WT samples (**Figure 6.7**). Neuronal markers such as Neural cell adhesion molecule 1 (*Ncam1*), doublecortin (*Dcx*), Microtubule Associate protein 2 (*Map2*), βIII-tubulin (*Tubb3*) and NeuN (*Rbfox3*), were instead upregulated at DIV8 and DIV12 compared to CA8 (**Figure 6.7**). This suggests that all three cell types, CYTO, KO and WT are able to follow through neuronal differentiation and mature appropriately from progenitor stage to neurons. At DIV8 and DIV12 there is also an upregulation of the GABA transporter 1 (*Slc6a1*), the vesicular glutamate transporter 2 (*Slc17a6*) and the ionotropic glutamate receptor NMDA1 (*Grin1*), indicating that the cells differentiate into both excitatory and inhibitory neurons. As stated, the differentiation protocol used in this thesis generates cortical-like neurons. However, whilst the vast majority of generated neurons (95%) will be glutamatergic excitatory neurons, this protocol also generates about 5% of GABAergic interneurons, that could be identified in the cultures by presence of the GABA transporter (vGAT) (Bibel et al., 2004). It was also interesting to determine if neurons express different cortical layer markers and if those are expressed to similar levels across the three cell-lines (**Figure 6.8a**). Remarkably, expression of cortical layer markers was not always homogeneous across genotypes. The strongest expressed markers were *Cux1* and *Cux2*, suggesting a skewed differentiation towards upper-layer neurons, which is common to WT, CYTO and KO. However, other markers were differentially expressed. For instance, at DIV8 and DIV12 *Foxp2* was significantly up-regulated in CYTO and in KO compared to WT neurons (CYTO *vs* WT: DIV8, p.adj = 1.08×10^{-16}; DIV12, p.adj = 1.01×10^{-21}; KO *vs* WT: DIV8, p.adj = 8.06×10^{-11}11; DIV12, p.adj = 1.13×10^{-12}) (**Figure 6.8b**). *Foxp2* is a marker for deep-layer neurons and is predominantly expressed by layer 6 neurons but can also be found in layer 5 (Kast et al., 2019). Interestingly, *Bcl11b*, which is also known as *Ctip2*, a cortical layer marker for subcortically

projecting deep layer neurons is downregulated in CYTO *vs* WT at DIV8 and DIV12 (CYTO *vs* WT: DIV8, p.adj= 1.69×10^{-6}; DIV12, p.adj= 9.74×10^{-5}; KO *vs* WT: DIV8, p.adj = 0.0161; DIV12, p.adj = 0.0631) (**Figure 6.8b**). *Rorb*, the classical layer 4 marker, is also downregulated in CYTO *vs* WT at DIV8 (CYTO *vs* WT: DIV8, p.adj = 1.24×10^{-7}) (**Figure 6.8b**). Surprisingly, although *Pax6* is differentially expressed between progenitors and neurons across genotypes, as expected, it is also differentially expressed in CYTO neurons compared to WT and to a lesser extent in KO (CYTO *vs* WT: DIV8, p.adj = 1×10^{-6}; DIV12, p.adj = 6.21×10^{-5}; KO *vs* WT: DIV8, p.adj = 0.0228) (**Figure 6.7b**) suggesting a potential delay in maturation of targeted neurons.

(a)

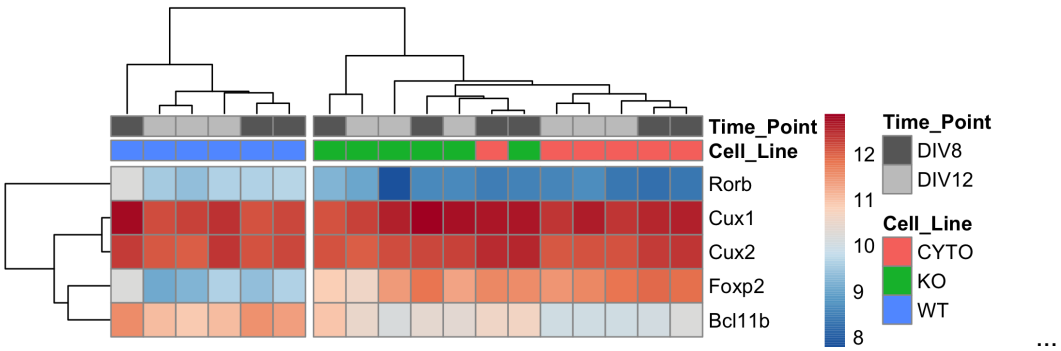


(b)

Pax6	CA8		DIV8		DIV12	
	$\log_2$ FC	p.adj	$\log_2$ FC	p.adj	$\log_2$ FC	p.adj
CYTOvsWT	0.943	0.391	3.59	1×10^{-06}	3.17	6.21×10^{-05}
KOvsWT	0.888	0.547	2.28	0.0228	2.01	0.068
CYTOvsKO	0.0557	0.979	1.32	1	1.16	0.401

Figure 6.7: Expression of selected neuronal marker genes. **(a)** Progenitor and neuronal marker genes show comparable expression pattern in CYTO, KO and WT cells, except for *Pax6*. Colour scale represents $\log_2$ fold change. **(b)** $\log_2$ fold change and p.adjusted values for*Pax6* for different comparisons at CA8, DIV8, DIV12. Comparisons with p. adjusted values < 0.001 are highlighted in grey.

(a)



...

(b)

Foxp2	CA8		DIV8		DIV12	
	$\log_2$ FC	p.adj	$\log_2$ FC	p.adj	$\log_2$ FC	p.adj
CYTOvsWT	-0.182	0.75	2.04	1.08×10^{-16}	2.32	1.01×10^{-21}
KOvsWT	-0.354	0.521	1.7	8.06×10^{-11}	1.83	1.13×10^{-12}
CYTOvsKO	0.172	0.752	0.332	1	0.487	0.324

Bcl11b	CA8		DIV8		DIV12	
	$\log_2$ FC	p.adj	$\log_2$ FC	p.adj	$\log_2$ FC	p.adj
CYTOvsWT	-0.431	0.223	-1.18	1.69×10^{-06}	-1.01	9.74×10^{-05}
KOvsWT	-0.617	0.068	-0.78	0.0161	-0.658	0.0631
CYTOvsKO	0.186	0.708	-0.401	1	-0.353	0.482

Rorb	CA8		DIV8		DIV12	
	$\log_2$ FC	p.adj	$\log_2$ FC	p.adj	$\log_2$ FC	p.adj
CYTOvsWT	-0.00407	1	-1.49	1.24×10^{-07}	-0.977	0.00227
KOvsWT	-0.462	0.372	-1.03	0.00339	-0.961	0.00765
CYTOvsKO	0.458	0.317	-0.466	1	-0.0166	0.989

Figure 6.8: Expression of selected cortical marker genes. (a) Expression of some cortical layer markers is different in CYTO, KO and WT cells. Colour scale represents $\log_2$ fold change. **(b)** $\log_2$ fold change and p.adjusted values for markers which are differentially expressed. Comparisons with p. adjusted values < 0.001 are highlighted in grey.

6.2.5 Differential Expression Analysis

To determine what are the most consistently up and down-regulated genes in CYTO and KO progenitors and neurons compared to WT, differential expression analysis was carried out at each time point (CA8, DIV8, DIV12) comparing each genotype (CYTO *vs* WT, KO *vs* WT and CYTO *vs* KO). The threshold for significance is arbitrary and for the following analysis was placed at p. adjusted $<$ 0.001 and $\log_2$ fold change cutoff $>$ 0.58, which corresponds to a fold change of 1.5. Results of differential expression analysis with less stringent p.adjusted cut-off ($<$ 0.01 and $<$ 0.05) and constant $\log_2$ fold change $>$ 0.58 are provided in the appendix (**Section 8.6**), (**Section 8.7**). Differentially expressed genes (DEG) are easily visualized in vulcano plots plotting $\log_2$ fold change *vs* -$\log_{10}$ p. adjusted value (**Figure 8.6**).

The biggest number of differentially expressed genes (DEG) is found at CA8, the progenitor stage, for both CYTO *vs* WT (101 up, 297 down) and KO *vs* WT (35 up, 132 down), with most of the differentially expressed genes being downregulated. At DIV8, the distribution of upregulated and downregulated genes is more even for CYTO *vs* WT (106 up, 143 down) but not for KO *vs* WT (39 up, 15 down). Finally, DIV12 is very similar to DIV8 for CYTO *vs* WT (115 up, 113 down) and KO *vs* WT (38 up 19 down). Not many genes are differentially expressed in CYTO *vs* KO (CA8: 11 up, 2 down; DIV8: 6 up, 0 down; DIV12: 7 up, 6 down) (**Figure 8.7a**).

Interestingly, comparing lists of DEGs of different analyses reveals that many of the genes that are changed in CYTO *vs* WT are also changed in KO *vs* WT across all timepoints (**Figure 8.7b**). This substantial overlap is also reflected by the small number of differentially expressed genes in CYTO *vs* KO. At CA8, 30.1% of CYTO *vs* WT are also KO *vs* WT DEGs, and 71.7% of KO *vs* WT are also significant CYTO *vs* WT genes. At DIV8, 20.6% of CYTO *vs* WT genes are also differentially expressed in KO *vs* WT, and 94.4% of KO *vs* WT are also significant in CYTO *vs* WT. At DIV12, 20.0% of CYTO *vs* WT DEGs are also common to KO *vs* WT, and 81.9% of KO *vs* WT are also significant in CYTO *vs* WT (**Figure 8.7b**). Most of the overlapping genes between

CYTO *vs* WT and KO *vs* WT are differentially expressed in the same direction. The proportion of upregulated/downregulated and overlapping DEGs across comparisons is comparable at different p.adjusted value thresholds (< 0.01 and < 0.05) (**Section 8.6**), (**Section 8.7**).

All the DEGs for the different comparisons are listed in the appendix: CYTO *vs* KO (**Appendix 8.3.3**), CYTO *vs* WT (**Appendix 8.3.1**) and KO *vs* WT (**Appendix 8.3.2**) including p.adjusted values and $\log_2$ fold change. An overview of the top 20 genes, based on p. adjusted value for each comparison, or less if there are fewer than 20 differentially expressed genes, is shown below (**Figure 6.11**). For simplicity, not all genes will be discussed and the following sections will focus on differentially expressed genes which are specific to CYTO and of interest for reasons explained below.

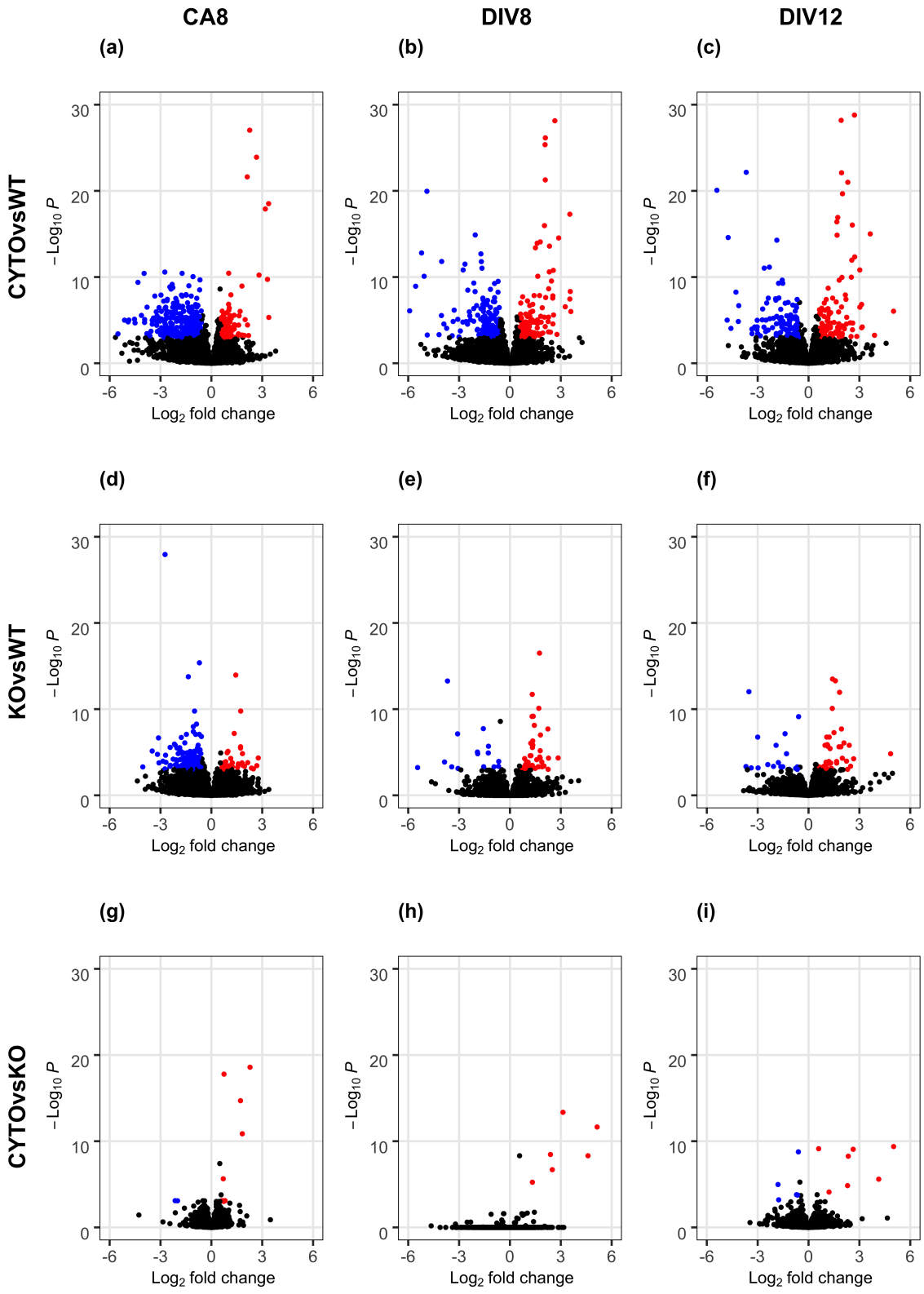


Figure 6.9: Vulcano plots showing differentially expressed genes for CYTO *vs* WT, KO *vs* WT and CYTO *vs* KO at CA8, DIV8 and DIV12. Downregulated genes are shown in blue, upregulated genes are shown in red. Threshold for significance was p.adj < 0.001 and $\log_2$ cutoff > 0.58.

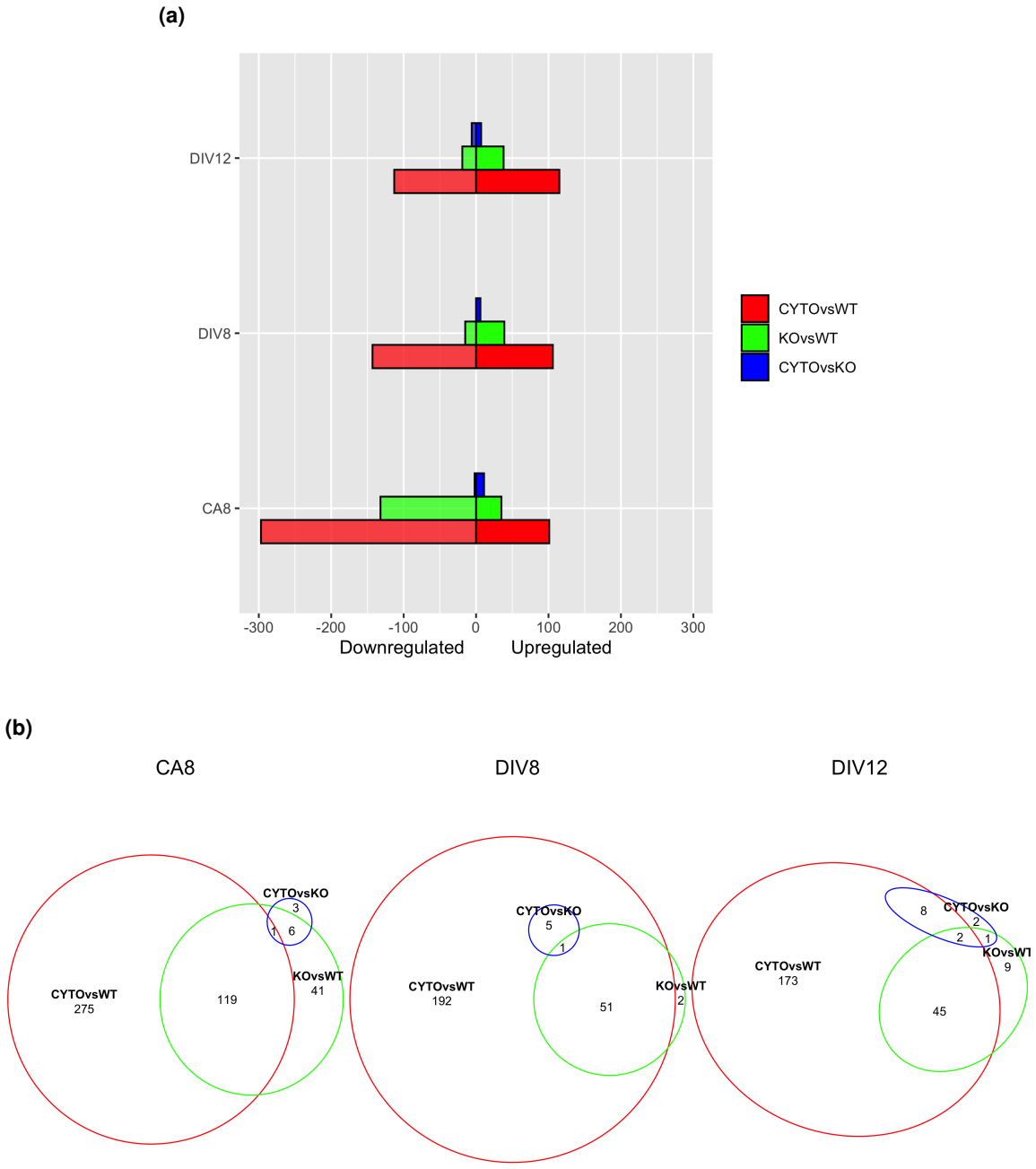


Figure 6.10: Number of differentially expressed genes. (a) Total upregulated and downregulated number of genes at each time point for CYTO *vs* WT, KO *vs* WT and CYTO *vs* KO. **(b)** Venn diagram showing common differentially expressed genes between CYTO, KO and WT cells at time points CA8, DIV8 and DIV12, at a significance threshold of p.adj < 0.001 and $\log_2$ cutoff > 0.58.

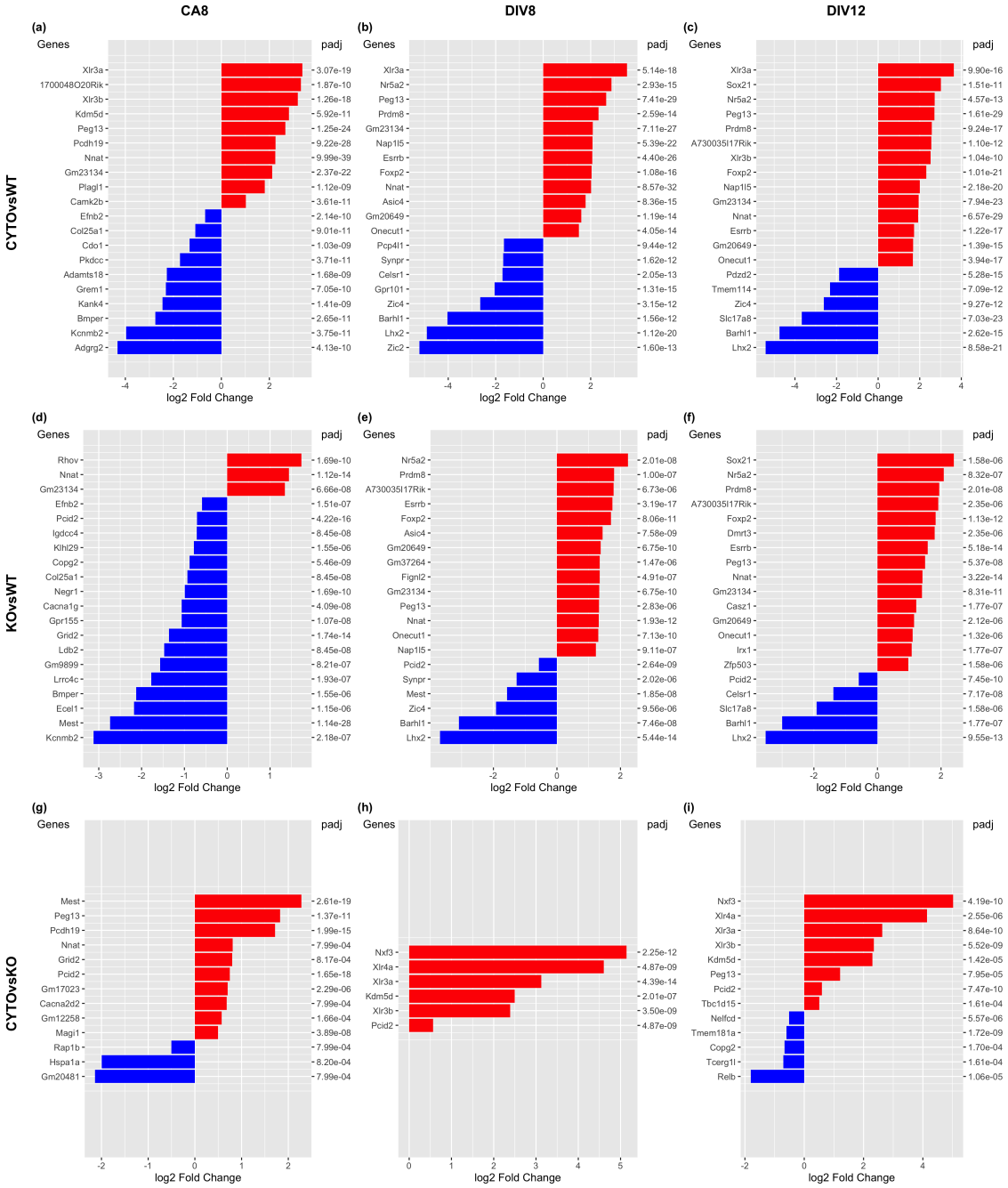


Figure 6.11: Top differentially expressed genes. Top 20 DEGs for each comparison (CYTO *vs* WT, KO *vs* WT, CYTO *vs* KO) at each time point (CA8, DIV8, DIV12), based on p.adjusted value.

6.2.6 Enrichment Analysis with DEGs

Differentially expressed genes resulting from the different comparisons presented above were subjected to an over-representation or enrichment analysis using the "Cluster Profiler" package on R (Yu et al., 2012). Over-representation analysis determines if genes belonging to a predetermined gene set are over-represented in a dataset. A gene set is a predefined list of genes that belongs to the same pathway or is associated with a particular biological process, for example a list of genes that is associated with a particular Gene Ontology (GO) term. For this analysis, DEGs were analysed for over-representation of GO terms. Analysis outputted more than 1000 terms across comparisons (**Figure 6.12a**). For simplicity, results were filtered by elimination of terms belonging to non relevant organ systems. Secondly, redundant terms were eliminated using the R package "GO.db" (Carlson, 2019), a set of annotation maps describing the entire Gene Ontology, which allows elimination of offspring terms. Filtered and simplified results of the analysis are summarised in (**Figure 6.12b**) and presented in the Appendix (**Appendix 8.4**). The majority of enriched terms appear at CA8 for both CYTO *vs* WT and KO *vs* WT. Selected significant terms of relevance are shown in (**Figure 6.12c**). Interestingly, several terms related to axon and dendrite morphogenesis resulted enriched. In particular, the terms "*dendrite development*", "*neuron projection arborization*", "*axon guidance*", "*neuron projection guidance*" and "*axonogenesis*" were identified at CA8 for CYTO *vs* WT and KO *vs* WT. Several axon guidance molecules are in fact differentially expressed and belong to these enriched terms. Amongst them are for instance several semaphorins (*Sema3a*, *Sema5a*, *Sema3c*, *Sema6a*) and netrin receptors (*Unc5a*, *Unc5c*, *Unc5d*) and netrin itself (*Ntn1*). *Unc5s* are dependence receptors which are required for neuronal survival when associated with netrin. The mentioned genes are differentially expressed in both directions, as shown in (**Figure 6.13a**). Although there is a substantial overlap between CYTO *vs* WT and KO *vs* WT in DEGs belonging to these terms, some genes are found only in CYTO *vs* WT, such as *Xlr3b*, *Unc5d*, *Unc5a*, *Ntn1* and *Epha10*. One of these, *Xlr3b*, which is one of the most upregulated genes in CYTO *vs* WT (**Figure 6.13a**) will be discussed later. At CA8 there are also other terms relating to neuronal differention: "*negative regulation of neuron differentiation*", "*positive regulation*

of neuron differentiation", "*neuron death*" and "*Wnt Signalling pathway*" (**Figure 6.14a**). At DIV8, for CYTO *vs* WT, several of the DEGs are associated with different neuronal compartments such as "*postsynaptic membrane*", "*neuron to neuron synapse*" or "*distal axon*". Interestingly, several genes belonging to the GABAergic synapse are also differentially expressed (**Figure 6.14b**).

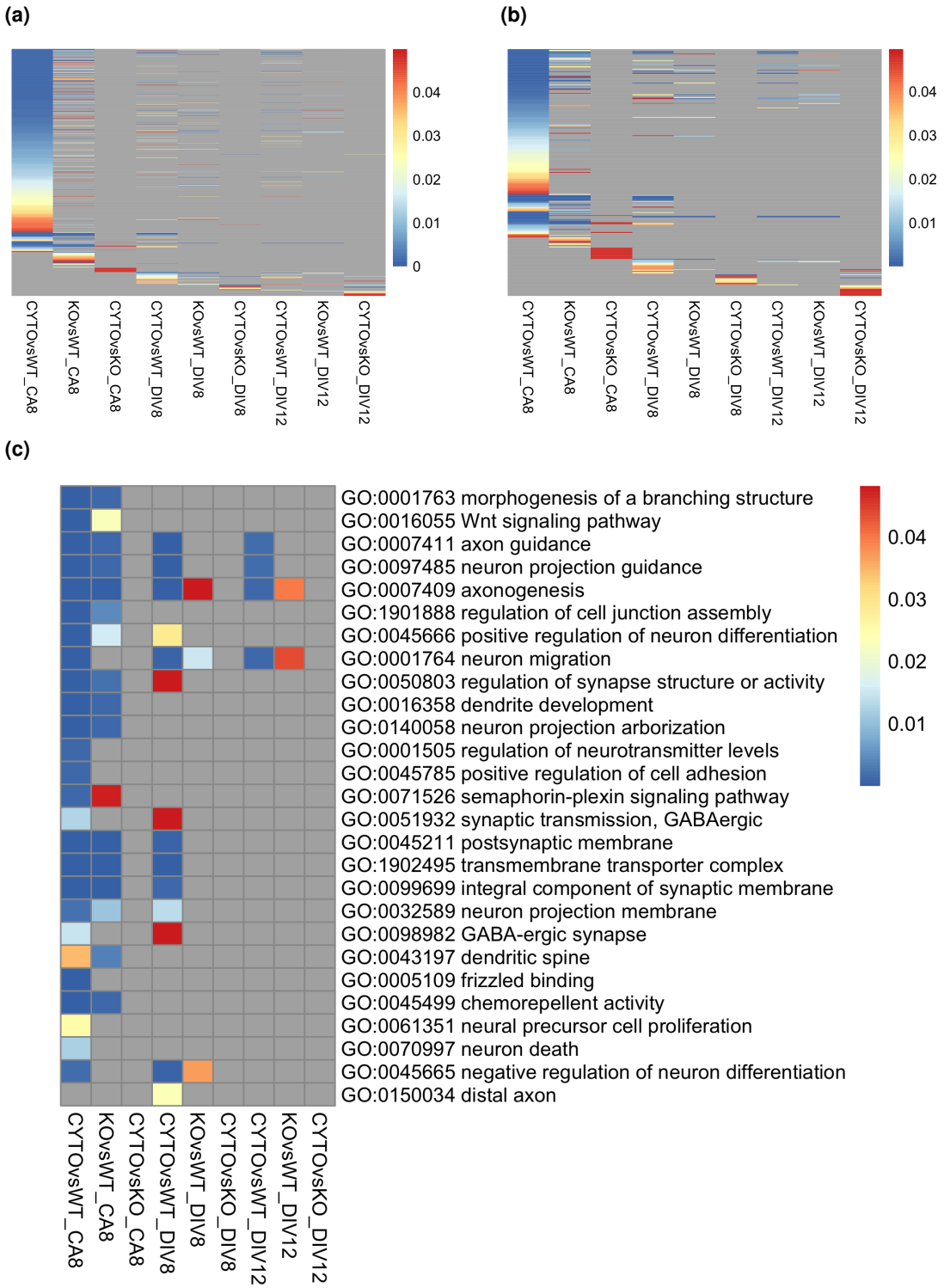
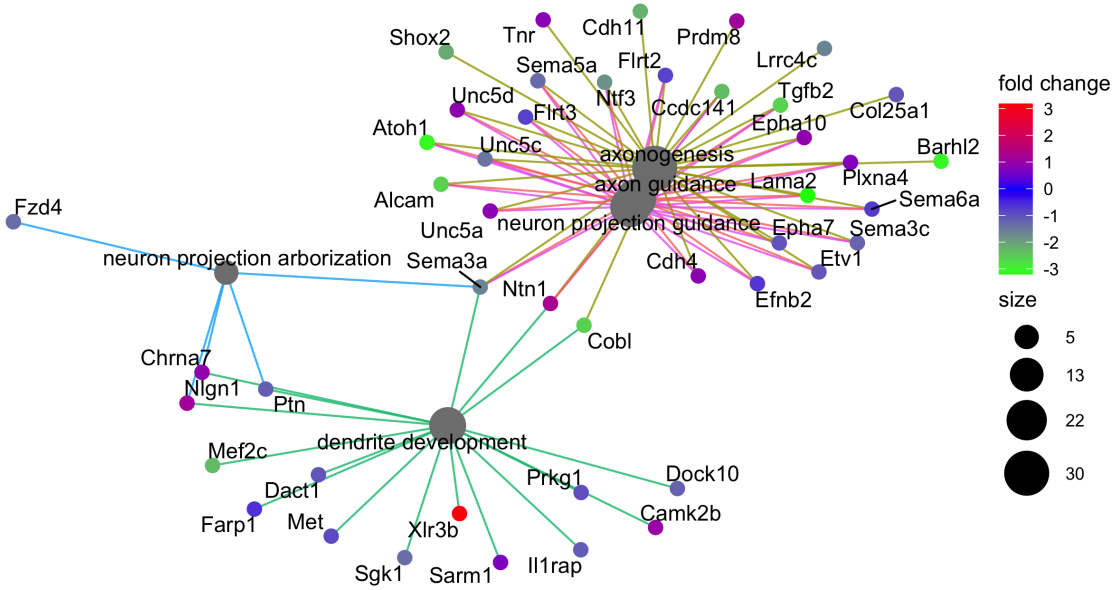


Figure 6.12: Overrepresentation Analysis of Gene Ontology terms in the DEG list. **(a)** Overview of significant GO terms, for the whole dataset, pre-filtering (1141 GO terms). **(b)** Significant GO terms, post filtering (321 GO terms). **(c)** Selected GO terms of relevance. Grey coloured box indicates term is not significant for a specific comparison. Threshold for significance was placed at p.adjusted < 0.05.

(a) CYTO *vs* WT, CA8



(b) KO *vs* WT, CA8

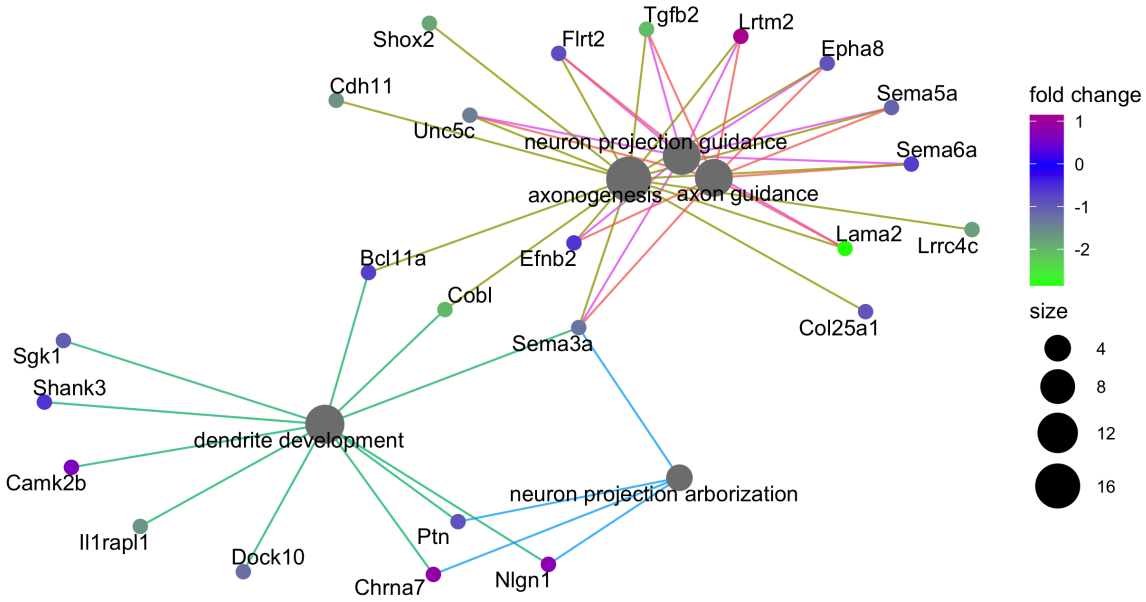
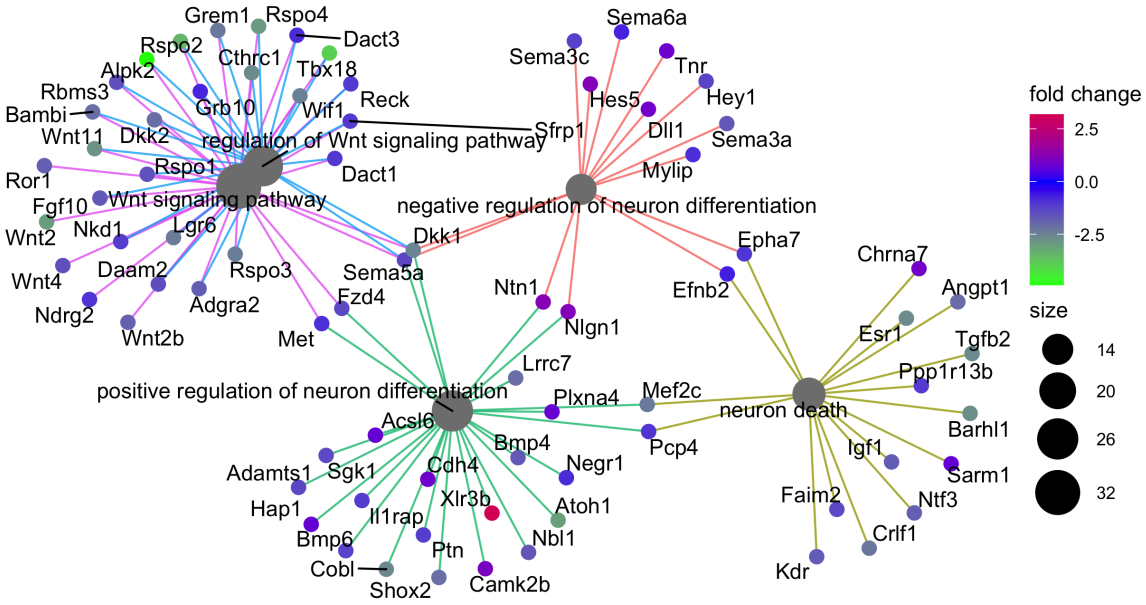


Figure 6.13: DEGs at CA8 involved in axon and dendrite morphogenesis. (a) Cnet plot of DEGs from CYTO *vs* WT at CA8, belonging to significantly enriched GO terms relating to axon and dendrite morphogenesis. **(b)** Cnet plot of DEGs from KO *vs* WT at CA8, belonging to significantly enriched GO terms relating to axon and dendrite morphogenesis. Color scale represents $\log_2$ fold change.

(a) CYTO *vs* WT, CA8



(b) CYTO *vs* WT, DIV8

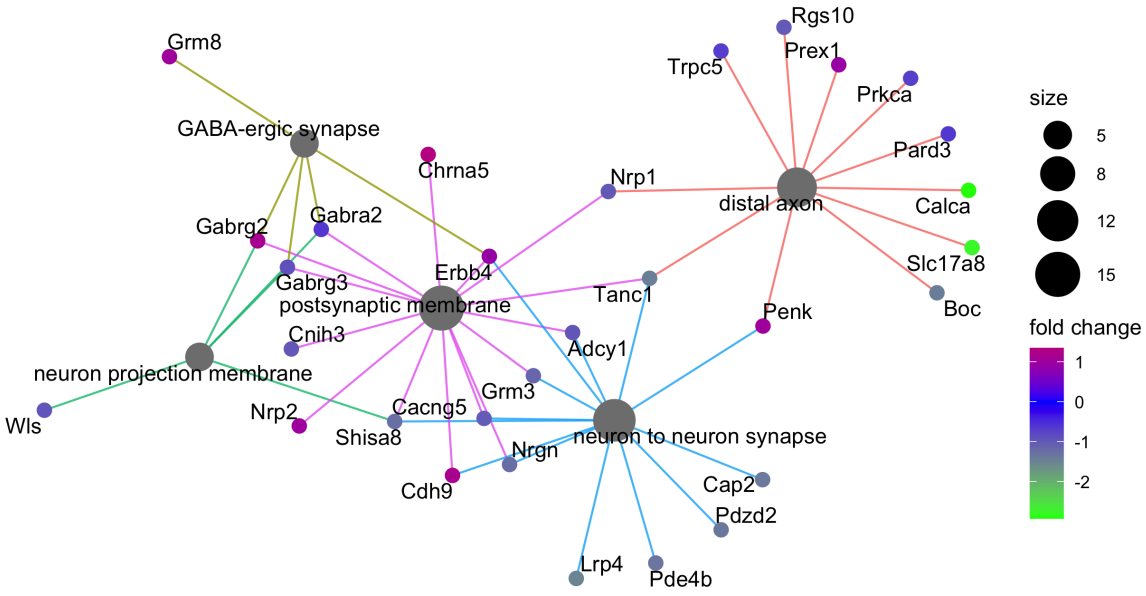


Figure 6.14: DEGs involved in neuronal differentiation and function. (a) Cnet plot of DEGs from CYTO *vs* WT at CA8, belonging to significantly enriched GO term relating to neuronal function. **(b)** Cnet plot of DEGs from CYTO *vs* WT at DIV8, belonging to significantly enriched terms relating to neuronal function. Color scale represents $\log_2$ fold change.

6.2.7 Gene Set Enrichment Analysis

To obtain a deeper insight into pathways that might be differentially regulated between the three genotypes, Gene Set Enrichment Analysis (GSEA) was carried out using the ”Cluster Profiler” package on R (Yu et al., 2012). GSEA uses all the genes in the dataset, not only the DEGs, and aggregates per gene statistics ($\log_2$ fold change) across genes within a gene set. In this way, GSEA can pick up small, but coordinated changes in a predefined gene set that would be missed when looking just at the DEGs (Subramanian et al., 2005). GSEA analysis will also determine if a pathway is ”activated” or ”suppressed” based on the $\log_2$ fold change ranking of the genes and positioning of the core enrichment genes within the ranked list. Importantly though, this classification is only based on the $\log_2$ fold change values of the genes and not on biological information of the specific genes. So, for instance, if several inhibitors of a pathway result upregulated, the pathway will be classified as ”activated”, not ”suppressed”. For this analysis, gene sets were chosen that have been associated with Gene Ontology (GO) terms and with KEGG pathways. When considering the whole dataset, more than one thousand significant GO terms (at a p.adjusted value threshold < 0.05) between ”Cellular Compartment” (GO:CC), ”Molecular Function” (GO:MF) and ”Biological Processes” (GO:BP) were identified (**Figure 6.15a**). Hence, a process of filtering and simplification of the dataset was applied, as previously done for the over-representation analysis. Interestingly, before filtering the vast majority of significant terms were found in the CYTO *vs* WT (at both CA8 and DIV8). Please note that the DIV12 data was eliminated for plotting because it did not have any significant GO terms for any comparison (**Figure 6.15**). Filtered terms were simplified, by elimination of offspring terms (**Figure 6.15b**). Finally, in a more biased approach, the list was scanned manually and selected interesting terms are presented here (**Figure 6.15c**). The ”*Wnt signaling pathway*” was suppressed for both CYTO *vs* WT and KO *vs* WT at both CA8 and DIV8. Interestingly, at DIV8 CYTO *vs* WT, the terms ”*glutamatergic neuron differentiation*” and ”*positive regulation of synaptic transmission, GABAergic*” are activated (**Figure 6.16a**), whilst *”regulation of nucleocytoplasmic transport”*, *”metallopeptidase activity”* and ”*neurotransmiter metabolic process*” are suppressed.

The GSEA analysis was also repeated for the KEGG database. Results were filtered, in a similar way as described before. Relevant results are presented in (**Figure 6.17**). Several interesting signalling pathways are altered, the majority also at CA8 and DIV8 for CYTO *vs* WT. Interestingly the *"estrogen signaling pathway"* was suppressed at CA8 for CYTO *vs* WT, as well as the *"PI3K-Akt signaling pathway"*, the *"Wnt signaling pathway"*, the *"Hippo signaling pathway"* and the *"Neuroactive ligand-receptor interaction"*, amongst others (**Figure 6.18b**). For KO *vs* WT at DIV8, *"steroid biosynthesis"* results activated. Although KEGG and GO do not always have the same nomenclature to define pathways, hence a direct comparison of terms is not straightforward, both analysis showed "*Wnt Signaling pathway*" to be significant for CYTO *vs* WT and KO *vs* WT at CA8 and DIV8.

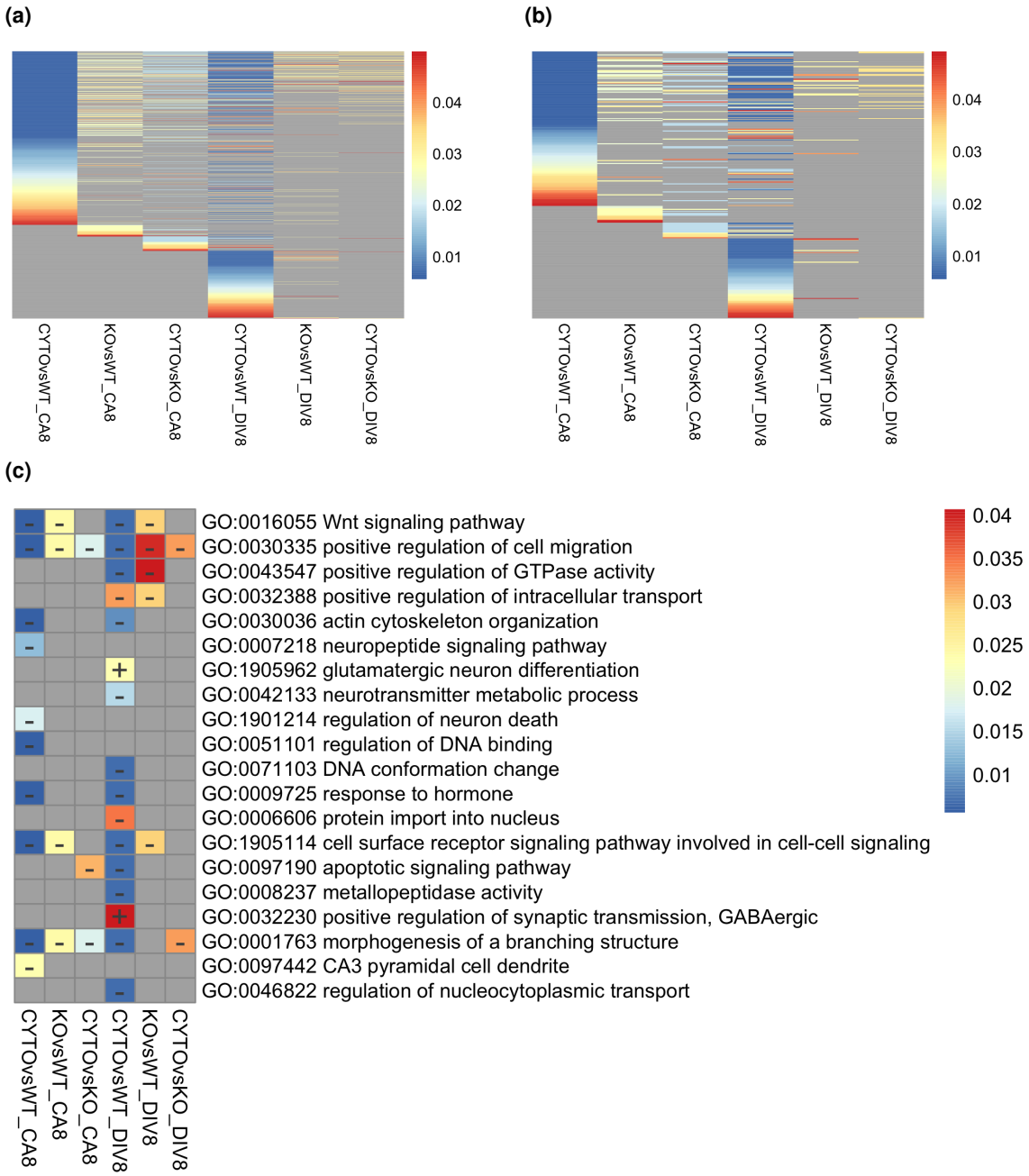
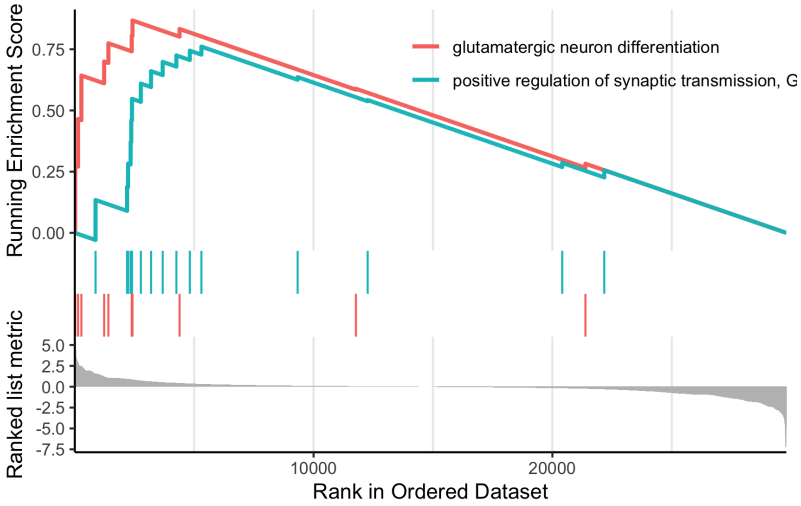


Figure 6.15: Gene Set Enrichment Analysis with Gene Ontology terms. **(a)** Overview of significant GO terms, for the whole dataset, pre-filtering (1241 GO terms). **(b)** Significant GO terms, post filtering (273 GO terms). **(c)** Selected GO terms of relevance. Grey coloured box indicates term is not significant for a specific comparison. (+) indicates ”activated” pathway; (-) indicates ”suppressed” pathway. Threshold for significance was placed at p.adjusted < 0.05.

(a) CYTO *vs* WT, DIV8



(b) CYTO *vs* WT, DIV8

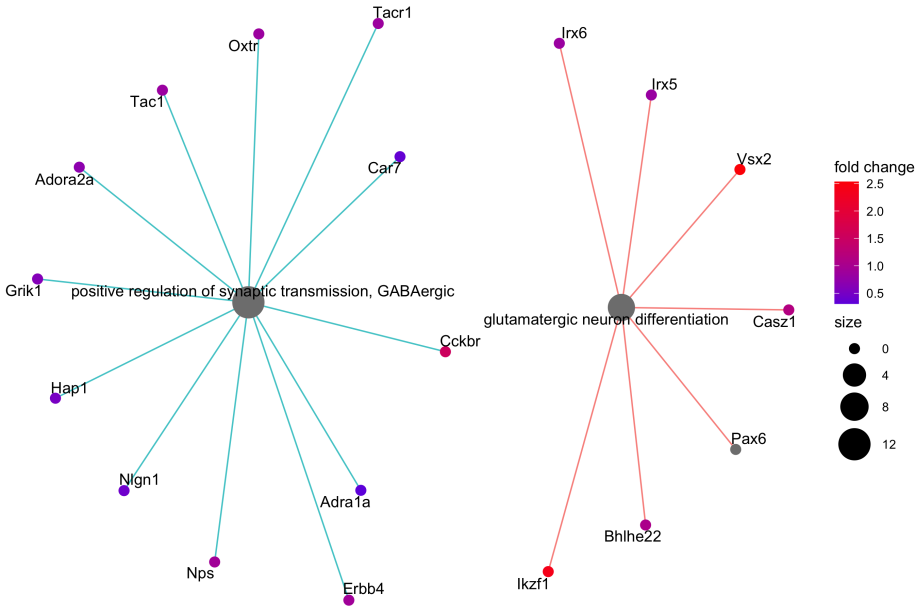


Figure 6.16: GO enrichment plots. **(a)** GSEA plot of selected enriched GO terms at DIV8 for CYTO *vs* WT. Ranking of genes is done based on $\log_2$ fold change. **(b)** Cnet plot of genes belonging to the selected GO terms. Scale of cnet plot is $\log_2$ fold change. Gene with grey dot (*Pax6*) is off the scale: $\log_2$ fold change > 2.5.

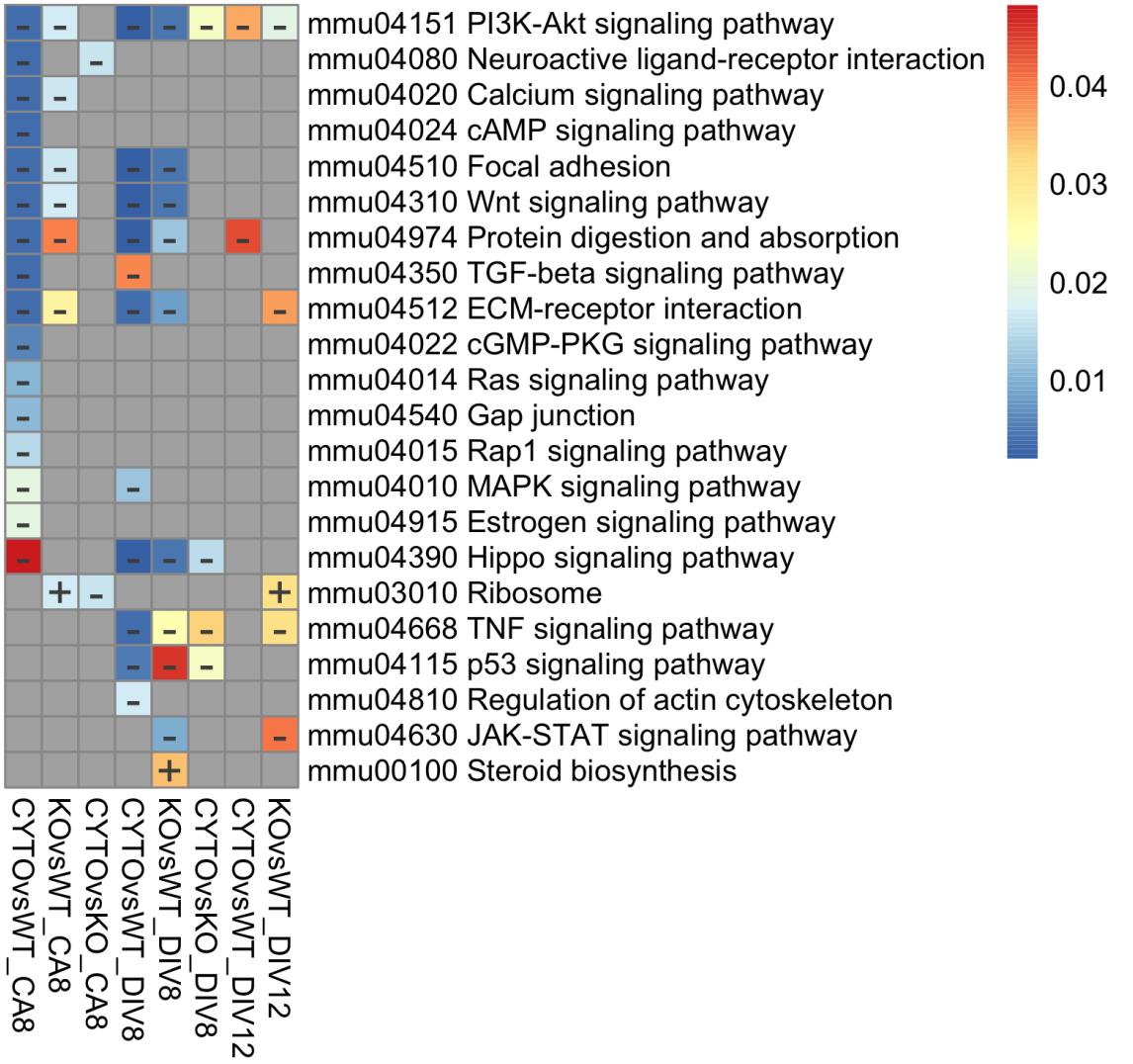
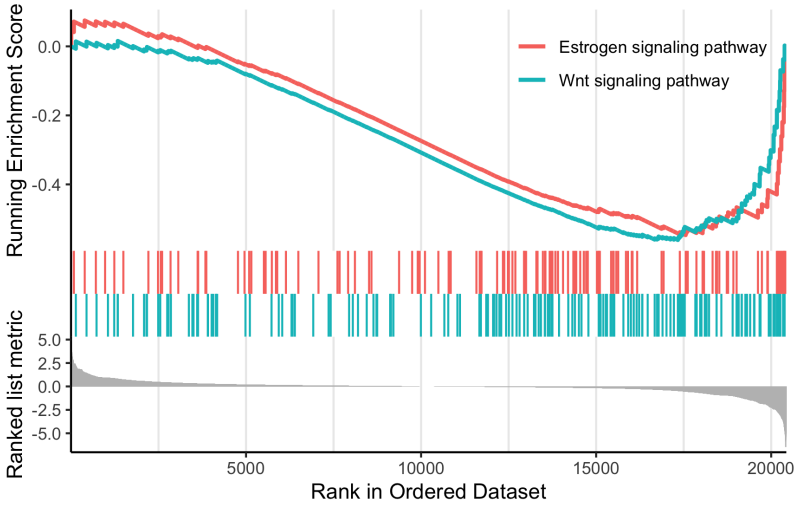


Figure 6.17: Gene Set Enrichment Analysis with KEGG. Selected significant pathways on KEGG at CA8, DIV8 and DIV12. Selected GO terms of relevance. Grey coloured box indicates term is not significant for a specific comparison. (+) indicates the pathway is "activated"; (-) indicates the pathway is "suppressed". Threshold for significance was placed at p.adjusted < 0.05.

(a) CYTO *vs* WT, CA8



(b) CYTO *vs* WT, CA8

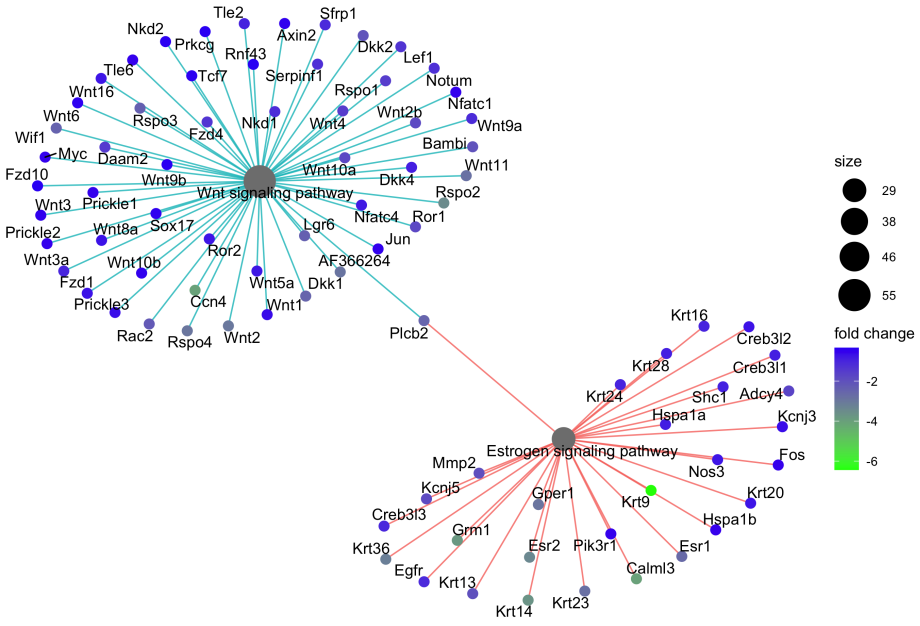


Figure 6.18: KEGG enrichment plots. **(a)** GSEA plot of selected KEGG pathways "*estrogen signalling*" and *Wnt signalling*" enriched at CA8 for CYTO *vs* WT. Ranking of genes is done based on $\log_2$ fold change. **(b)** Cnet plot of genes belonging to the selected KEGG pathways. Scale of cnet plot is $\log_2$ fold change.

Selected CYTO DEGs

Because of the interest in uncovering the function of the cytoplasmic domain of PCDH19 and the substantial overlap between CYTO *vs* WT and KO *vs* WT differentially expressed genes, we decided to focus on CYTO-specific differentially expressed genes.

The X-linked lymphocyte regulated (*Xlr*) genes *Xlr3a*, *Xlr3b* and *Xlr4a* were significantly upregulated in CYTO *vs* WT and CYTO *vs* KO, across all timepoints (*Xlr3a* - CYTO *vs* WT: CA8, p.adj = 3.07×10^{-19}; DIV8, p.adj = 5.14×10^{-18}; DIV12, p.adj = 9.9×10^{-16} (**Figure 6.19a**); *Xlr3b* - CYTO *vs* WT: CA8, p.adj = 1.26×10^{-18}; DIV8, p.adj = 1.67×10^{-11}; DIV12, p.adj = 1.04×10^{-10} (**Figure 6.19b**); *Xlr4a* - CYTO *vs* WT: CA8, p.adj = 4.85×10^{-06}; DIV8, p.adj = 3.54×10^{-08}; DIV12, p.adj = 9.24×10^{-07}; CYTO *vs* KO: CA8, p.adj = NA; DIV8, p.adj = 4.87×10^{-09}; DIV12, p.adj = 2.55×10^{-06} (**Figure 6.19c**) *Xlr4b* approached the significance threshold in CYTO compared to WT and KO cells (**Figure 6.19d**)). The *Xlr* genes are a cluster of X-linked imprinted genes in the mouse (Raefski and O'Neill, 2005), of which the human orthologs are the *FAM9* family (Martinez-Garay et al., 2002). Not much is known about the function of these genes in brain development and function. However, it has been described that CUX1, a transcription factor expressed mainly in upper-layer cortical neurons (Nieto et al., 2004), binds and represses expression of *Xlr3b* and *Xlr4b* in order to regulate dendritic morphogenesis and spine number in layer II/III neurons in the mouse (Cubelos et al., 2010) . Overexpression of *Xlr4b* via *in utero* electroporation leads to aberrant spine morphology and reduced number of spines (Cubelos et al., 2010). Moreover, in another study, increased expression of *Xlr3b* in the brain was correlated to behavioural deficits in a mouse model of Turner Syndrome (Davies et al., 2005).

Nxf3, nuclear RNA Export Factor 3 RNA binding protein, is also upregulated in CYTO *vs* WT and CYTO *vs* KO (CYTO *vs* WT: CA8, p.adj = 0.271; DIV8, p.adj = 2.8×10^{-07}; DIV12, p.adj = 0.000131; CYTO *vs* KO: CA8, p.adj = NA; DIV8, p.adj = 2.25×10^{-12}; DIV12, p.adj = 4.19×10^{-10} (**Figure 6.19f**)). *Nxf3* can mediate transport of RNA outside the nucleus (Yang et al., 2001).

Interestingly, *Nxf3* has been linked to mental retardation (Jun et al., 2001)

Kdm5d, Lysine Demethylase 5D, Histone demethylase (**Figure 6.19e**) male specific (Y-chromosome demethylase) is upregulated in CYTO only (CYTO *vs* WT: CA8, p.adj = 5.92×10^{-11}; DIV8, p.adj = 1.53×10^{-08}; DIV12, p.adj = 4.66×10^{-06}; CYTO *vs* KO: CA8, p.adj = NA; DIV8, p.adj = 2.01×10^{-07}; DIV12, p.adj = 1.42×10^{-05}). *Kdm5d* belongs to a protein family that also includes *Kdm5a*, *Kdm5b*, *Kdm5C*. Mutations in *KDM5A*, *KDM5B*, and *KDM5C* are associated with intellectual disability, with genetic variants in *KDM5C* associated with a disorder known as Mental Retardation, X-linked, Syndromic, Claes-Jensen type (Gonçalves et al., 2014).

In line with the results from the GSEA analysis, which found "*estrogen signalling pathway* " to be suppressed in CYTO *vs* WT at CA8 (**Figure 6.18a**), *Esr1*, the estrogen receptor 1 (alpha), is downregulated in CYTO *vs* WT at CA8 (CYTO *vs* WT: CA8, p.adj = 0.000126) but in none of the other time points or comparisons (**Figure 6.19g**). Interestingly, as mentioned, *Esr1* has been shown to modulate gene expression together with PCDH19 (Pham et al., 2017). Together with *Esr1*, other estrogen related receptors or estrogen regulated proteins are also altered either in CYTO cells alone or in both CYTO and KO cells. *Essrb*, estrogen-related receptor beta, for example, is a nuclear orphan receptor, which is found upregulated in CYTO *vs* WT and in KO *vs* WT at DIV8 and DIV12 (CYTO *vs* WT: DIV8, p.adj = 4.4×10^{-26}; DIV12, p.adj = 1.22×10^{-17}; KO *vs* WT: DIV8, p.adj = 3.19×10^{-17}; DIV12, p.adj = 5.18×10^{-14}) (**Figure 6.19h**). *Essrb* can reprogram mouse fibroblasts into induced pluripotent stem cells, replacing *c-Myc* and *Klf4* of the 4 traditional Yamanaka factors *Oct4*, *Sox2*, *c-Myc* and *Klf4* (Feng et al., 2009) reviewied by (Divekar et al., 2016). *Essrg*, estrogen-related receptor gamma, is upregulated in CYTO *vs* WT but not in KO *vs* WT at DIV12 (CYTO *vs* WT: DIV12, p.adj = 6.31×10^{-05}). *Essrg* is a hormone independent orphan-receptor but does bind to estrogen-response elements (EREs) and can activate reporter genes controlled by EREs (Hong et al., 1999). Other estrogen regulated genes such as *Rerg*, RAS-like, estrogen-regulated, and *Greb1*, growth-inhibitor and gene regulated by estrogen in breast cancer protein, are also differentially expressed. *Rerg* is altered in both CYTO and KO at

CA8 (CYTO *vs* WT: CA8, p.adj = 8.2×10^{-08}; KO *vs* WT: CA8, p.adj = 7.28×10^{-06}) whilst *Greb1* was dysregulated only in CYTO at DIV8 (CYTO *vs* WT: DIV8, p.adj = 0.000537).

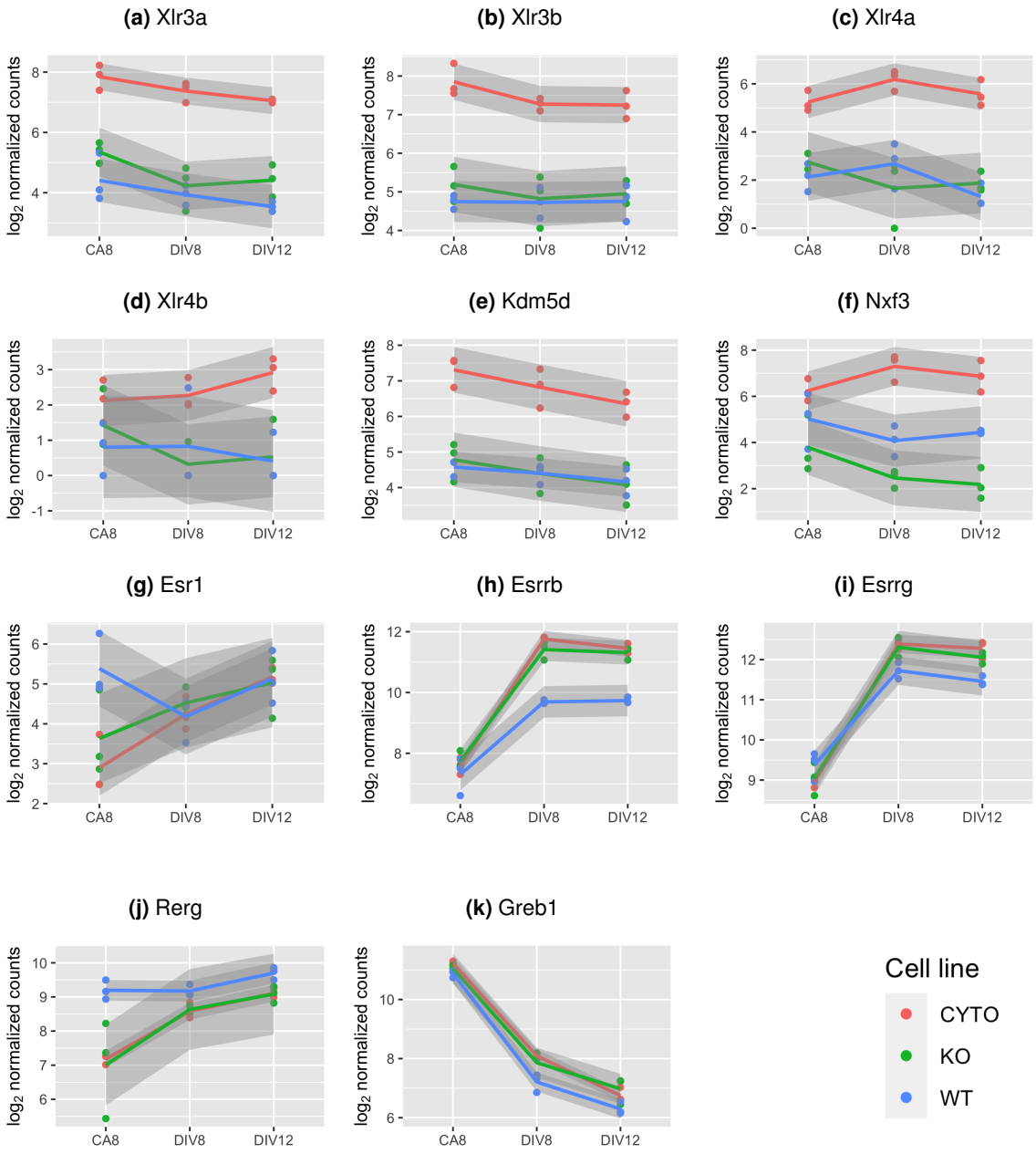


Figure 6.19: Expression of selected genes at CA8, DIV8 and DIV12. Expression of selected genes during differentiation in CYTO (red), KO (green) and WT (blue) cells, plotted as $\log_2$ normalised counts.

6.2.8 Transcriptional changes induced by neuronal activity

As PCDH19 is processed in response to neuronal activity, it was interesting to investigate if stimulating neuronal activity in CYTO, KO and WT cells would have different effects on transcription. To address this, DIV12 neurons were treated with 10 μM Bicuculline to induce neuronal activity. Bicuculline is a competitive antagonist of GABA-A receptors, and therefore promotes activity of neurons by easing inhibitory action of GABAergic neurons, which should comprise about 5% of the cells in culture. Neurons were treated for 4 hours before lysis and RNA extraction (**Figure 6.20a**). This experiment was started before the processing data (**Chapter 3**) were generated, hence it was yet to be determined that NMDA treatments could stimulate processing. Principal Component Analysis (PCA) (**Figure 6.20b**) was done to see how similar cells were. PC1 explains 32% of the variance, whilst PC2 explains 26% of the variance. Correlation analysis shows that the CYTO BIC samples cluster together, whilst clustering of other samples is less obvious (**Figure 6.20c**).

Differential expression analysis was carried out comparing bicuculline treated (DIV12 + BIC) *vs* untreated (DIV12) samples within each cell line (CYTO DIV 12 + BIC *vs* CYTO DIV12; KO DIV12 + BIC *vs* KO DIV 12; WT DIV 12 + BIC *vs* WT DIV12) (**Figure 6.21**). In this analysis we can investigate how the different cell lines respond to prolonged induced neuronal activity. For CYTO, WT and KO, a set of common genes that responds to treatment was identified. The vast majority of genes that are differentially expressed in WT neurons in response to bicuculline are also altered in KO and CYTO, and might therefore represent a ”core” response (82.7%). Amongst these ”core” set of DEGs there are several upregulated aminoacyl t-RNA synthetases (ARSs) enzymes that mediate transfer of amino acids on their respective tRNAs (*Cars, Iars, Lars*), possibly reflecting changes in protein translation. CYTO and KO have many more differentially expressed genes than WT and with a substantial overlap between them. 48.8% of CYTO DEGs are also KO DEGs and 61.8% of KO DEGs are also CYTO DEGs (**Figure 6.21c**).

Gene Set Enrichment Analysis (GSEA) was conducted on this dataset in order to extract potential differences between genotypes. Using the GO database, there were no enriched data sets. Using the KEGG database, there were 2 significant pathways of relevance for CYTO DIV12 *vs* CYTO DIV12 BIC that were "*neuroactive ligand-receptor interaction*", which resulted suppressed and "*RNA transport*", which was activated (**Figure 6.22a**). Genes belonging to these pathways are shown in a cnetplot (**Figure 6.22b**). The "*neuroactive ligand-receptor interaction*" KEGG pathway was also suppressed for KO *vs* WT.

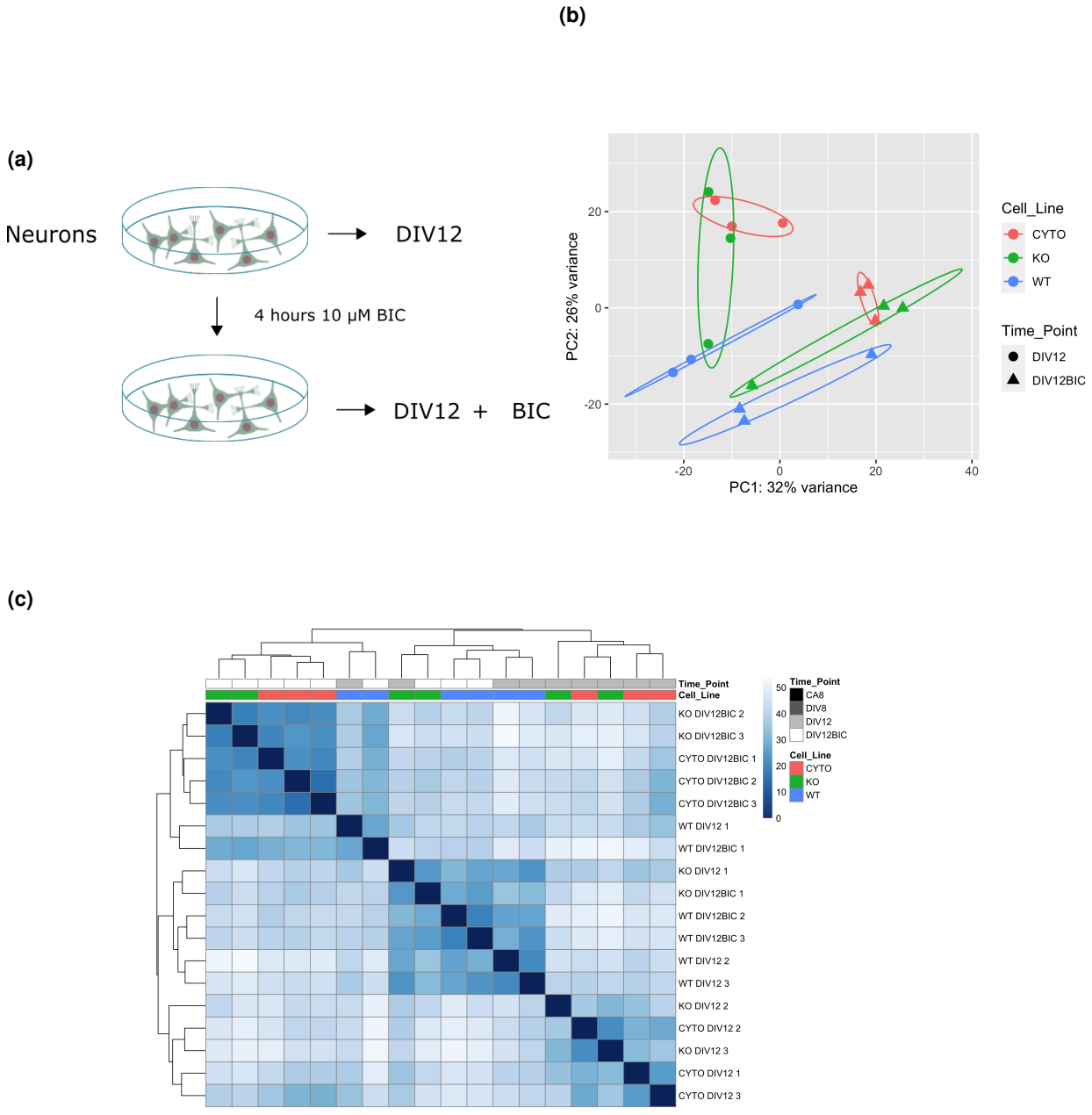


Figure 6.20: Bicuculline treated and untreated DIV12 samples. **(a)** Experimental design. **(b)** Principal Component Analysis. Colour of dot indicates cell-line (CYTO, KO, WT), shape of dot indicates treatment (DIV12, no treatment, DIV12 + BIC, treatment). **(c)** Correlation Matrix. Blue scale represents distance between samples, with darker blue corresponding to smaller distances.

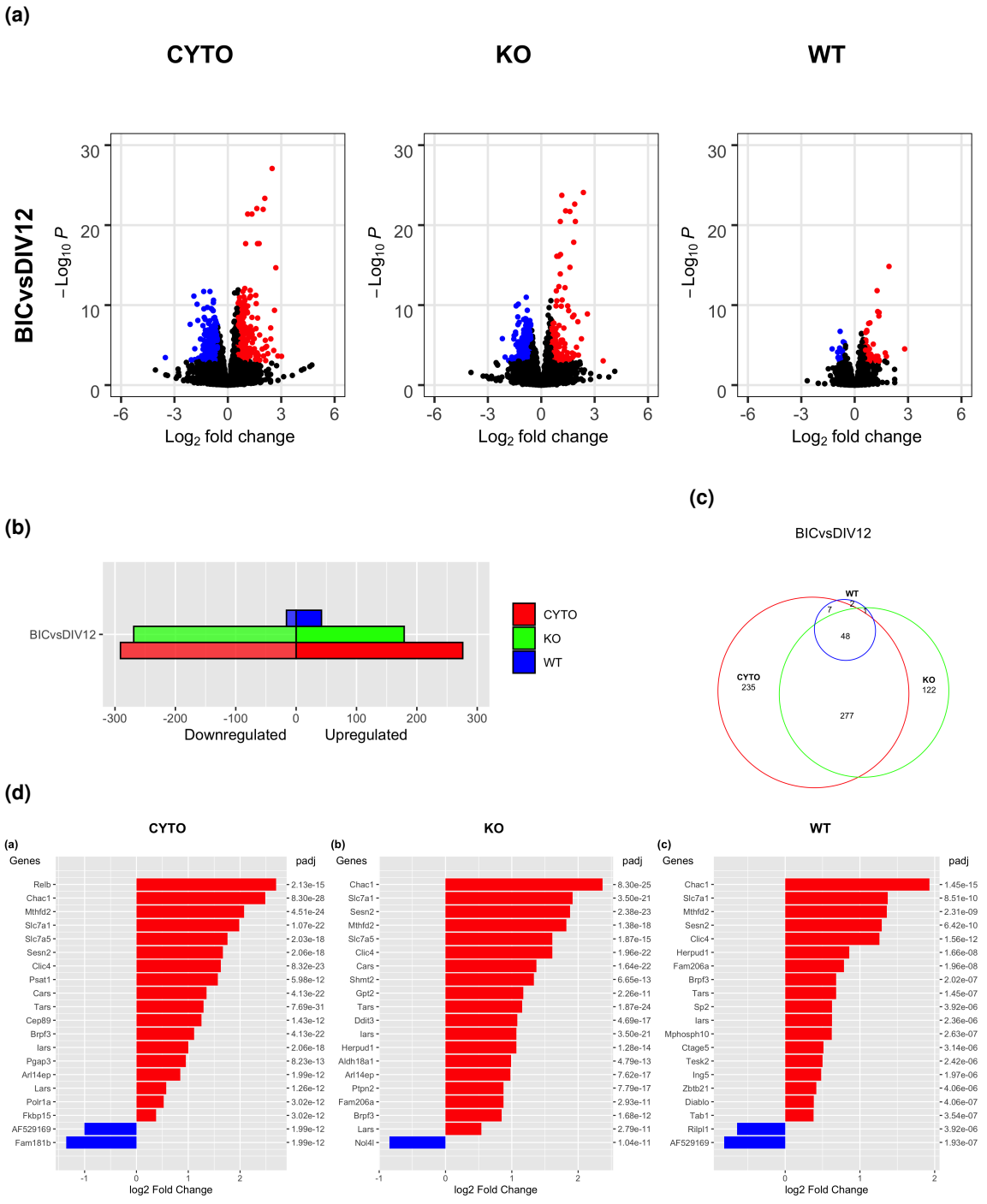
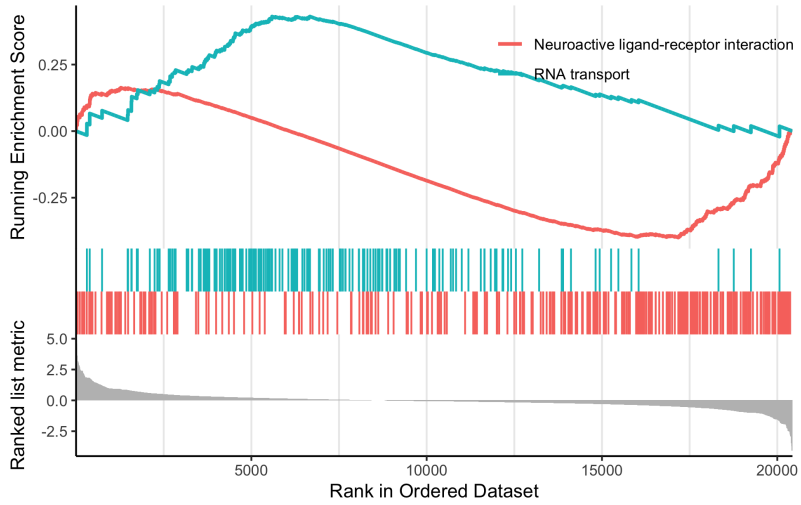


Figure 6.21: Differentially expressed genes for bicuculline treated (DIV12 + BIC) *vs* untreated (DIV12) neurons for each cell line (CYTO, KO, WT). **(a)** Downregulated genes are shown in blue, upregulated genes are shown in red. Threshold for p.adj value is 0.001, $\log_2$ fold change cut off > 0.58. **(b)** Total upregulated and downregulated number of genes for each comparison. **(c)** Overlapping genes up or downregulated in response to bicuculline in WT, CYTO and KO neurons at DIV12. **(d)** 20 most significant genes for each comparison based on p.adjusted value.

(a) CYTO, DIV12 + BIC *vs* DIV12



(b) CYTO, DIV12 + BIC *vs* DIV12

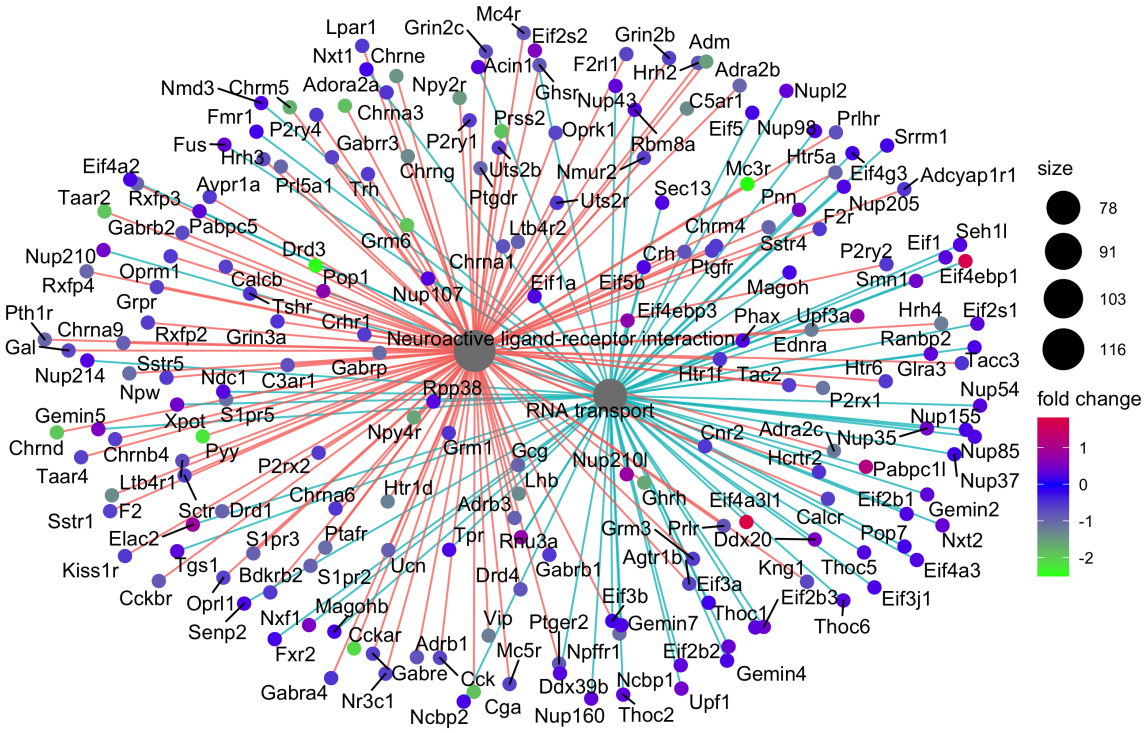


Figure 6.22: Gene Set Enrichment Analysis with KEGG for CYTO DIV12 + BIC *vs* DIV12. **(a)** GSEA with KEGG for two significant pathways for CYTO DIV12 + BIC *vs* CYTO DIV12, "*neuroactive ligand-receptor interaction*" and "*RNA transport*". Ranking of genes is done based on $\log_2$ fold change. **(b)** Genes present in the dataset belonging to the above pathways. Colour of dot indicates $\log_2$ fold change.

6.2.9 *In vivo* overexpression of PCDH19-CYTO

Several members of the protocadherin family have been implicated in dendrite and spine morphogenesis, reflecting a role in synapse formation and circuit function. For instance, PCDH19 downregulation via *in utero* electroporation of shRNAs in the hippocampus results in decreased total length of dendrites. Interestingly, whilst apical dendrite length was reduced, basal dendrite length was increased (Bassani et al., 2018). Overexpression of PCDH17 in primary cortical neurons results in decreased spine density and abnormal spine morphology, with increased stubby spines and decreased mushroom-shaped ones (Chang et al., 2018). *Pcdh10*$^{+/-}$ mice have increased spine density in the amygdala due to elevated number of elongated filipodia-like spines, which are an immature type of spine (Schoch et al., 2017). Overexpression of a dominant-negative form of α and γ-Pcdh, encoding a myristoylated form of the cytoplasmic domain constant region, (membrane-attached ICD) leads to defects both in dendrite morphogenesis and in spine density (Suo et al., 2012), possibly due to sequestering interactors.

As mentioned, several members of the X-linked lymphocyte regulated (*Xlr*) gene family are amongst the most differentially expressed genes in CYTO *vs* WT neurons and CYTO *vs* KO and play a role in upper-layer neuron dendrite morphogenesis. In fact, as previously stated, the transcription factor CUX1 binds and represses expression of *Xlr4b* and *Xlr3b* to regulate dendritic morphogenesis and spine number layer II/III cortical neurons (Cubelos et al., 2010). Overexpression of *Xlr4b* leads to aberrant spine morphology and reduced number of spines (Cubelos et al., 2010). Other genes, involved in similar processes, were also differentially expressed. For example *Sema3a*, a secreted chemorepellent important for axon growth repulsion, *Sema3c* and *Sema6a* are all downregulated at CA8 CYTO *vs* WT samples, as previously shown. *Sema3a* is also known to be involved in dendrite morphogenesis (Polleux et al., 2000). In addition, *Foxp2*, which regulates neurite outgrowth in primary neurons (Vernes et al., 2011), is also upregulated in CYTO *vs* WT.

Given the evidence in the literature supporting a role for protocadherins in dendritic and spine morphogenesis and the results from the RNA sequencing, it was decided to investigate the effect of PCDH19-CYTO overexpression *in vivo* by using *in utero* electroporation. PCDH19 is expressed in the adult cortex predominantly in layers V and II/III. Given time constraints and unforeseen circumstances, this analysis, which initially focused on spines and dendrites in layer II/III neurons of the cortex, could not be completed. However, preliminary results are presented nonetheless to showcase the viability of this method to investigate the role of the intracellular domain of PCDH19. *In utero* electroporation (IUE) was performed by Dr. Cristina Llinares-Benadero, as described in methods (**Section 2.1.4**), with plasmids expressing a tagged version of PCDH19-ICD and EGFP (pCIG-PCDH19-CYTO-HA) or just EGFP (pCIG). IUE was performed at E15.5 on C57BL6J WT animals (plasmids described in (**Table 2.2**) (**Figure 6.23a**). EGFP expression is necessary to localise the electroporated area and to trace the dendritic arbour and the spines of targeted neurons. Electroporated brains were collected at P60 for downstream processing. For each litter, half of the embryos were electroporated with the experimental condition in the left hemisphere and the other half with control plasmid in the right hemisphere. All brains were processed via immunohistochemistry to confirm experimental group via detection of the HA-tag that is only present in the pCIG-PCDH19-CYTO-HA (**Figure 6.23a**). Dendritic morphology will be analysed for whole neurons (**Figure 6.24a**), whilst a spine analysis will be performed on representative second-order apical and basal dendrites belonging to the traced neurons (**Figure 6.24b**).

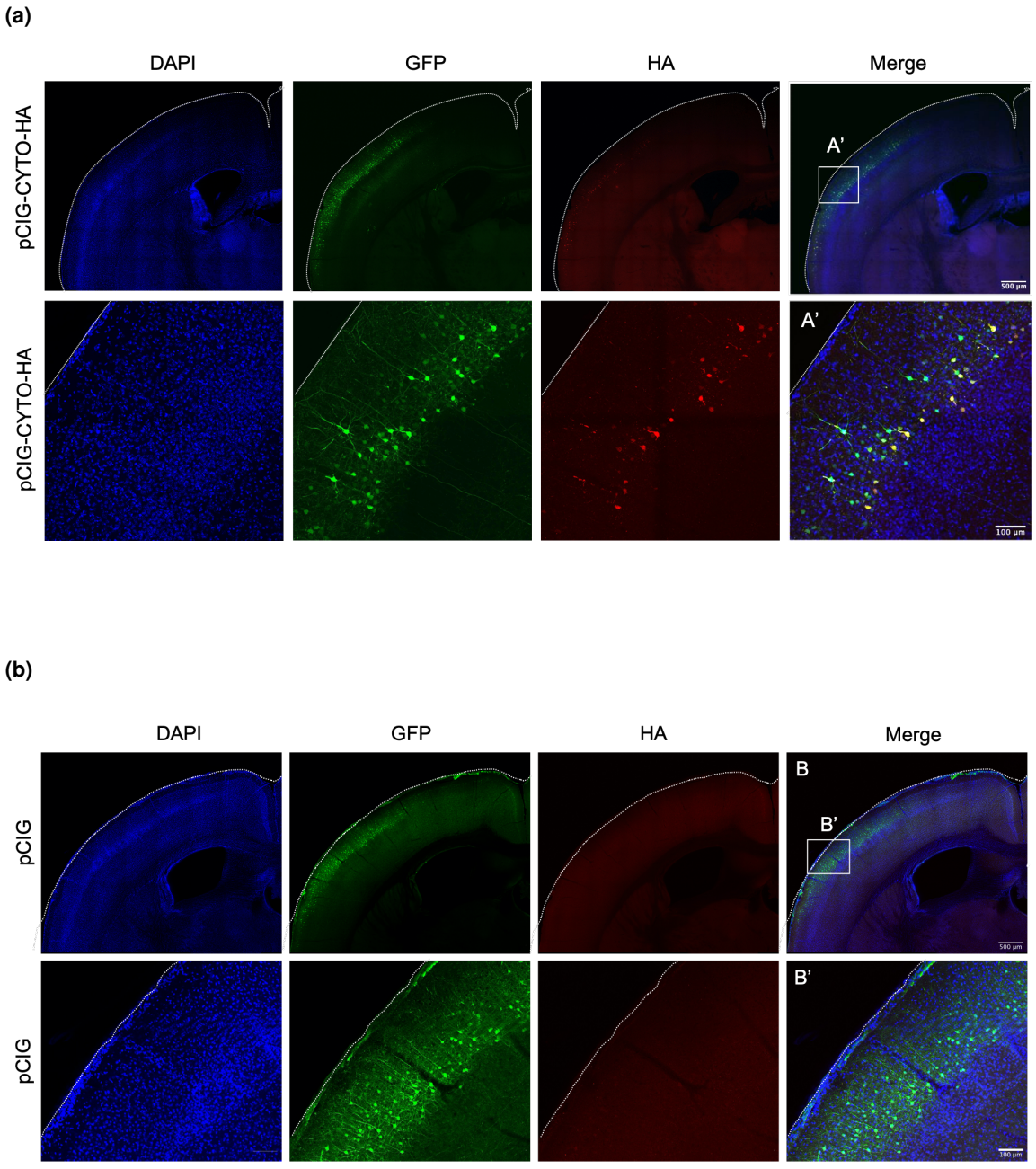


Figure 6.23: PCDH19-CYTO-HA overexpression *in vivo*. **(a)** *In utero* electroporation of PCDH19-CYTO-HA at E15.5 -> P60. **(b)** *In utero* electroporation of pCIG at E15.5 -> P60. Scale bars: A-B, 500 μm; A'-B', 100 μm.

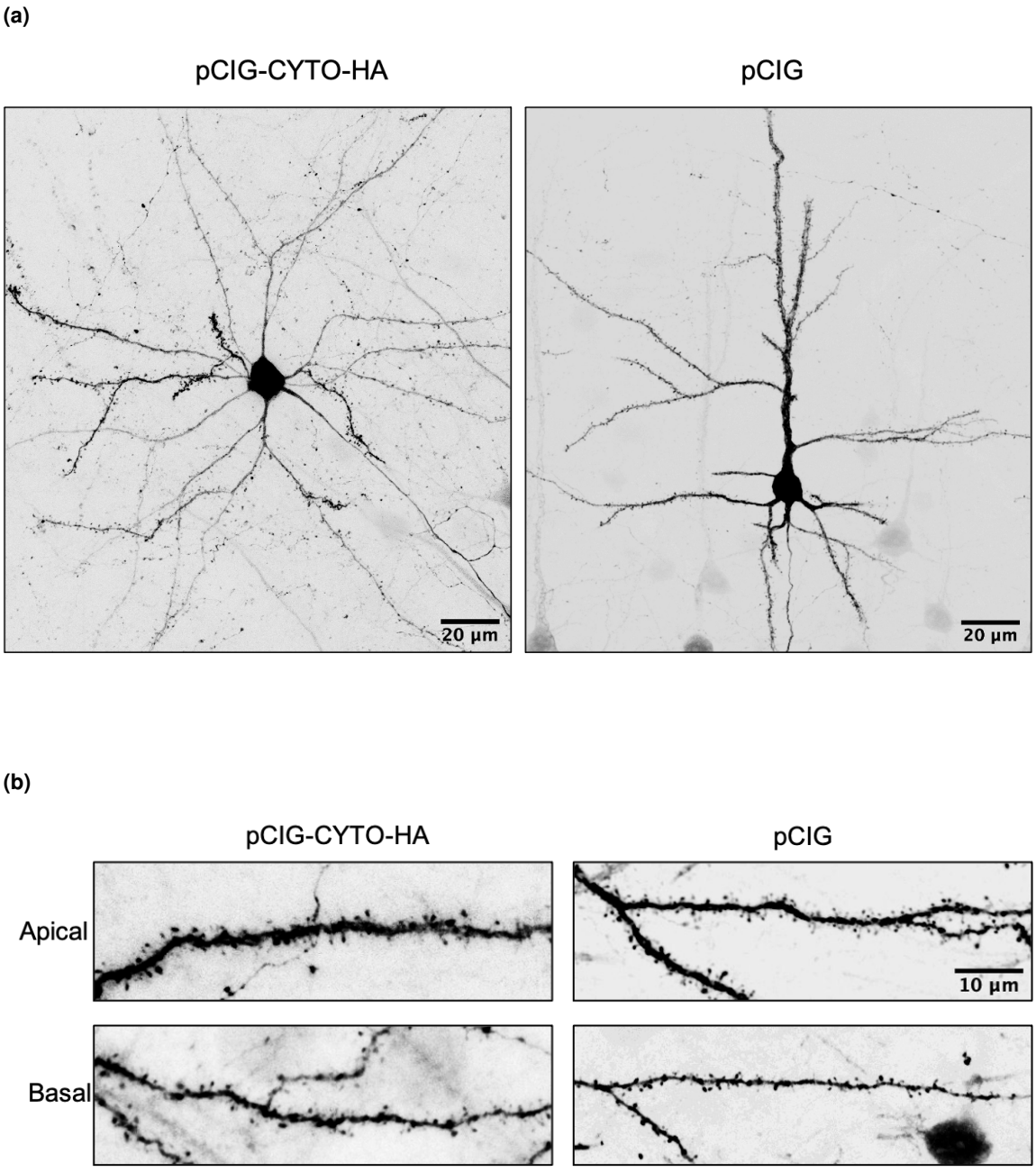


Figure 6.24: PCDH19-CYTO overexpression effect on morphology of layer II/III neurons *in vivo*. **(a)** pCIG-PCDH19-CYTO-HA or pCIG IUE layer II/III cortical neurons imaged at 40X for dendrite tracing. Scale bar: 20 μm. **(b)** Representative 63X imaging of apical and basal dendritic spines following pCIG-PCDH19-CYTO-HA or pCIG IUE. Scale bar: 10 μm.

6.3 Discussion

This chapter aimed to investigate the *in vitro* transcriptional profile of progenitors and neurons that lack PCDH19 or overexpress its cytoplasmic domain. Confirmation of PCDH19-CYTO overexpression was obtained both by differential expression analysis and by splicing analysis, showing increased transcription of exons 2-6 that encode the cytoplasmic domain. The overexpression of PCDH19-CYTO was most pronounced in CA8 progenitors, presumably because at this stage the *Rosa26* promoter is more active than the *Pcdh19* promoter, giving rise to the biggest difference between PCDH19-ICD and endogenous levels. As neuronal maturation progresses endogenous *Pcdh19* expression increases, while *Rosa26* promoter activity decreases, hence the overexpression of CYTO becomes more subtle at DIV8 and DIV12. Interestingly, γ-secretase, the protease that would generate PCDH19 CTF2, seems to be more active in progenitors than in neurons, based on the observation that γ-secretase generated ICD fragments of APP, Notch1, N-Cadherin, EphrinB and p75-NTR are all reduced in adult brain compared to embryonic brain (Frånberg et al., 2010). Potentially, this could also explain why the biggest effects of CYTO overexpression are seen at CA8, as progenitors might be in the right physiological state to respond transcriptionally to increased levels of PCDH19-ICD. Although PCDH19-CYTO overexpression was further confirmed by the sequencing data, PCDH19-KO could not be validated this way, as *Pcdh19* mRNA was found in PCDH19-KO progenitors and neurons at the same levels than in WT cells. Since PCDH19 protein cannot be detected in PCDH19-KO ES cells or ESC-derived neurons, as verified by western blot in the previous chapter, these results indicate that *Pcdh19* mRNA is escaping non-sense mediated decay (NSMD). Non-sense mediated decay efficiency depends on a variety of factors, mainly location of the premature termination codon within the gene. If the premature termination codon is very close to the start codon it can reduce efficiency of non-sense mediated decay (Lindeboom et al., 2016), (Popp and Maquat, 2016). In the targeted PCDH19-KO ESCs the STOP codon was very close to the start of the first exon and much before the first exon/exon boundary, potentially reducing efficiency of NSMD.

Differential expression analysis showed that PCDH19-ICD overexpression leads to changes at the transcriptional level, which are most pronounced at the neuronal progenitor stage (CA8) and diminish during neuronal maturation, maybe reflecting relative levels of PCDH19-CYTO overexpression, or perhaps a higher susceptibility of progenitors to transcriptional changes brought about by this fragment. Overexpression of CYTO led to bigger changes than complete loss of the protein, somewhat in line with the fact that complete lack of PCDH19 in mice and humans does not induce disease. Surprisingly, there was a high overlap of differentially expressed genes (DEGs) across timepoints between CYTO and KO cells. CYTO and KO cells are therefore fairly similar when compared to WT cells. As mentioned previously, cadherins are known to exert dominant-negative effects when expressed without their extracellular domain (Nieman et al., 1999), (Heggem and Bradley, 2003), (Gil-Sanz et al., 2013). Therefore, although no obvious defects were observed when generating the ES cell line, and despite the fact that CYTO cells were able to differentiate into neurons, a potential dominant negative effect of PCDH19-CYTO overexpression by sequestering PCDH19 cytoplasmic interactors cannot be ruled out. This could interfere with the normal function of PCDH19 and, to some extent, mimic the lack of the protein. These effects would probably be more subtle, as PCDH19 KO does not have huge effects of its own. The similarities observed in PCDH19-KO and PCDH19-CYTO cells could also be explained by a dosage-dependent effect whereby too much or too little PCDH19 leads to similar detrimental effects. For instance, hypothetically, if PCDH19 or PCDH19-CYTO belonged to a complex which functioned only with a precise stoichiometry, both overexpression and removal of PCDH19, could have the same disruptive effect on complex assembly (Bergendahl et al., 2019), (Sopko et al., 2006).

Another explanation for the similarities between CYTO and KO cells could be their 'history'. CYTO and KO cells were generated, starting from the same E14 cells, in two different ways: PCDH19-CYTO cells were created by ZFN targeting of the *Rosa26* locus whilst PCDH19-KO were created by CRISPR/Cas9 of the *Pcdh19* locus. Two separate methods were utilised for clone selection: CYTO cells were selected via antibiotic resistance and picked manually, KO cells

were sorted via FACS as single-cells into a 96-well plate. Moreover, CYTO cells had to go through a second round of nucleofection and colony picking in order to remove the selection cassette, which considerably increased the length of the process. Although the two targeting processes are technically different, both CYTO and KO cells were maintained in culture for a substantial number of passages beyond that of WT cells and they were then pre-cultured on MEFs for 5-6 passages before differentiation. Hence, passage number of KO ESCs was considerably higher than WT, and CYTO passage numbers were higher than KO. One of the downsides of ESC use is that prolonged culture of ESCs can lead to epigenetic instability of the cells (Gaztelumendi and Nogués, 2014), (Humpherys et al., 2001). In fact, some imprinted genes like *H19*, *Igf2*, *Peg1*/*Mest* and *Meg1*/*Grb10* have been shown to lose methylation in culture and can show differential expression in different ESCs lines which is reflected by changes in DNA methylation (Dean et al., 1998). Changes can go in both directions, with both loss of and increased methylation affecting maternal and paternal alleles differently. Surprisingly, changes can happen after only 10 passages *in vitro* (Lee et al., 2018).

Some of the most differentially expressed genes that were identified in this study, including *Mest*, *Xlr3a*, *Xlr3b, Peg13* and *Nnat* are, in fact, known imprinted genes, although other previously described imprinted genes such as *Igf2* were not changed in this dataset. It therefore remains a possibility that KO and CYTO ES cells were in a different methylation status than the E14 controls and that the substantial overlap between CYTO *vs* WT and KO *vs* WT differentially expressed genes potentially reflects changes due to the prolonged culture of cells. Retrospectively, a more appropriate control for the RNAseq analysis would have been an untargeted clone, with the same 'history' as CYTO or KO ESCs clones, but not carrying the insertion or mutation - in this way it would have been possible to control for effects of prolonged culture.

Given these observations, interpretation of the obtained results is complicated because of the potentially confounding effects of prolonged culture and/or genetic engineering of cells, which cannot be controlled for because of the different 'history' of WT cells. Nonetheless, it

cannot be excluded, and in fact it is highly probable, that some of the effects seen are due to the overexpression or lack of PCDH19. Selected genes of interest, such as the mentioned *Xlr* genes that are overexpressed only in CYTO, will have to be validated with alternative methods. Normally, candidate genes of interest identified via RNAseq analysis are validated by qPCR in the same system. Doing qPCR on the same CYTO and KO cells though would not resolve if the differences are due to overexpression/lack of PCDH19 or to the epigenetic landscape of the cells. In order to address this, a different system would have to be used. For example via nucleofection of PCDH19-CYTO-HA in primary neuron cultures followed by qPCR for the candidate genes. Details and proposal of other approaches for potential future experiments will be discussed in (**Section 7.6**).

Despite the discussed limitations of this experiment, valuable information could be extracted by pathway analysis. Enrichment analysis revealed the presence of many differentially expressed genes for both CYTO and KO associated with neurite outgrowth. In particular several Semaphorins and several Netrin receptors were differentially expressed. As part of this list we also found *Xlr3b*, which was CYTO specific, and has a role in dendrite morphogenesis for layer 2/3 neurons. Via GSEA, other interesting information could be extrapolated from the data. GSEA showed many more significant GO terms and KEGG pathways for CYTO *vs* WT at CA8 than for any other comparison, reflecting the number of DEGs and the substantial overexpression of PCDH19-CYTO at this time point. Interestingly, several terms or pathways were associated with previously described functions of PCDH19. For example the KEGG "*estrogen signaling pathway*", previously reported to be modulated by PCDH19 (Pham et al., 2017) and altered in *PCDH19*-epilepsy patients (Tan et al., 2015) (Trivisano et al., 2017), was significant at CA8 for CYTO *vs* WT and was found to be suppressed. Moreover, the Estrogen Receptor Alpha (*Esr1*), which is part of the estrogen signalling pathway and was shown to regulate gene expression with PCDH19, was downregulated in CYTO *vs* WT at CA8. Of the previously reported genes that were altered following PCDH19 interaction with NONO: *AKR1C3*, *APOD*, *ENC1* and *OXTR* none of them were significantly differentially expressed in either of the two PCDH19 mutant cell lines. It is

important to note, though, that those experiments were carried out in human cancer cell-lines, as opposed to mESC-derived progenitors and neurons.

Interestingly, "*regulation of synaptic transmission, GABAergic*" was also found to be activated, somewhat in line with the previously reported interaction between PCDH19 and the GABA-A receptor (Bassani et al., 2018), (Serratto et al., 2020). In fact, PCDH19 can bind via its cytoplasmic domain to the GABA-A receptor α-1 subunit and modulate GABAergic transmission, by altering surface levels of the receptor. Although not significant in the differential expression analysis, *Oxtr*, the oxytocin receptor, mentioned above as one of the genes modulated by PCDH19 (Pham et al., 2017), appears in the GSEA as part of this GO term, with a $\log_2$ fold change of 0.85.

The "*Wnt signaling pathway*" was suppressed in both CYTO and KO, which is interesting because there is a long-standing relationship between protocadherins and Wnt signalling. For instance, via its intracellular domain, PcdhγC3, can bind and sequester Axin1 at the membrane, a component of the canonical Wnt pathway, competing with Dishevelled, resulting in reduced levels of phosphorylated LRP6, hence inhibiting Wnt signalling (Mah et al., 2016). More recently, the δ-Pcdhs, including PCDH19, have been linked to Wnt signalling via interaction with Ryk receptor (Biswas et al., 2020). δ-Pcdhs can also inhibit Wnt signalling, as in δ-Pcdh zebrafish mutants Wnt signalling is upregulated and cell proliferation is increased in the developing neuroepithelium (Biswas et al., 2020). Although our findings seem to be in opposition to what has been described before, as discussed, GSEA predicts activation or suppression of a pathway based on the $\log_2$ fold change and not on the function of its components within a particular pathway. On the other hand, other studies have shown that Wnt/β-catenin activity can safe-guard epigenetic stability in E14 cells by maintaining methylation of DNA (Theka et al., 2019), hence changes in Wnt signalling can also reflect the status of the epigenetic landscape of ES cells.

The effect of stimulating neuronal activity on DIV12 CYTO, KO and WT neurons on the transcriptome was also investigated. Activity-dependent processing has been reported for other

cadherins (Uemura et al., 2006a) and has been shown to happen for PCDH19 as well (**Chapter 3**). Neurons were stimulated using Bicuculline, a GABA-A antagonist. Overall, PCDH19-CYTO and PCDH19-KO behave similarly in response to BIC treatment and the effect on transcription is much bigger, compared to WT cells. This analysis compared the transcriptome of treated *versus* untreated neurons $\underline{within}$ the same cell line, hence it represents a different analysis in which the changes observed do not necessarily reflect the differences arisen when comparing $\underline{across}$ cell lines. CYTO and KO cells seem to be more susceptible to changes induced by activity or less able to buffer the effects of prolonged stimulation. Interestingly, there was a ”core” set of genes common to all three cell types which represented more than 80% of the WT DEGs, a set of genes, which is consistently changed after 4 hours of bicuculline treatment.

Targeting ESCs and generating ESC-derived progenitors and neurons is a good strategy to investigate transcriptional changes via RNAseq, as it provides a more homogeneous population than would be obtained, for example, by using primary cortical neurons. PCDH19 is only expressed in certain neuronal subtypes (**Figure 1.6**), and currently there are no available tools to isolate PCDH19-expressing neurons. However, any interesting findings obtained this way need to be further investigated in a more physiological setting, if we are to understand the role of PCDH19 and its cytoplasmic domain *in vivo*. Therefore, and based on the results of the RNAseq, it was decided to investigate *in vivo* the effect of PCDH19-CYTO overexpression on neuronal dendritic morphology and spines by *in utero* electroporation. Unfortunately, due to a lack of time and several experimental issues, this study could not be completed in time for the submission of this thesis, but it is ongoing and will be complemented with an electrophysiological analysis in the future.

6.3.1 Conclusion

Transcriptome analysis of PCDH19-CYTO, PCDH19-KO and WT progenitors and neurons was carried out, which confirmed PCDH19-CYTO overexpression and the presence of *Pcdh19* mRNA in the KO cells. Overall, CYTO overexpression leads to the biggest changes at all stages, but

particularly at CA8. Enrichment analysis revealed many of the DEGs belong to pathways involved in neurite morphogenesis. Moreover, CYTO and KO neurons were more susceptible than WT to changes induced by neuronal activity, when stimulated via bicuculline treatment. However, results need to be carefully interpreted because of a potential alteration in the methylation status of targeted ESCs compared to controls. Following up from the RNA sequencing results, an investigation into the role of CYTO overexpression on the morphology of layer II/III neurons is currently ongoing.

Chapter 7

General discussion

This thesis investigated the proteolytic processing and potential nuclear function of the epilepsy-linked cell-adhesion molecule protocadherin-19 (PCDH19). Using a mixture of *in vitro* and *ex vivo* approaches it was determined that PCDH19 is proteolytically processed and that PCDH19 can interact with nuclear import proteins. An *in vitro* model was developed and tested to investigate the effect of the overexpression of PCDH19 intracellular domain on gene expression, which resulted with the identification of several differentially expressed genes involved in neurite outgrowth and synaptic function. Moreover, an *in vivo* method was developed to assess the function of PCDH19-CYTO on neuronal morphology.

7.1 Main findings on proteolytic processing of PCDH19

PCDH19 processing was investigated *in vitro* using mouse embryonic fibroblasts and mESC-derived neurons as a model. It was determined that PCDH19 can be processed into at least one extracellular 70 kDa N-terminal fragment (PCDH19-NTF) and two intracellular C-terminal fragments (PCDH19-CTF1 and PCDH19-CTF2) just over and below 50 kDa respectively. Two proteases potentially involved in this processing, ADAM10 and the γ-secretase complex, were selected via candidate approach based on a review of the existing literature. In MEFs, increased

intracellular calcium, brought about by ionomycin treatment, led to the accumulation of the first cytoplasmic fragment (PCDH19-CTF1) and, based on the observations that PCDH19-CTF1 intracellular fragment is reduced in ADAM10-KO MEFs and in the presence of ADAM10 inhibitors, it was demonstrated that PCDH19 can be cleaved by ADAM10. Results obtained were further validated using N-Cadherin processing as a read-out for successful inhibitor treatments, as N-Cadherin processing has been well characterized (Uemura et al., 2006b), (Reiss et al., 2005). Whilst PCDH19-CTF1 is the most prominent fragment generated post calcium-influx in MEFs, in mESC-derived-neurons the dominant fragment is PCDH19-CTF2. Moreover, in mESC-derived neurons, although treatment with ADAM10 inhibitors reduces the generation of PCDH19-CTF1, it does not impact production or accumulation of CTF2, suggesting the involvement of other proteases in this cellular context. Therefore, whilst in MEFs CTF2 seems to be immediately degraded, in neurons, CTF2 is more stable, suggesting a neuronal-specific function of the fragment. Using a specific γ-secretase inhibitor, it was determined that in mESC-derived neurons this complex is responsible for the production of CTF2, and that PCDH19-CTF2 is, at least in part, generated from PCDH19-CTF1. Finally, it was found that in mESC-derived neurons, proteolytic processing of PCDH19 is activity-dependent and can be stimulated through the activation of the NMDA receptor, with the effect being blocked by simultaneous treatment with NMDA receptor antagonists.

7.2 Main findings on nuclear function of PCDH19

Mass spectrometry analysis carried out on E11 forebrain and adult cortex samples derived from C57BL6 and *Pcdh19* KO animals led to the identification of several potential novel binding partners of PCDH19. Known interactors of PCDH19 were also identified, strengthening the validity of the findings. A gene ontology enrichment analysis of the list of potential interactors suggests that PCDH19 plays a role at excitatory synapses, but more work is necessary in order to understand the function of PCDH19 at the synapse. Furthermore, it was shown that PCDH19 is involved in nucleo-cytoplasmatic transport and can interact with nuclear import and export proteins. The interaction between PCDH19 and one of the identified importins was further validated *in vitro* and

would explain the nuclear localisation of overexpressed PCDH19-CYTO, due to the nuclear localisation signals in the intracellular portion of the protein. In order to investigate the potential nuclear function of the intracellular domain of PCDH19, two mESC lines were generated: one PCDH19-KO cell-line via CRISPR/Cas9 and one PCDH19-CYTO-HA cell line via ZFN targeting. PCDH19-KO and PCDH19-CYTO cells were validated based on expression levels of PCDH19 and differentiated into cortical-like neurons. RNA sequencing analysis on progenitors and neurons derived from PCDH19-CYTO, PCDH19-KO and WT mESCs identified many differentially expressed genes across three time points, with the biggest differences seen at the progenitor stage. Many of the differentially expressed genes in both CYTO and KO were genes with known functions in the process of neurite outgrowth and synapse assembly and plasticity. A preliminary attempt to investigate *in vivo* the effect of PCDH19-CYTO overexpression on the morphology of cortical neurons is ongoing in the Martinez-Garay lab.

7.3 Proposed function of PCDH19-CTF2

Speculating on the preliminary findings gathered throughout this project, we propose a novel nuclear function for PCDH19. We suggest that in cortical neurons PCDH19 is processed in response to neuronal activity. Several proteases that reside at the post-synaptic density are activated in response to neuronal activity and their localisation is changed in response to activity. For example, ADAM10 levels at the post-synaptic density are altered by clathrin-mediated endocytosis, and whilst LTP stimulates endocytosis of ADAM10, LTD results in increased membrane reinsertion of the protease (Marcello et al., 2013) (Malinverno et al., 2010). Similarly, γ-secretase is located at synapses (Schedin-Weiss et al., 2016) (Restituito et al., 2011) and can also modulate synaptic plasticity, as γ-secretase inhibitors can inhibit LTP (Chen and Behnisch, 2013).

We propose that the intracellular fragment PCDH19-CTF2, which is generated by γ-secretase proteolytic processing, is transported into the nucleus. Nuclear translocation of PCDH19-CTF2 would happen via interaction with importin-α, which is kept at the post-synaptic density via binding

to the glutamate receptor until a calcium influx event (Jeffrey et al., 2009) (**Figure 7.1**). We believe that PCDH19-CTF2 has a nuclear function and can modulate gene expression in response to neuronal activity via interaction with transcription factors, ultimately serving as a mechanism to mediate synaptic plasticity. Similar mechanisms to influence transcription either directly or indirectly have been described for other synaptic cell-adhesion molecules (Nagappan-Chettiar et al., 2017) (Inoue et al., 2009). For instance, the product of N-Cadherin processing, N-Cad-CTF2, keeps CREB-dependent transcription in check, by sequestering of CBP, which is needed for CREB mediated gene transcription (Marambaud et al., 2003), (Uemura et al., 2006a). It cannot be excluded that PCDH19 processing might also have an impact on synaptic plasticity via generation of C-terminal fragments that could sequester binding partners in the cytoplasm, for example via interactions through the WIRS domain or through the CM1/CM2. Nonetheless, a cytoplasmic function of CTF2 is less likely, as we have observed in several systems the consistent accumulation of PCDH19-CYTO in the nucleus, suggesting that PCDH19-CTF2 does not linger in the cytoplasm, but is readily translocated into the nucleus. Although the function of PCDH19 during cortical development was not directly investigated in this thesis, because PCDH19 is expressed at different points during corticogenesis, it is also possible that PCDH19 processing is important for circuit assembly and function. Proteolytic processing by γ-secretase of several cell-adhesion molecules is essential for circuit formation. For instance PS1 cleaves DCC, coordinating the interplay between DCC/Netrin and Slit/Robo signalling in the formation of motor circuits (Bai et al., 2011), reviewed by (Bai and Pfaff, 2011).

The case of DSCAM1

Down syndrome cell adhesion molecule (DSCAM) receptors are part of the immunoglobulin superfamily, and although structurally unrelated to protocadherins, they have evolved a parallel function to clustered protocadherins in mediating dendritic self-avoidance during neuronal circuit assembly (Jin and Li, 2019). In the invertebrate *Drosophila melanogaster*, which lacks protocadherins, DSCAM is extensively alternatively spliced, generating extraordinary molecular diversity (Schmucker et al., 2000). In humans, only two paralogous exists, DSCAM and DSCAML (Ya-

makawa et al., 1998), which are not subjected to the same alternative splicing as in *Drosophila*. DSCAM is located on chromosome 21, and in Down Syndrome it is expressed in 3 copies. DSCAM gene dosage is thought to be essential, and when altered, it is believed to contribute to the pathology (Amano et al., 2004). DSCAM mediates both homophilic and heterophilic binding to neighbouring cells and is involved in several processes during development. A recent report investigated the proteolytic processing and nuclear function of DSCAM (Sachse et al., 2019), and the mechanisms they describe are reminiscent of what we have observed in this thesis. DSCAM is processed by γ-secretase and the DSCAM-ICD was shown to bind importin 5 which mediates its translocation to the nucleus. DSCAM-ICD modulates transcription of several genes, including resulting in differential expression of many genes involved in neuronal circuit formation and function. One of the upregulated genes was the netrin receptor *Unc5a*, which, interestingly, was also found to be upregulated in PCDH19-ICD overexpressing cells in our RNA-seq experiment. Overexpression of DSCAM-ICD was found to have an impact on neurite outgrowth and synapse assembly, as DSCAM-ICD transfected primary neurons have less synapses (Sachse et al., 2019). The authors speculated that translocation of DSCAM-ICD to the nucleus increases the levels of secreted factors capable of inhibiting synapse formation and of axon repulsive factors.The basic molecular mechanisms underpinning the neuronal function of DSCAM are similar to what we have observed for PCDH19, and might therefore be helpful to increase our understanding of the function of PCDH19-ICD during development and *in vivo*.

7.4 Findings in the context of *PCDH19*-epilepsy

Both increased and decreased proteolytic processing can underpin certain forms of epilepsy. For instance, disruption of proteolytic activity of specific proteases can lead to epileptic phenotypes and both BACE-1 KO and ADAM10 conditional KO mice present with seizures (Hitt et al., 2010), (Prox et al., 2013). On the other hand, in some cases of intractable epilepsy, an upregulation of matrix metalloproteinases (MMPs) has been reported in human tissue (Konopka et al., 2013) and MMP inhibitors, such as IPR-179, have been demonstrated to have an anti-epileptic effect in

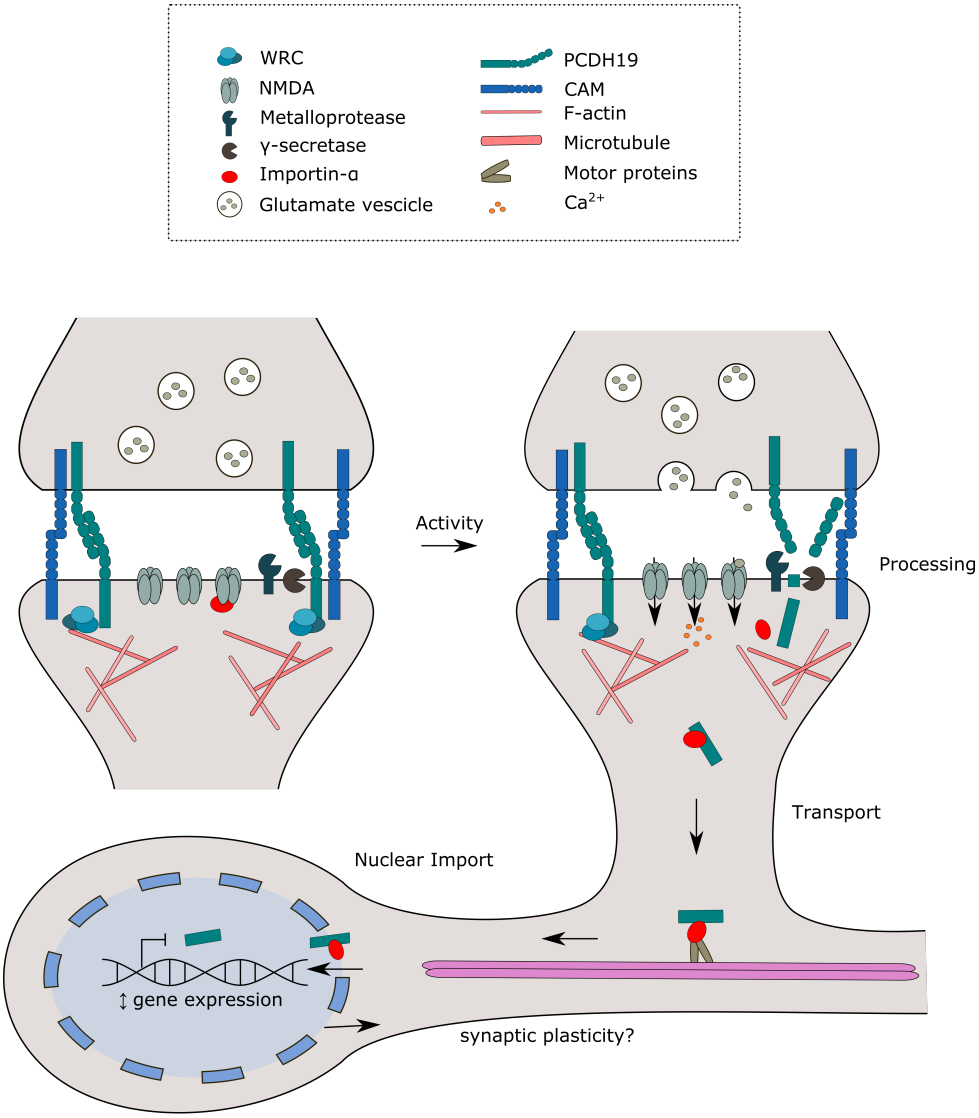


Figure 7.1: PCDH19 processing and gene regulatory function in neurons. Hypothesised role of PCDH19 in regulating synaptic plasticity in neurons. After γ-secretase processing, CTF2 is released from the plasma membrane and retrogradely transported by importin-α into the nucleus. In the proposed model PCDH19-CTF2 can modulate gene expression which ultimately impacts on synaptic plasticity and dendritic morphology.

animal models (Broekaart et al., 2021). Of course, proteases have multiple substrates and it is difficult to tease apart individual effects, with phenotypes being usually due to a cumulative effect on several substrates. For example, postnatal disruption of ADAM10 using a CaMKIIα-Cre deleter causes seizures, which are caused by disrupted synaptic function due to defective shedding of essential synaptic adhesion proteins such as N-Cadherin, Nectin and APP, amongst others (Prox et al., 2013). The result of ADAM10 cKO is altered dendritic spine morphology, reduction in glutamatergic NMDA receptors and reduction of dendritic spines (Prox et al., 2013).

It is possible that defects in PCDH19 processing, or rather, an imbalance in the downstream effects of PCDH19 processing could augment the *PCDH19*-epilepsy phenotype, particularly if PCDH19 cleavage happened only after *trans* homophilic interactions with a PCDH19 expressing neighbouring cell. In a PCDH19 heterozygous setting, although mosaic synapses might be less likely, PCDH19 molecules could be accumulating at synapses, unable to be processed and increasing any synaptic defects. Alternatively, PCDH19 could be excluded from PCDH19 WT-KO synapses and not processed after activity, resulting in reduced levels of PCDH19-CTF2 with a consequent impaired downstream effect on plasticity. Once we understand the functional role of PCDH19 processing in synaptic biology we will start to understand what are the consequences of altered levels of PCDH19 and its processing in the heterozygous brains and if PCDH19 processing could be contributing to the disease phenotype of *PCDH19*-epilepsy, which is currently explained by the ”cellular interference hypothesis”, as a wiring/ miscommunication of cells. First of all, in order to understand if PCDH19 processing is contributing to the pathology it needs to be determined if and when processing is happening *in vivo*, and at what stages during development and/or in adulthood. We also need to determine the levels of the processed fragments and if they are altered between WT and mutant animals (of course taking into account the amount of starting material, the full length protein). We have seen that in neurons, PCDH19 processing is triggered by neuronal activity but it remains to be investigated what triggers processing, if it happens early during neuronal development, when there is yet no activity. A possibility is that PCDH19 processing could be stimulated or triggered by homophilic binding with other PCDH19 molecules on

neighbouring cells. In the case of Notch-Delta signalling, cleavage only happens after binding, as the normally buried extracellular cleavage site is exposed by a substantial conformational change after ligand-receptor interaction (De Strooper et al., 1999), (Gordon et al., 2007). Investigations into the behaviour of progenitors in the PCDH19 heterozygous brain have established that, on top of their striking columnar segregation, PCDH19 WT and PCDH19 KO progenitors within the same heterozygous brain behave differently (Dr.Jessica Griffithis, unpublished). In particular, at E11.5 PCDH19 KO progenitors divide asymmetrically more often than PCDH19 WT progenitors, producing more neurons, earlier. These *in vivo* findings have been confirmed by the *in vitro* study of *PCDH19*-epilepsy patient-derived hIPSCs, which shows that loss of PCDH19 accelerates neurogenesis (Homan et al., 2018). PCDH19 processing in neuronal progenitors could therefore be a mechanism of activating downstream signalling pathways that instruct cell-fate decision and promote a switch between symmetric and asymmetric division. PCDH19 WT and PCDH19 KO progenitors in the PCDH19 heterozygous brain could be behaving differently because of the activation/inactivation of different downstream signalling pathways, which, because of the PCDH19 cellular mosaic, could fail to be activated by lack of PCDH19 homophilic binding.

7.5 Limitations of mESC-derived neuron culture

This thesis used mouse embryonic stem cell (ESC) culture as a model, a technique that has both advantages and limitations. One of the main advantages of ESCs is that they can provide an unlimited source of neurons and protocols have been developed to differentiate ESCs toward multiple neuronal fates including, and not limited to, motor neurons (Wichterle et al., 2002), cortical GABAergic interneurons (Maroof et al., 2010), (Maroof et al., 2013) and cortical glutamatergic neurons (Bibel et al., 2004). Via mESC neuronal differentiation homogenous cultures are produced, which simplify experimental design and interpretation. mESC-derived neurons resemble their *in vivo* counterparts as they become electrically active after several days in culture and can form functional networks as seen by multielectrode array (Ban et al., 2007). mESCs can be easily

manipulated *in vitro* at the genomic level, for example via nucleofection of CRISPR/Cas9 and gRNAs for the target of interest, in order to create genetically engineered stable lines, with specific features. For example, the *Rosa*-TomRITVA mESC line has been engineered to express the TVA receptor and can be used for *in vitro* circuit tracing experiments with the pseudotyped modified rabies virus (Garcia et al., 2012). The Tau-GFP line is useful as once the mESCs are differentiated into neurons, it is possible to visualise them via GFP fluorescence which is expressed from the *Mapt* (*tau*) locus (Tucker et al., 2001). ESC-derived neurons can be transfected with viruses or constructs in order to induce genetic modifications, to manipulate gene expression, to overexpress proteins of interest or to induce reporter activity. ESC-derived neurons are also very easily treated with compounds, in a scalable way and can therefore be used for drug screening applications. Overall ESCs and ESC-derived neurons constitute a simple, easily manipulatable system. Nonetheless, ESCs can, under certain conditions, be epigenetically unstable (Dean et al., 1998), (Humpherys et al., 2001), (Lee et al., 2018) and accumulate chromosomal abnormalities *in vitro* from early passages (Gaztelumendi and Nogués, 2014), which will impact downstream analysis if not controlled for correctly. Most of the features described above for mouse ESCs also apply to human cells which are available, both embryonically derived and reprogrammed induced pluripotent stem cells (iPSCs). One advantage of iPSCs is that they can be derived from patients with particular genomic conditions. In fact, iPSCs from *PCDH19*-epilepsy patients have been derived (Homan et al., 2018) and have been a valuable tool of investigation. The *PCDH19*-epilepsy iPSCs cells were differentiated into cortical-neurons using an adherent monolayer culture protocol, in the traditional 2D way (Shi et al., 2012), but other protocols have been generated to grow iPSCs in suspended 3D culture in order to generate organoids. Organoids can form self-organising complex structures that recapitulate developmental milestones and display both layer organisation and connectivity (Gordon et al., 2021). Organoids can be used to model interactions between different brain regions, for example they have been used to study interneuron migration by fusion of dorsal and ventral forebrain organoids (Bagley et al., 2017). Remarkably, functional neuronal circuits can be formed *in vitro*, for example by fusion of a cortical and a striatal organoid (Miura et al., 2020). In order to understand and model human neurodevelopmental disorders *in vitro*, hIPSC are currently

the most appropriate tool, given the development of the described different protocols. The protocol used in this thesis generates a generic ”cortical-like” population of neurons. Therefore, it does not recapitulate cortical architecture nor does it include all the cellular diversity of the cerebral cortex. Nonetheless, because of the homogeneity of the cells produced, it is a valuable tool for bulk analysis, such as lysis for downstream applications like western blot and bulk RNA-sequencing, as used in this thesis. Although, with the advancement of single-cell RNA-sequencing technology, the advantages of adopting more complex *in vitro* systems for the investigation of human neurodevelopmental disorders are ever increasing.

7.6 Future directions

The next sections will discuss specific proposed experiments for future investigation which would be a direct continuation of the experiments presented in this thesis.

Follow-up on proteolytic processing

Although we have observed nuclear localisation of overexpressed PCDH19-ICD in several systems (HEK293 cells, ESC-derived neurons and layer II/III *in utero* electroporated neurons), the nuclear translocation of CTF2, the fragment with supposed nuclear function that is generated by proteolytic processing, remains to be experimentally observed. In order to address this question and determine if CTF2 can in fact translocate to the nucleus different approaches could be used.

A more sophisticated fractionation method could be implemented in order to effectively separate nuclear, cytoplasmic and membrane fractions. Cells, either MEFs or mESC-derived neurons, that have been used throughout this thesis, could be treated with compounds that increase intracellular calcium, in order to stimulate processing. Lysing and fractionation could be done at several intervals post treatment in order to detect fragment movement through different cellular compartments. This would be a relatively simple experiment as it does not require any new cell lines.

Alternatively, it would be interesting to be able to visualise nuclear translocation of PCDH19 via immunocytochemistry. Because PCDH19 antibodies are not reliable for immunocytochemistry or immunohistochemistry, a tagged version of PCDH19 should be used. As seen in the subcellular localisation experiments (**Chapter 3**), overexpression of PCDH19 via transfection in HEK293 cells leads to accumulation of PCDH19 in the ER, which could compromise experimental success. Moreover, overexpressing proteins always requires careful interpretation of results, so whenever possible, it is much more informative to study the endogenous protein. In order to do this, a new mESC line could be generated, in which PCDH19 would be tagged with a tag that is easily detected via ICC. This could be achieved by CRISPR/Cas9 targeting of the 3' end of PCDH19, to insert an HA tag or another type of tag such as MYC. A mESC line of this sort would be in fact very useful, not only for this experiment. The PCDH19-HA mESC line could be differentiated into cortical-like neurons, as previously done, treated with ionomycin, to induce processing, and neurons could be processed via ICC to visualise possible changes in HA signal in different cellular compartments. This cell system could also be used to test the effect of different compounds, such as protease inhibitors or stimulators of neuronal activity, on the subcellular localisation of PCDH19-HA. This experiment can also be done in combination with inhibitors of protein synthesis, to confirm that translocation is due to processing and not to accumulation of newly synthesised protein.

This question of PCDH19-CTF2 nuclear translocation could also be tackled with a different approach. PCDH19 could be C-terminally fused to a transactivator domain for a reporter plasmid, in order to detect nuclear translocation post processing via activation of a reporter. For example, PCDH19 could be fused to a Gal4 DNA binding domain and a VP16 transactivation domain. Gal4 would specifically bind to a UAS (upstream activator sequence) placed upstream of luciferase in a reporter vector, as previously described (Karlström et al., 2002), (Sachse et al., 2019). Cells would be transfected in parallel with the PCDH19-Gal4/ VP16 construct and a luciferase reporter plasmid and treated with ionomycin in order to stimulate processing, which would generate a

PCDH19-CTF2-Gal4/VP16 fusion fragment. If processing is followed by nuclear translocation of the generated fragment, it would be reflected by increased luciferase signal. A non-cleavable IFNaR-Gal4DBD-VP16 receptor and a constitutively cleaved IFNaR2-Gal4bd-VP16 receptor or a constitutively active control Gal4bd-VP16 could be used as negative and positive controls respectively, as previously published (Sachse et al., 2019).

Another open question derived from the results obtained on PCDH19 processing concerns the ectodomain shedding of PCDH19 and the role of ADAM10 or other proteases. Our results show that ADAM10 seems to cleave PCDH19, but other sheddases must be involved, as in mESC-derived neurons production of PCDH19-CTF2 is not affected by treatment with an ADAM10 specific inhibitor. Broad-spectrum inhibitors could initially be used to narrow down what families of proteases are involved in PCDH19 processing, utilising the same *in vitro* system described in this thesis. PCDH19 processing might be mediated by more than two proteases, there could be several pathways, leading to slightly different processed fragments. For instance, calcium activated proteases like calpain could be playing a role, especially as processing is triggered by increased intracellular calcium. Calpains are calcium-dependent cysteine proteases, which can be found in synaptic terminals of neurons and can be activated by calcium entry through activation of NMDA receptors (Vanderklish et al., 1995). Calpains can cleave many known substrates, including cell-adhesion molecules. In fact, N-Cadherin can also be cleaved by calpain, in addition to being cleaved by ADAM10 and γ-secretase (Jang et al., 2009). Calpains therefore have an important role at the synapse and in modulating synaptic plasticity. We have preliminarily tested the effect of calpain inhibitors on PCDH19 processing in mESC-derived neurons and have seen a slight accumulation of CTF1, suggesting that calpain and γ-secretase could be acting in parallel (data not shown).

It could also be interesting to generate mutant constructs of PCDH19 with mutated amino acids near the supposed cleavage site close to the transmembrane domain, in order to pin point the cleavage site. If the cleavage site is identified, it might also be interesting to generate a

construct with an uncleavable version of PCDH19. This uncleavable PCDH19 could be used as a negative control for the nuclear translocation experiments or to generate stable ESC-line and investigate the effect of impaired PCDH19 processing in cortical-like neurons.

Experiments need to be done in order to investigate the role of PCDH19 in synapse formation and in activity dependent processing. For instance, it would be interesting to see if synaptic strength is altered after PCDH19 processing. N-Cadherin processing at synapses alters strength of synaptic contacts and influences the size of dendritic spines (Malinverno et al., 2010). Considering that PCDH19 can form a complex with N-Cadherin (Biswas et al., 2010), that N-Cadherin is found in synapses (Malinverno et al., 2010) and that at least a fraction of PCDH19 is also localised to synapses (Pederick et al., 2016), (Hayashi et al., 2017), (Bassani et al., 2018), (Mincheva-Tasheva et al., 2021) it would be very interesting to determine if the N-Cadherin/PCDH19 complex is present in glutamatergic synapses and whether the N-Cadherin/PCDH19 complex could alter the strength of synaptic contact between neurons. An appealing possibility would be an effect of the complex on the susceptibility of individual versus complexed proteins to activity-dependent processing.

In (**Chapter 3**), the focus was on the cytotoplasmic fragments of PCDH19, as we were interested in determining a potential nuclear function, but it is important to note that the extracellular domain, PCDH19-NTF, which is released by shedding into the extracellular space could also have a function. Extracellular fragments of other protocadherins can bind to cell-adhesion molecules and receptors on neighbouring cells. For example, N-Cadherin-NTF can bind to FGFR and can stimulate neurite outgrowth (McCusker and Alfandari, 2009), (Utton et al., 2001). PCDH19-NTF could bind PCDH19-FL on neighbouring cells, antagonising the standard PCDH19-PCDH19 homophillic adhesion.

Follow-up on the interactome study

Regarding the mass-spectrometry analysis, the focus for this thesis was on the identified nuclear transporters, the importins, but the dataset contains a lot of unexplored data and several potential interactors of PCDH19 that might be interesting to investigate. For instance, many of the proteins identified in the mass spectrometry analysis were synaptic proteins and, in particular, components of the excitatory synapse. The most recent paper on PCDH19 suggests that PCDH19 is important for synapse formation (Mincheva-Tasheva et al., 2021). Moreover, in the mass spectrometry analysis we also identified many members of the cadherin superfamily. In order to understand how PCDH19 works as a cell adhesion molecule, we need to take into consideration other cadherins as well. In fact, heterophilic *cis* interaction between cadherins creates complexes with different adhesion properties. γ-Pcdhs, for example, can form zipper-like arrays with different types of γ-Pcdhs. *Trans* interactions between different δ-Pcdhs and how these impact on cell adhesion are starting to be elucidated (Bisogni et al., 2018) and although δ-Pcdhs do not interact in *cis* via the extracellular domain (Harrison et al., 2020), there is the possibility of interactions happening via the transmembrane or the cytoplasmic domain, but these have not been described yet. Regarding the results on PCDH19 interactions with importins, a mutant construct of PCDH19-ICD lacking the main NLS, could be cloned to see if nuclear localisation is lost when removing the presumed NLS, as we assume that the interaction between PCDH19 and KPNA1 happens via the NLS. As other importins were also identified by MS, other co-immunoprecipitation experiments with PCDH19 and PCDH19-CYTO and these other importins could also be performed.

Follow-up on RNA sequencing results

The RNA sequencing of CYTO, KO and WT progenitors and neurons generated a substantial amount of data which could be useful for future investigations. First of all, it provides a helpful characterization of the transcriptome of E14 WT cells at different stages during the differentiation protocol. Hence it provides a convenient reference for future experiments in the same cellular system. As discussed in the previous chapter, a more appropriate control could have been used

for the RNA-sequencing analysis, namely WT ESCs with a similar 'history' to the targeted clones. WT cells should have been subjected to the same processes of FACS/colony picking in parallel to the targeted clones and kept in culture for the equivalent number of passages that was needed to generate the CYTO and KO clones. Moreover, WT cells would have also had to be grown on MEFs as this could have had an impact on the transcriptome of targeted clones. One way to address this problem could be to generate new WT data, for use as a control data-set for the differential expression analysis, after a mock targeting process. But ideally, this would have been done in parallel, to limit experimental variability.

Nonetheless, many genes were differentially expressed and amongst them, as validated by *Pcdh19* expression in CYTO, some genes must be altered due to the overexpression of CYTO. As RNAseq follow up experiments, some of the target genes would have to be confirmed by qPCR, to validate the model. This could be carried out on the same mESC-derived neurons, PCDH19-CYTO, PCDH19-KO and WT cells, or in a different system. Nucleofection of primary cortical neurons with PCDH19-CYTO-HA and control empty vector could be performed and followed by sequencing or qPCR of selected target genes. Morphological analysis could also be done *in vitro*. Another way of addressing this question would be to perform *in utero* electroporation, as previosly done of pCIG-PCDH19-CYTO-HA, followed instead by isolation of the electroporated cells by FACS, and RNA extraction for sequencing or qPCR of candidate genes. It would be interesting to determine the degree of overlap in differentially expressed genes obtained by *in vitro* and *in vivo* overexpression of PCDH19-CYTO. If the morphological analysis of layer II/III neurons results in some differences due to PCDH19-CYTO overexpression, it could also be interesting to attempt *in utero* electroporation of constructs containing genes that were upregulated in the RNA sequencing analysis, for example encoding for *Xlr* genes and see if that overexpression can mimic the phenotype brought about by overexpression of PCDH19-CYTO-HA.

7.7 Open questions on PCDH19

PCDH19 is involved in diverse cellular processes, from neurogenesis and cell-sorting of progenitors in the developing cortex to cortical neuron migration, circuit and synapse formation and modulation of network activity. The precise mechanisms by which PCDH19 is involved in these processes remains unexplored and somewhat controversial. For instance, although migration defects have been reported in PCDH19-KO cortical explants, this is not reflected by differences in cortical lamination (Pederick et al., 2016), (Galindo-Riera et al., 2020). Regarding synapse formation, the latest report suggests there are defects in synaptic assembly between PCDH19-WT and PCDH19-KO neurons in a mosaic co-culture (Mincheva-Tasheva et al., 2021), but these observations were not backed up by findings by Dr. Jessica Griffiths from the Martinez-Garay lab who did not observe any changes in number of synaptic contacts in a different *in vitro* system (unpublished). The discrepancy could be explained by the fact that in the first case, the experiment was carried out analysing single isolated PCDH19 WT neurons in a PCDH19 KO culture or a single PCDH19 KO neurons in a WT culture. In the other case, synaptic connectivity was analysed in a 1:1 WT:KO mosaic co-culture. In order to understand the pathophysiology of *PCDH19*-epilepsy, further investigations are necessary on cortical connectivity and function, using both *in vitro* and *in vivo* mosaic systems. This could be achieved in different ways. Several experiments that address these questions and have not been discussed in this thesis are currently ongoing in the lab and include *in vivo* calcium imaging of cortical activity and circuit tracing using the modified rabies virus, both done in a mouse model of *PCDH19*-epilepsy.

7.8 Concluding remarks

This thesis describes the novel proteolytic processing of PCDH19 as a mechanism of membrane-to-nucleus signal transduction in neurons, and begins to unravel the nuclear function of PCDH19 expanding the existing knowledge which was limited to the interaction between the paraspeckle protein NONO and PCDH19 in modulating ER-α dependent gene expression (Pham

et al., 2017). For the first time, it was also shown that proteolytic processing of PCDH19 can be triggered by neuronal activity and it can be mediated through the NMDA receptor. We proposed that the activity-dependent processing of PCDH19 is a mechanism necessary for the modulation of synaptic plasticity, via downstream effects on gene expression mediated by the intracellular fragment PCDH19-CTF2. Further work will be necessary in order to unravel the mechanisms of proteolytic processing *in vivo*, both in adulthood and during development, and to understand how PCDH19 processing contributes to the pathology of *PCDH19*-epilepsy.

Bibliography

Aaku-Saraste, E., Hellwig, A., and Huttner, W. B. (1996). Loss of occludin and functional tight junctions, but not ZO-1, during neural tube closure - Remodeling of the neuroepithelium prior to neurogenesis. *Developmental Biology*, 180(2):664–679.

Alberi, L., Liu, S., Wang, Y., Badie, R., Smith-Hicks, C., Wu, J., Pierfelice, T. J., Abazyan, B., Mattson, M. P., Kuhl, D., Pletnikov, M., Worley, P. F., and Gaiano, N. (2011). Activity-Induced Notch Signaling in Neurons Requires Arc/Arg3.1 and Is Essential for Synaptic Plasticity in Hippocampal Networks. *Neuron*, 69(3):437–444.

Amano, K., Sago, H., Uchikawa, C., Suzuki, T., Kotliarova, S. E., Nukina, N., Epstein, C. J., and Yamakawa, K. (2004). Dosage-dependent over-expression of genes in the trisomic region of Ts1 Cje mouse model for Down syndrome. *Human Molecular Genetics*, 13(13):1333–1340.

Anders, S., Reyes, A., and Huber, W. (2012). Detecting differential usage of exons from RNA-seq data. *Genome Research*, 22(10):2008–2017.

Anderson, S. A., Eisenstat, D. D., Shi, L., and Rubenstein, J. L. (1997). Interneuron migration from basal forebrain to neocortex: Dependence on Dlx genes. *Science*, 278(5337):474–476.

Angevine, J. B. J. and Sidman, R. L. (1961). Autoradiographic study of cell migration during histogenesis of cerebral cortex in mouse. *Nature*, 192:766–768.

Arikkath, J. and Reichardt, L. F. (2008). Cadherins and catenins at synapses: roles in synaptogenesis and synaptic plasticity. *Trends in Neurosciences*, 31(9):487–494.

Arlotta, P., Molyneaux, B. J., Chen, J., Inoue, J., Kominami, R., and MacKlis, J. D. (2005). Neuronal subtype-specific genes that control corticospinal motor neuron development in vivo. *Neuron*, 45(2):207–221.

Asada-Utsugi, M., Uemura, K., Kubota, M., Noda, Y., Tashiro, Y., Uemura, T. M., Yamakado, H., Urushitani, M., Takahashi, R., Hattori, S., Miyakawa, T., Ageta-Ishihara, N., Kobayashi, K., Kinoshita, M., and Kinoshita, A. (2021). Mice with cleavage-resistant N-cadherin exhibit synapse anomaly in the hippocampus and outperformance in spatial learning tasks. *Molecular Brain*, 14(1):1–16.

Auffret, A., Gautheron, V., Repici, M., Kraftsik, R., Mount, H. T., Mariani, J., and Rovira, C. (2009). Age-dependent impairment of spine morphology and synaptic plasticity in hippocampal CA1 neurons of a presenilin 1 transgenic mouse model of Alzheimer's disease. *Journal of Neuroscience*, 29(32):10144–10152.

Baas, P. W., Deitch, J. S., Black, M. M., and Banker, G. A. (1988). Polarity orientation of microtubules in hippocampal neurons: Uniformity in the axon and nonuniformity in the dendrite. *Proceedings of the National Academy of Sciences of the United States of America*, 85(21):8335–8339.

Bagley, J. A., Reumann, D., Bian, S., Lévi-Strauss, J., and Knoblich, J. A. (2017). Fused cerebral organoids model interactions between brain regions. *Nature Methods*, 14(7):743–751.

Bai, G., Chivatakarn, O., Bonanomi, D., Lettieri, K., Franco, L., Xia, C., Stein, E., Ma, L., Lewcock, J. W., and Pfaff, S. L. (2011). Presenilin-dependent receptor processing is required for axon guidance. *Cell*, 144(1):106–118.

Bai, G. and Pfaff, S. L. (2011). Protease Regulation: The Yin and Yang of Neural Development and Disease. *Neuron*, 72(1):9–21.

Bai, N., Hayashi, H., Aida, T., Namekata, K., Harada, T., Mishina, M., and Tanaka, K. (2013). Dock3 interaction with a glutamate-receptor NR2D subunit protects neurons from excitotoxicity. *Molecular Brain*, 6(1):1–11.

Ban, J., Bonifazi, P., Pinato, G., Broccard, F. D., Studer, L., Torre, V., and Maria Elisabetta, R. (2007). Embryonic Stem Cell-Derived Neurons Form Functional Networks. *Stem Cells*, 25:738–749.

Bao, J., Lin, H., Ouyang, Y., Lei, D., Osman, A., Kim, T. W., Mei, L., Dai, P., Ohlemiller, K. K., and Ambron, R. T. (2004). Activity-dependent transcription regulation of PSD-95 by neuregulin-1 and Eos. *Nature Neuroscience*, 7(11):1250–1258.

Bassani, S., Cwetsch, A. W., Gerosa, L., Serratto, G. M., Folci, A., Hall, I. F., Mazzanti, M., Cancedda, L., and Passafaro, M. (2018). The female epilepsy protein PCDH19 is a new GABAAR-binding partner that regulates GABAergic transmission as well as migration and morphological maturation of hippocampal neurons. *Human Molecular Genetics*, 27(6):1027–1038.

Beel, A. J. and Sanders, C. R. (2008). Substrate specificity of γ-secretase and other intramembrane proteases. *Cellular and Molecular Life Sciences*, 65(9):1311–1334.

Ben-Shem, A., Fass, D., and Bibi, E. (2007). Structural basis for intramembrane proteolysis by rhomboid serine proteases. *Proceedings of the National Academy of Sciences of the United States of America*, 104(2):462–466.

Benavides-Piccione, R., Hamzei-Sichani, F., Ballesteros-Yáñez, I., Defelipe, J., and Yuste, R. (2006). Dendritic size of pyramidal neurons differs among mouse cortical regions. *Cerebral Cortex*, 16(7):990–1001.

Benito, E., Valor, L. M., Jimenez-minchan, M., Huber, W., and Barco, A. (2011). cAMP Response Element-Binding Protein Is a Primary Hub of Activity-Driven Neuronal Gene Expression. *The Journal of Neuroscience*, 31(50):18237–18250.

Bergendahl, L. T., Gerasimavicius, L., Miles, J., Macdonald, L., Wells, J. N., Welburn, J. P., and Marsh, J. A. (2019). The role of protein complexes in human genetic disease. *Protein Science*, 28(8):1400–1411.

Bibel, M., Richter, J., Lacroix, E., and Barde, Y. A. (2007). Generation of a defined and uniform population of CNS progenitors and neurons from mouse embryonic stem cells. *Nature Protocols*, 2(5):1034–1043.

Bibel, M., Richter, J., Schrenk, K., Tucker, K. L., Staiger, V., Korte, M., Goetz, M., and Barde, Y. A. (2004). Differentiation of mouse embryonic stem cells into a defined neuronal lineage. *Nature Neuroscience*, 7(9):1003–1009.

Bisogni, A. J., Ghazanfar, S., Williams, E. O., Marsh, H. M., Yang, J. Y., and Lin, D. M. (2018). Tuning of delta-protocadherin adhesion through combinatorial diversity. *eLife*, 7:1–27.

Biswas, S., Emond, M. R., Chenoweth, K., and Jontes, J. D. (2020). δ-protocadherins control neural progenitor cell proliferation by antagonizing Ryk and Wnt/β-catenin signaling.

Biswas, S., Emond, M. R., Duy, P. Q., Hao, L. T., Beattie, C. E., and Jontes, J. D. (2014). Protocadherin-18b interacts with Nap1 to control motor axon growth and arborization in zebrafish. *Molecular Biology of the Cell*, 25(5):633–642.

Biswas, S., Emond, M. R., and Jontes, J. D. (2010). Protocadherin-19 and N-cadherin interact to control cell movements during anterior neurulation. *Journal of Cell Biology*, 191(5):1029–1041.

Bode, W., Gomis-Rüth, F. X., and Stöckler, W. (1993). Astacins, serralysins, snake venom and matrix metalloproteinases exhibit identical zinc-binding environments (HEXXHXXGXXH and Met-turn) and topologies and should be grouped into a common family, the 'metzincins'. *FEBS Letters*, 331(1-2):134–140.

Bolduc, D. M., Montagna, D. R., Gu, Y., Selkoe, D. J., and Wolfe, M. S. (2016). Nicastrin functions to sterically hinder γ-secretase-substrate interactions driven by substrate transmembrane domain. *Proceedings of the National Academy of Sciences of the United States of America*, 113(5):E509–E518.

Bond, C. S. and Fox, A. H. (2009). Paraspeckles: Nuclear bodies built on long noncoding RNA. *Journal of Cell Biology*, 186(5):637–644.

Bonn, S., Seeburg, P. H., and Schwarz, M. K. (2007). Combinatorial Expression of - and - Protocadherins Alters Their Presenilin-Dependent Processing. *Molecular and Cellular Biology*, 27(11):4121–4132.

Borcel, E., Palczynska, M., Krzisch, M., Dimitrov, M., Ulrich, G., Toni, N., and Fraering, P. C. (2016). Shedding of neurexin 3β ectodomain by ADAM10 releases a soluble fragment that affects the development of newborn neurons. *Scientific Reports*, 6(December):1–12.

Bot, N., Schweizer, C., Halima, S. B., and Fraering, P. C. (2011). Processing of the synaptic cell adhesion molecule neurexin-3β by Alzheimer disease α- and γ-secretases. *Journal of Biological Chemistry*, 286(4):2762–2773.

Bouillot, S., Tillet, E., Carmona, G., Prandini, M. H., Gauchez, A. S., Hoffmann, P., Alfaidy, N., Cand, F., and Huber, P. (2011). Protocadherin-12 cleavage is a regulated process mediated by ADAM10 protein: Evidence of shedding up-regulation in pre-eclampsia. *Journal of Biological Chemistry*, 286(17):15195–15204.

Brasch, J., Goodman, K. M., Noble, A. J., Rapp, M., Bahna, F., Dandey, V. P., Bepler, T., Berger, B., Maniatis, T., Potter, C. S., Carragher, B., and Honig, B. (2019). Visualization of clustered protocadherin neuronal self-recognition complexes. *Nature*, 569(7755):280–283.

Bray, S. J. (2006). Notch signalling: A simple pathway becomes complex. *Nature Reviews Molecular Cell Biology*, 7(9):678–689.

Brodmann, K. (1909). *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Barth.

Broekaart, D. W., Bertran, A., Jia, S., Korotkov, A., Senkov, O., Bongaarts, A., Mills, J. D., Anink, J. J., Seco, J., Baayen, J. C., Idema, S., Chabrol, E., Becker, A. J., Wadman, W. J., Tarragó, T., Gorter, J. A., Aronica, E., Prades, R., Dityatev, A., and van Vliet, E. A. (2021). The matrix

metalloproteinase inhibitor IPR-179 has antiseizure and antiepileptogenic effects. *Journal of Clinical Investigation*, 131(1).

Brummer, T., Pigoni, M., Rossello, A., Wang, H., Noy, P. J., Tomlinson, M. G., Blobel, C. P., and Lichtenthaler, S. F. (2018). The metalloprotease ADAM10 (a disintegrin and metalloprotease 10) undergoes rapid, postlysis autocatalytic degradation. *FASEB Journal*, 32(7):3560–3573.

Brunkan, A. L., Martinez, M., Walker, E. S., and Goate, A. M. (2005). Presenilin endoproteolysis is an intramolecular cleavage. *Molecular and Cellular Neuroscience*, 29(1):65–73.

Buchanan, S. M., Schalm, S. S., and Maniatis, T. (2010). Proteolytic processing of protocadherin proteins requires endocytosis. *Proceedings of the National Academy of Sciences of the United States of America*, 107(41):17774–17779.

Butler, M. G., Rafi, S. K., Hossain, W., Stephan, D. A., and Manzardo, A. M. (2015). Whole exome sequencing in females with autism implicates novel and candidate genes. *International Journal of Molecular Sciences*, 16(1):1312–1335.

Caviness, V. S. (1976). Patterns of cell and fiber distribution in the neocortex of the reeler mutant mouse.

Chang, H., Hoshina, N., Zhang, C., Ma, Y., Cao, H., Wang, Y., Wu, D. D., Bergen, S. E., Landén, M., Hultman, C. M., Preisig, M., Kutalik, Z., Castelao, E., Grigoroiu-Serbanescu, M., Forstner, A. J., Strohmaier, J., Hecker, J., Schulze, T. G., Müller-Myhsok, B., Reif, A., Mitchell, P. B., Martin, N. G., Schofield, P. R., Cichon, S., Nöthen, M. M., Walter, H., Erk, S., Heinz, A., Amin, N., Van Duijn, C. M., Meyer-Lindenberg, A., Tost, H., Xiao, X., Yamamoto, T., Rietschel, M., Li, M., Backlund, L., Frisén, L., Lavebratt, C., Schalling, M., Ösby, U., Mühleisen, T. W., Leber, M., Degenhardt, F., Treutlein, J., Mattheisen, M., Maaser, A., Meier, S., Herms, S., Hoffmann, P., Lacour, A., Witt, S. H., Streit, F., Lucae, S., Maier, W., Schwarz, M., Vedder, H., Kammerer-Ciernioch, J., Pfennig, A., Bauer, M., Hautzinger, M., Wright, A., Fullerton, J. M., Montgomery, G. W., Medland, S. E., Gordon, S. D., Becker, T., Schumacher, J., and Propping, P. (2018). The protocadherin 17 gene affects cognition, personality, amygdala structure and function,

synapse development and risk of major mood disorders. *Molecular Psychiatry*, 23(2):400–412.

Chappuis-Flament, S., Wong, E., Hicks, L. D., Kay, C. M., and Gumbiner, B. M. (2001). Multiple cadherin extracellular repeats mediate homophilic binding and adhesion. *Journal of Cell Biology*, 154(1):231–243.

Chen, B., Brinkmann, K., Chen, Z., Pak, C. W., Liao, Y., Shi, S., Henry, L., Grishin, N. V., Bogdan, S., and Rosen, M. K. (2014). The WAVE Regulatory Complex Links Diverse Receptors to the Actin Cytoskeleton. *Cell*, 156(1-2):195–207.

Chen, Y. Q. and Behnisch, T. (2013). The role of γ-secretase in hippocampal synaptic transmission and activity-dependent synaptic plasticity. *Neuroscience Letters*, 554:16–21.

Chu, V. T., Weber, T., Wefers, B., Wurst, W., Sander, S., Rajewsky, K., and Kühn, R. (2015). Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nature Biotechnology*, 33(5):543–548.

Coleman, H. A., Labrador, J. P., Chance, R. K., and Bashaw, G. J. (2010). The Adam family metalloprotease Kuzbanian regulates the cleavage of the roundabout receptor to control axon repulsion at the midline. *Development*, 137(14):2417–2426.

Conant, K., Allen, M., and Lim, S. T. (2015). Activity dependent CAM cleavage and neurotransmission. *Frontiers in Cellular Neuroscience*, 9(AUGUST):1–15.

Concordet, J. P. and Haeussler, M. (2018). CRISPOR: Intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Research*, 46(W1):W242–W245.

Cooper, S. R., Jontes, J. D., and Sotomayor, M. (2016). Structural determinants of adhesion by protocadherin-19 and implications for its role in epilepsy. *eLife*, 5:1–22.

Cubelos, B., Sebastián-Serrano, A., Beccari, L., Calcagnotto, M. E., Cisneros, E., Kim, S., Dopazo, A., Alvarez-Dolado, M., Redondo, J. M., Bovolenta, P., Walsh, C. A., and Nieto, M. (2010).

Cux1 and Cux2 regulate dendritic branching, spine morphology, and synapses of the upper layer neurons of the cortex. *Neuron*, 66(4):523–535.

Cullot, G., Boutin, J., Toutain, J., Prat, F., Pennamen, P., Rooryck, C., Teichmann, M., Rousseau, E., Lamrissi-Garcia, I., Guyonnet-Duperat, V., Bibeyran, A., Lalanne, M., Prouzet-Mauléon, V., Turcq, B., Ged, C., Blouin, J. M., Richard, E., Dabernat, S., Moreau-Gaudry, F., and Bedel, A. (2019). CRISPR-Cas9 genome editing induces megabase-scale chromosomal truncations. *Nature Communications*, 10(1):1–14.

Cwetsch, A. W., Narducci, R., Bolla, M., Pinto, B., Perlini, L., Bassani, S., Passafaro, M., and Cancedda, L. (2020). Mosaic expression of X-linked PCDH19 Protein by in Utero Electroporation in Rats Replicates Human Cortical and Hippocampal Developmental Abnormalities, Associated Core Behaviors Related to Autism, and Cognitive Impairment. *Bioarchive*, 10.18372/0.

D'Arcangelo, G., Miao, G. G., Chen, S. C., Scares, H. D., Morgan, J. I., and Curran, T. (1995). A protein related to extracellular matrix proteins deleted in the mouse mutant reeler. *Nature*, 374(6524):719–723.

Davies, W., Isles, A., Smith, R., Karunadasa, D., Burrmann, D., Humby, T., Ojarikre, O., Biggin, C., Skuse, D., Burgoyne, P., and Wilkinson, L. (2005). Xlr3b is a new imprinted candidate for X-linked parent-of-origin effects on cognitive function in mice. *Nature Genetics*, 37(6):625–629.

de Lange, I. M., Rump, P., Neuteboom, R. F., Augustijn, P. B., Hodges, K., Kistemaker, A. I., Brouwer, O. F., Mancini, G. M., Newman, H. A., Vos, Y. J., Helbig, K. L., Peeters-Scholte, C., Kriek, M., Knoers, N. V., Lindhout, D., Koeleman, B. P., van Kempen, M. J., and Brilstra, E. H. (2017). Male patients affected by mosaic PCDH19 mutations: five new cases. *Neurogenetics*, 18(3):147–153.

De Strooper, B., Annaert, W., Cupers, P., Saftig, P., Craessaerts, K., Mumm, J. S., Schroeter, E. H., Schrijvers, V., Wolfe, M. S., Ray, W. J., Goate, A., and Kopan, R. (1999). A presenilin-1-dependent γ-secretase-like protease mediates release of notch intracellular domain. *Nature*, 398(6727):518–522.

De Strooper, B., Saftig, P., Craessaerts, K., Vanderstichele, H., Guhde, G., Annaert, W., Von Figura, K., and Van Leuven, F. (1998). Deficiency of presenilin-1 inhibits the normal cleavage of amyloid precursor protein. *Nature*, 391(January):387–390.

Dean, W., Bowden, L., Aitchison, A., Klose, J., Moore, T., Meneses, J. J., Reik, W., and Feil, R. (1998). Altered imprinted gene methylation and expression in completely ES cell-derived mouse fetuses: Association with aberrant phenotypes. *Development*, 125(12):2273–2282.

Depienne, C., Bouteiller, D., Keren, B., Cheuret, E., Poirier, K., Trouillard, O., Benyahia, B., Quelin, C., Carpentier, W., Julia, S., Afenjar, A., Gautier, A., Rivier, F., Meyer, S., Berquin, P., Hélias, M., Py, I., Rivera, S., Bahi-Buisson, N., Gourfinkel-An, I., Cazeneuve, C., Ruberg, M., Brice, A., Nabbout, R., and Leguern, E. (2009). Sporadic infantile epileptic encephalopathy caused by mutations in PCDH19 resembles dravet syndrome but mainly affects females. *PLoS Genetics*, 5(2).

Depienne, C. and Leguern, E. (2012). PCDH19-related infantile epileptic encephalopathy: An unusual X-linked inheritance disorder. *Human Mutation*, 33(4):627–634.

Dibbens, L. M., Tarpey, P. S., Hynes, K., Bayly, M. A., Scheffer, I. E., Smith, R., Bomar, J., Sutton, E., Vandeleur, L., Shoubridge, C., Edkins, S., Turner, S. J., Stevens, C., O'Meara, S., Tofts, C., Barthorpe, S., Buck, G., Cole, J., Halliday, K., Jones, D., Lee, R., Madison, M., Mironenko, T., Varian, J., West, S., Widaa, S., Wray, P., Teague, J., Dicks, E., Butler, A., Menzies, A., Jenkinson, A., Shepherd, R., Gusella, J. F., Afawi, Z., Mazarib, A., Neufeld, M. Y., Kivity, S., Lev, D., Lerman-Sagie, T., Korczyn, A. D., Derry, C. P., Sutherland, G. R., Friend, K., Shaw, M., Corbett, M., Kim, H. G., Geschwind, D. H., Thomas, P., Haan, E., Ryan, S., McKee, S., Berkovic, S. F., Futreal, P. A., Stratton, M. R., Mulley, J. C., and Gécz, J. (2008). X-linked protocadherin 19 mutations cause female-limited epilepsy and cognitive impairment. *Nature Genetics*, 40(6):776–781.

Divekar, S. D., Tiek, D. M., Fernandez, A., and Riggins, R. B. (2016). Estrogen-related receptor β (ERRβ) - renaissance receptor or receptor renaissance? *Nuclear receptor signaling*, 14:e002.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T. R. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1):15–21.

Eiselleova, L., Peterkova, I., Neradil, J., Slaninova, I., Hampl, A., and Dvorak, P. (2008). Comparative study of mouse and human feeder cells for human embryonic stem cells. *International Journal of Developmental Biology*, 52(4):353–363.

Emond, M. R., Biswas, S., Blevins, C. J., and Jontes, J. D. (2011). A complex of protocadherin-19 and n-cadherin mediates a novel mechanism of cell adhesion. *Journal of Cell Biology*, 195(7):1115–1121.

Emond, M. R., Biswas, S., and Jontes, J. D. (2009). Protocadherin-19 is essential for early steps in brain morphogenesis.

Emond, M. R., Biswas, S., Morrow, M. L., and Jontes, J. D. (2021). Proximity-dependent Proteomics Reveals Extensive Interactions of Protocadherin-19 with Regulators of Rho GTPases and the Microtubule Cytoskeleton. *Neuroscience*, 452:26–36.

Esumi, S., Kakazu, N., Taguchi, Y., Hirayama, T., Sasaki, A., Hirabayashi, T., Koide, T., Kitsukawa, T., Hamada, S., and Yagi, T. (2005). Monoallelic yet combinatorial expression of variable exons of the protocadherin-α gene cluster in single neurons. *Nature Genetics*, 37(2):171–176.

Fazzari, P., Snellinx, A., Sabanov, V., Ahmed, T., Serneels, L., Gartner, A., Shariati, S. A. M., Balschun, D., and De Strooper, B. (2014). Cell autonomous regulation of hippocampal circuitry via Aph1b-γ-secretase/neuregulin 1 signalling. *eLife*, 3:1–16.

Feng, B., Jiang, J., Kraus, P., Ng, J. H., Heng, J. C. D., Chan, Y. S., Yaw, L. P., Zhang, W., Loh, Y. H., Han, J., Vega, V. B., Cacheux-Rataboul, V., Lim, B., Lufkin, T., and Ng, H. H. (2009). Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. *Nature Cell Biology*, 11(2):197–203.

Ferber, E. C., Kajita, M., Wadlow, A., Tobiansky, L., Niessen, C., Ariga, H., Daniel, J., and Fujita, Y. (2008). A role for the cleaved cytoplasmic domain of E-cadherin in the nucleus. *Journal of Biological Chemistry*, 283(19):12691–12700.

Fleck, D., Voss, M., Brankatschk, B., Giudici, C., Hampel, H., Schwenk, B., Edbauer, D., Fukumori, A., Steiner, H., Kremmer, E., Haug-Kröper, M., Rossner, M. J., Fluhrer, R., Willem, M., and Haass, C. (2016). Proteolytic processing of neuregulin 1 type III by three intramembrane-cleaving proteases. *Journal of Biological Chemistry*, 291(1):318–333.

Fox, A. H., Fox, A. H., Lam, Y. W., Lam, Y. W., Leung, A. K. L., Leung, A. K. L., Lyon, C. E., Lyon, C. E., Andersen, J., Andersen, J., Mann, M., Mann, M., Lamond, A. I., and Lamond, A. I. (2002). Paraspeckles: a novel nuclear domain. *Current Biology*, 12(1):13–25.

Frade, J. M. (2005). Nuclear translocation of the p75 neurotrophin receptor cytoplasmic domain in response to neurotrophin binding. *Journal of Neuroscience*, 25(6):1407–1411.

Frånberg, J., Karlström, H., Winblad, B., Tjernberg, L. O., and Frykman, S. (2010). γ-Secretase dependent production of intracellular domains is reduced in adult compared to embryonic rat brain membranes. *PLoS ONE*, 5(3).

Franco, S. J., Martinez-Garay, I., Gil-Sanz, C., Harkins-Perry, S. R., and Müller, U. (2011). Reelin Regulates Cadherin Function via Dab1/Rap1 to Control Neuronal Migration and Lamination in the Neocortex. *Neuron*, 69(3):482–497.

Fujitani, M., Zhang, S., Fujiki, R., Fujihara, Y., and Yamashita, T. (2017). A chromosome 16p13.11 microduplication causes hyperactivity through dysregulation of miR-484/protocadherin-19 signaling. *Molecular Psychiatry*, 22(3):364–374.

Gaitan, Y. and Bouchard, M. (2006). Expression of the δ-protocadherin gene Pcdh19 in the developing mouse embryo. *Gene Expression Patterns*, 6(8):893–899.

Galindo-Riera, N., Newbold, S. A., Sledziowska, M., Griffiths, J., Mire, E., and Martinez-Garay,

I. (2020). Cellular and behavioral characterization of Pcdh19 mutant mice: subtle molecular changes, increased exploratory behavior and an impact of social environment. *bioRxiv*.

Garcia, I., Huangc, L., Ungd, K., and Arenkiel, B. R. (2012). Tracing Synaptic Connectivity onto Embryonic Stem Cell- Derived Neurons. *Stem Cells*, 30(10):2140–2151.

Gaztelumendi, N. and Nogués, C. (2014). Chromosome Instability in mouse Embryonic Stem Cells. *Scientific Reports*, 4:1–8.

Gerosa, L., Francolini, M., Bassani, S., and Passafaro, M. (2018). The Role of Protocadherin 19 (PCDH19) in Neurodevelopment and in the Pathophysiology of Early Infantile Epileptic Encephalopathy-9 (EIEE9). *Developmental Neurobiology*, 19:75–84.

Gil-Sanz, C., Franco, S. J., Martinez-Garay, I., Espinosa, A., Harkins-Perry, S., and Müller, U. (2013). Cajal-Retzius cells instruct neuronal migration by coincidence signaling between secreted and contact-dependent guidance cues. *Neuron*, 79(3):461–477.

Gil-Sanz, C., Landeira, B., Ramos, C., Costa, M. R., and Müller, U. (2014). Proliferative defects and formation of a double cortex in mice lacking Mltt4 and Cdh2 in the dorsal telencephalon. *Journal of Neuroscience*, 34(32):10475–10487.

Gonçalves, T. F., Gonçalves, A. P., Fintelman Rodrigues, N., dos Santos, J. M., Pimentel, M. M. G., and Santos-Rebouças, C. B. (2014). KDM5C mutational screening among males with intellectual disability suggestive of X-Linked inheritance and review of the literature. *European Journal of Medical Genetics*, 57(4):138–144.

Goodman, K. M., Rubinstein, R., Dan, H., Bahna, F., Mannepalli, S., Ahlsén, G., Thu, C. A., Sampogna, R. V., Maniatis, T., Honig, B., and Shapiro, L. (2017). Protocadherin cis-dimer architecture and recognition unit diversity. *Proceedings of the National Academy of Sciences of the United States of America*, 114(46):E9829–E9837.

Gordon, A., Yoon, S.-j., Tran, S. S., Makinson, C. D., Park, J. Y., Andersen, J., Valencia, A. M., Horvath, S., Xiao, X., Huguenard, J. R., Pașca, S. P., and Geschwind, D. H. (2021). Long-

term maturation of human cortical organoids matches key early postnatal transitions. *Nature Neuroscience*, 24(3):331–342.

Gordon, W. R., Vardar-Ulu, D., Histen, G., Sanchez-Irizarry, C., Aster, J. C., and Blacklow, S. C. (2007). Structural basis for autoinhibition of Notch. *Nature Structural and Molecular Biology*, 14(4):295–300.

Gotz, M., Stoykova, A., and Gruss, P. (1998). Pax6 Controls Radial Glia Differentiation in the Cerebral Cortex. *Neuron*, 21:1031–1044.

Gowrishankar, K., Zeidler, M. G., and Vincenz, C. (2004). Release of a membrane-bound death domain by γ-secretase processing of the p75NTR homolog NRADD. *Journal of Cell Science*, 117(18):4099–4111.

Gray, E. G. (1959). Electron microscopy of synaptic contacts on dendrite spines of the cerebral cortex. *Nature*, 183(4675):1592–1593.

Haas, I. G., Frank, M., Véron, N., and Kemler, R. (2005). Presenilin-dependent processing and nuclear function of γ-protocadherins. *Journal of Biological Chemistry*, 280(10):9313–9319.

Hallaq, R., Volpicelli, F., Cuchillo-Ibanez, I., Hooper, C., Mizuno, K., Uwanogho, D., Causevic, M., Asuni, A., To, A., Soriano, S., Giese, K. P., Lovestone, S., and Killick, R. (2015). The Notch intracellular domain represses CRE-dependent transcription. *Cellular Signalling*, 27(3):621–629.

Hambsch, B., Grinevich, V., Seeburg, P. H., and Schwarz, M. K. (2005). γ-Protocadherins, presenilin-mediated release of C-terminal fragment promotes locus expression. *Journal of Biological Chemistry*, 280(16):15888–15897.

Hansen, D. V., Lui, J. H., Parker, P. R., and Kriegstein, A. R. (2010). Neurogenic radial glia in the outer subventricular zone of human neocortex. *Nature*, 464(7288):554–561.

Hanson, J. E. and Madison, D. V. (2007). Presynaptic Fmr1 genotype influences the degree of

synaptic connectivity in a mosaic mouse model of fragile X syndrome. *Journal of Neuroscience*, 27(15):4014–4018.

Harris, K. D. and Shepherd, G. M. (2015). The neocortical circuit: Themes and variations. *Nature Neuroscience*, 18(2):170–181.

Harrison, O. J., Brasch, J., Katsamba, P. S., Ahlsen, G., Noble, A. J., Dan, H., Sampogna, R. V., Potter, C. S., Carragher, B., Honig, B., and Shapiro, L. (2020). Family-wide Structural and Biophysical Analysis of Binding Interactions among Non-clustered δ-Protocadherins. *Cell Reports*, 30(8):2655–2671.e7.

Harrison-Uy, S. J. and Pleasure, S. J. (2012). Wnt signaling and forebrain development. *Cold Spring Harbor Perspectives in Biology*, 4(7):1–11.

Hartmann, D., De Strooper, B., Serneels, L., Craessaerts, K., Herreman, A., Annaert, W., Umans, L., Lübke, T., Illert, A. L., Von Figura, K., and Saftig, P. (2002). The disintegrin/metalloprotease ADAM 10 is essential for Notch signalling but not for α-secretase activity in fibroblasts. *Human Molecular Genetics*, 11(21):2615–2624.

Hatanaka, Y. and Murakami, F. (2002). In vitro analysis of the origin, migratory behavior, and maturation of cortical pyramidal cells. *Journal of Comparative Neurology*, 454(1):1–14.

Hayashi, S., Inoue, Y., Hattori, S., Kaneko, M., Shioi, G., Miyakawa, T., and Takeichi, M. (2017). Loss of X-linked Protocadherin-19 differentially affects the behavior of heterozygous female and hemizygous male mice. *Scientific Reports*, 7(1):1–15.

Hayashi, S., Inoue, Y., Kiyonari, H., Abe, T., Misaki, K., Moriguchi, H., Tanaka, Y., and Takeichi, M. (2014). Protocadherin-17 Mediates Collective Axon Extension by Recruiting Actin Regulator Complexes to Interaxonal Contacts. *Developmental Cell*, 30(6):673–687.

Hayashi, S. and Takeichi, M. (2015). Emerging roles of protocadherins: From self-avoidance to enhancement of motility. *Journal of Cell Science*, 128(8):1455–1464.

Heggem, M. A. and Bradley, R. S. (2003). The cytoplasmic domain of xenopus NF-protocadherin interacts with TAF1/SET. *Developmental Cell*, 4(3):419–429.

Hemming, M. L., Elias, J. E., Gygi, S. P., and Selkoe, D. J. (2009). Identification of β-secretase (BACE1) substrates using quantitative proteomics. *PLoS ONE*, 4(12).

Hertel, N., Krishna-K, Nuernberger, M., and Redies, C. (2008). A cadherin-based code for the divisions of the mouse basal ganglia. *Journal of Comparative Neurology*, 508(4):511–528.

Hirabayashi, Y., Itoh, Y., Tabata, H., Nakajima, K., Akiyama, T., Masuyama, N., and Gotoh, Y. (2004). The Wnt/β-catenin pathway directs neuronal differentation of cortical neural precursor cells. *Development*, 131(12):2791–2801.

Hirano, S., Kimoto, N., Shimoyama, Y., Hirohashi, S., and Takeichi, M. (1992). Identification of a neural α-catenin as a key regulator of cadherin function and multicellular organization. *Cell*, 70(2):293–301.

Hirano, S. and Takeichi, M. (2012). Cadherins in brain morphogenesis and wiring. *Physiological Reviews*, 92(2):597–634.

Hitt, B. D., Jaramillo, T. C., Chetkovich, D. M., and Vassar, R. (2010). BACE1-/- mice exhibit seizure activity that does not correlate with sodium channel level or axonal localization. *Molecular Neurodegeneration*, 5(1):1–14.

Homan, C. C., Pederson, S., To, T. H., Tan, C., Piltz, S., Corbett, M. A., Wolvetang, E., Thomas, P. Q., Jolly, L. A., and Gecz, J. (2018). PCDH19 regulation of neural progenitor cell differentiation suggests asynchrony of neurogenesis as a mechanism contributing to PCDH19 Girls Clustering Epilepsy. *Neurobiology of Disease*, 116(May):106–119.

Hong, H., Yang, L., and Stallcup, M. R. (1999). Hormone-independent transcriptional activation and coactivator binding by novel orphan nuclear receptor ERR3. *Journal of Biological Chemistry*, 274(32):22618–22626.

Hongisto, H., Vuoristo, S., Mikhailova, A., Suuronen, R., Virtanen, I., Otonkoski, T., and Skottman, H. (2012). Laminin-511 expression is associated with the functionality of feeder cells in human embryonic stem cell culture. *Stem Cell Research*, 8(1):97–108.

Hoshina, N., Tanimura, A., Yamasaki, M., Inoue, T., Fukabori, R., Kuroda, T., Yokoyama, K., Tezuka, T., Sagara, H., Hirano, S., Kiyonari, H., Takada, M., Kobayashi, K., Watanabe, M., Kano, M., Nakazawa, T., and Yamamoto, T. (2013). Protocadherin 17 regulates presynaptic assembly in topographic corticobasal ganglia circuits. *Neuron*, 78(5):839–854.

Hsia, H. E., Tüshaus, J., Brummer, T., Zheng, Y., and Scilabra, S. D. (2019). Functions of ' A disintegrin and metalloproteases (ADAMs)' in the mammalian nervous system. *Cellular and Molecular Life Sciences*, 76(16):3055–3081.

Huang, Z. J. and Paul, A. (2019). The diversity of GABAergic neurons and neural communication elements. *Nature Reviews Neuroscience*, 20(September):563–572.

Huenniger, K., Krämer, A., Soom, M., Chang, I., Köhler, M., Depping, R., Kehlenbach, R. H., and Kaether, C. (2010). Notch1 signaling is mediated by importins alpha 3, 4, and 7. *Cellular and Molecular Life Sciences*, 67(18):3187–3196.

Hulpiau, P. and van Roy, F. (2009). Molecular evolution of the cadherin superfamily. *International Journal of Biochemistry and Cell Biology*, 41(2):349–369.

Humpherys, D., Eggan, K., Akutsu, H., Hochedlinger, K., Rideout, W. M., Biniszkiewicz, D., Yanagimachi, R., and Jaenisch, R. (2001). Epigenetic instability in ES cells and cloned mice. *Science*, 293(5527):95–97.

Inoue, E., Deguchi-Tawarada, M., Togawa, A., Matsui, C., Arita, K., Katahira-Tayama, S., Sato, T., Yamauchi, E., Oda, Y., and Takai, Y. (2009). Synaptic activity prompts γ-secretase-mediated cleavage of EphA4 and dendritic spine formation. *Journal of Cell Biology*, 185(3):551–564.

Itoh, Y. (2015). Membrane-type matrix metalloproteinases: Their functions and regulations. *Matrix Biology*, 44-46:207–223.

Jan, Y. N. and Jan, L. Y. (2010). Branching out: Mechanisms of dendritic arborization. *Nature Reviews Neuroscience*, 11(5):316–328.

Janes, P. W., Saha, N., Barton, W. A., Kolev, M. V., Wimmer-Kleikamp, S. H., Nievergall, E., Blobel, C. P., Himanen, J. P., Lackmann, M., and Nikolov, D. B. (2005). Adam meets Eph: An ADAM substrate recognition module acts as a molecular switch for ephrin cleavage in trans. *Cell*, 123(2):291–304.

Jang, Y. N., Jung, Y. S., Soo, H. L., Moon, C. H., Kim, C. H., and Eun, J. B. (2009). Calpain-mediated N-cadherin proteolytic processing in brain injury. *Journal of Neuroscience*, 29(18):5974–5984.

Jarriault S, Brou C, Logeat F, Schroeter Eh, Kopan R, and Israel A (1995). Signalling downstream of activated mammalian Notch [see comments]. *Nature*, 377(6547):355–358.

Jeffrey, R. A., Ch'ng, T. H., O'Dell, T. J., and Martin, K. C. (2009). Activity-dependent anchoring of importin α at the synapse involves regulated binding to the cytoplasmic tail of the NR1-1a subunit of the NMDA receptor. *Journal of Neuroscience*, 29(50):15613–15620.

Jin, Y. and Li, H. (2019). Revisiting Dscam diversity: lessons from clustered protocadherins. *Cellular and Molecular Life Sciences*, 76(4):667–680.

Jorissen, E., Prox, J., Bernreuther, C., Weber, S., Schwanbeck, R., Serneels, L., Snellinx, A., Craessaerts, K., Thathiah, A., Tesseur, I., Bartsch, U., Weskamp, G., Blobel, C. P., Glatzel, M., De Strooper, B., and Saftig, P. (2010). The disintegrin/metalloproteinase ADAM10 is essential for the establishment of the brain cortex. *Journal of Neuroscience*, 30(14):4833–4844.

Juberg, R. C. and Hellman, C. D. (1971). A new familial form of convulsive disorder and mental retardation limited to females. *The Journal of Pediatrics*, 79(5):726–732.

Jun, L., Frints, S., Duhamel, H., Herold, A., Abad-Rodrigues, J., Dotti, C., Izaurralde, E., Marynen, P., and Froyen, G. (2001). NXF5, a novel member of the nuclear RNA export factor

family, is lost in a male patient with a syndromic form of mental retardation. *Current Biology*, 11(18):1381–1391.

Kadowaki, M., Nakamura, S., Machon, O., Krauss, S., Radice, G. L., and Takeichi, M. (2007). N-cadherin mediates cortical organization in the mouse brain. *Developmental Biology*, 304(1):22–33.

Karlström, H., Bergman, A., Lendahl, U., Näslund, J., and Lundkvist, J. (2002). A sensitive and quantitative assay for measuring cleavage of presenilin substrates. *Journal of Biological Chemistry*, 277(9):6763–6766.

Kast, R. J., Lanjewar, A. L., Smith, C. D., and Levitt, P. (2019). FOXP2 exhibits projection neuron class specific expression, but is not required for multiple aspects of cortical histogenesis. *eLife*, 8:1–20.

Kawaguchi, D., Yoshimatsu, T., Hozumi, K., and Gotoh, Y. (2008). Selection of differentiating cells by different levels of delta-like 1 among neural precursor cells in the developing mouse telencephalon. *Development*, 135(23):3849–3858.

Kelleher, R. J. and Shen, J. (2017). Presenilin-1 mutations and Alzheimer's disease. *Proceedings of the National Academy of Sciences of the United States of America*, 114(4):629–631.

Kim, S. Y., Chung, H. S., Sun, W., and Kim, H. (2007). Spatiotemporal expression pattern of non-clustered protocadherin family members in the developing rat brain. *Neuroscience*, 147(4):996–1021.

Kim, S. Y., Yasuda, S., Tanaka, H., Yamagata, K., and Kim, H. (2011). Non-clustered protocadherin. *Cell Adhesion and Migration*, 5(2):97–105.

Kolc, K. L., Møller, R. S., Sadleir, L. G., Scheffer, I. E., Kumar, R., and Gecz, J. (2020). PCDH19 Pathogenic Variants in Males : Expanding the Phenotypic Spectrum. *Adv Exp Med Bio*.

Kolc, K. L., Sadleir, L. G., Scheffer, I. E., Ivancevic, A., Roberts, R., and Gecz, J. (2019). A systematic review and meta-analysis of 271 PCDH19-variant individuals identifies psychiatric

comorbidities , and association of seizure onset and disease severity. *Molecular Psychiatry*, pages 241–251.

Komada, M., Saitsu, H., Kinboshi, M., Miura, T., Shiota, K., and Ishibashi, M. (2008). Hedgehog signaling is involved in development of the neocortex. *Development*, 135(16):2717–2727.

Konopka, A., Grajkowska, W., Ziemiańska, K., Roszkowski, M., Daszkiewicz, P., Rysz, A., Marchel, A., Koperski, Ł., Wilczyński, G. M., and Dzwonek, J. (2013). Matrix metalloproteinase-9 (MMP-9) in human intractable epilepsy caused by focal cortical dysplasia. *Epilepsy Research*, 104(1-2):45–58.

Kosugi, S., Hasebe, M., Tomita, M., and Yanagawa, H. (2009). Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. *Proceedings of the National Academy of Sciences of the United States of America*, 106(25):10171–10176.

Lammich, S., Kojro, E., Postina, R., Gilbert, S., Pfeiffer, R., Jasionowski, M., Haass, C., and Fahrenholz, F. (1999). Constitutive and regulated α-secretase cleavage of Alzheimer's amyloid precursor protein by a disintegrin metalloprotease. *Proceedings of the National Academy of Sciences of the United States of America*, 96(7):3922–3927.

Lang, C., Barco, A., Zablow, L., Kandel, E. R., Siegelbaum, S. A., and Zakharenko, S. S. (2004). Transient expansion of synaptically connected dendritic spines upon induction of hippocampal long-term potentiation. *Proceedings of the National Academy of Sciences of the United States of America*, 101(47):16665–16670.

Lee, J., Matsuzawa, A., Shiura, H., Sutani, A., and Ishino, F. (2018). Preferable in vitro condition for maintaining faithful DNA methylation imprinting in mouse embryonic stem cells. *Genes to Cells*, 23(3):146–160.

Lefebvre, J. L., Kostadinov, D., Chen, W. V., Maniatis, T., and Sanes, J. R. (2012). Protocadherins mediate dendritic self-avoidance in the mammalian nervous system. *Nature*, 488(7412):517–521.

Lenge, M., Marini, C., Canale, E., Napolitano, A., De Masi, S., Trivisano, M., Mei, D., Longo, D., Rossi Espagnet, M. C., Lucenteforte, E., Barba, C., Specchio, N., and Guerrini, R. (2020). Quantitative MRI-Based Analysis Identifies Developmental Limbic Abnormalities in PCDH19 Encephalopathy. *Cerebral Cortex*, pages 1–12.

Leung, L. C., Urbančič, V., Baudet, M. L., Dwivedy, A., Bayley, T. G., Lee, A. C., Harris, W. A., and Holt, C. E. (2013). Coupling of NF-protocadherin signaling to axon guidance by cue-induced translation. *Nature Neuroscience*, 16(2):166–173.

Leyva-Díaz, E. and López-Bendito, G. (2013). In and out from the cortex: Development of major forebrain connections. *Neuroscience*, 254:26–44.

Liao, Y., Smyth, G. K., and Shi, W. (2014). FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, 30(7):923–930.

Lichtenthaler, S. F., Lemberg, M. K., and Fluhrer, R. (2018). Proteolytic ectodomain shedding of membrane proteins in mammals—hardware, concepts, and recent developments. *The EMBO Journal*, 37(15):1–24.

Light, S. E. and Jontes, J. D. (2017). δ-Protocadherins: Organizers of neural circuit assembly. *Seminars in Cell and Developmental Biology*, 69:83–90.

Lim, L., Mi, D., Llorca, A., and Marín, O. (2018). Development and Functional Diversification of Cortical Interneurons. *Neuron*, 100(2):294–313.

Lindeboom, R. G., Supek, F., and Lehner, B. (2016). The rules and impact of nonsense-mediated mRNA decay in human cancers. *Nature Genetics*, 48(10):1112–1118.

Llinares-Benadero, C. and Borrell, V. (2019). Deconstructing cortical folding: genetic, cellular and mechanical determinants. *Nature Reviews Neuroscience*, 20(3):161–176.

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12):1–21.

Lu, P., Bai, X. C., Ma, D., Xie, T., Yan, C., Sun, L., Yang, G., Zhao, Y., Zhou, R., Scheres, S. H., and Shi, Y. (2014). Three-dimensional structure of human γ-secretase. *Nature*, 512(7513):166–170.

Lv, X., Ren, S.-q., Zhang, X.-j., Shen, Z., Ghosh, T., Xianyu, A., Gao, P., Li, Z., Lin, S., Yu, Y., Zhang, Q., Groszer, M., and Shi, S.-h. (2019). TBR2 coordinates neurogenesis expansion and precise microcircuit organization via Protocadherin 19 in the mammalian cortex. *Nature Communications*, 10(3946):1–15.

Magg, T., Schreiner, D., Solis, G. P., Bade, E. G., and Hofer, H. W. (2005). Processing of the human protocadherin Fat1 and translocation of its cytoplasmic domain to the nucleus. *Experimental Cell Research*, 307(1):100–108.

Mah, K. M., Houston, D. W., and Weiner, J. A. (2016). The γ 3-Protocadherin-C3 isoform inhibits canonical Wnt signalling by binding to and stabilizing Axin1 at the membrane. *Scientific Reports*, 6:1–17.

Malatesta, P., Hartfuss, E., and Götz, M. (2000). Isolation of radial glial cells by fluorescent-activated cell sorting reveals a neural lineage. *Development*, 127(24):5253–5263.

Malinverno, M., Carta, M., Epis, R., Marcello, E., Verpelli, C., Cattabeni, F., Sala, C., Mulle, C., Di Luca, M., and Gardoni, F. (2010). Synaptic localization and activity of ADAM10 regulate excitatory synapses through N-cadherin cleavage. *Journal of Neuroscience*, 30(48):16343–16355.

Marambaud, P., Shioi, J., Serban, G., Georgakopoulos, A., Sarner, S., Nagy, V., Baki, L., Wen, P., Efthimiopoulos, S., Shao, Z., Wisniewski, T., and Robakis, N. K. (2002). A presenilin-1/γ-secretase cleavage releases the E-cadherin intracellular domain and regulates disassembly of adherens junctions. *EMBO Journal*, 21(8):1948–1956.

Marambaud, P., Wen, P. H., Dutt, A., Shioi, J., Takashima, A., Siman, R., and Robakis, N. K. (2003). A CBP binding transcriptional repressor produced by the PS1/ ϵ-cleavage of N-Cadherin is inhibited by PS1 FAD mutations. *Cell*, 114(5):635–645.

Marcello, E., Saraceno, C., Musardo, S., Vara, H., De La Fuente, A. G., Pelucchi, S., Di Marino, D., Borroni, B., Tramontano, A., Pérez-Otaño, I., Padovani, A., Giustetto, M., Gardoni, F., and Di Luca, M. (2013). Endocytosis of synaptic ADAM10 in neuronal plasticity and Alzheimer's disease. *Journal of Clinical Investigation*, 123(6):2523–2538.

Marín, O. and Rubenstein, J. L. (2001). A long, remarkable journey: Tangential migration in the telencephalon. *Nature Reviews Neuroscience*, 2(11):780–790.

Maroof, A. M., Brown, K., Shi, S.-h., Studer, L., and Anderson, S. A. (2010). Prospective Isolation of Cortical Interneuron Precursors from Mouse Embryonic Stem Cells. *The Journal of Neuroscience*, 30(13):4667–4675.

Maroof, A. M., Keros, S., Tyson, J. A., Ying, S.-w., Ganat, Y. M., Merkle, F. T., Liu, B., Goulburn, A., Stanley, E. G., Elefanty, A. G., Widmer, H. R., Eggan, K., Goldstein, P. A., Anderson, S. A., and Studer, L. (2013). Directed Differentiation and Functional Maturation of Cortical Interneurons from Human Embryonic Stem Cells. *Stem Cell*, 12(5):559–572.

Martinez-Garay, I., Gil-Sanz, C., Franco, S. J., Espinosa, A., Molnár, Z., and Mueller, U. (2016). Cadherin 2/4 signaling via PTP1B and catenins is crucial for nucleokinesis during radial neuronal migration in the neocortex. *Development (Cambridge)*, 143(12):2121–2134.

Martinez-Garay, I., Jablonka, S., Sutajova, M., Steuernagel, P., Gal, A., and Kutsche, K. (2002). A new gene family (FAM9) of low-copy repeats in Xp22.3 expressed exclusively in testis: Implications for recombinations in this region. *Genomics*, 80(3):259–267.

Matus, A., Bernhardt, R., and Hugh-Jonest, T. (1981). High molecular weight microtubule-associated proteins are preferentially associated with dendritic microtubules in brain (tubulin/immunohistochemistry/immunoelectrophoresis/postsynaptic density). *Cell Biology*, 78(5):3010–3014.

May, P., Bock, H. H., Nimpf, J., and Herz, J. (2003). Differential glycosylation regulates processing of lipoprotein receptors by γ-secretase. *Journal of Biological Chemistry*, 278(39):37386–37392.

McCusker, C., Cousin, H., Neuner, R., and Alfandari, D. (2009). Extracellular Cleavage of Cadherin-11 by ADAM Metalloproteases Is Essential for Xenopus Cranial Neural Crest Cell Migration CatherineSophie Mokas, John R. Mills Cristina Garreau,. *Molecular Biology of the Cell*, 20:78–89.

McCusker, C. D. and Alfandari, D. (2009). Life after proteolysis. Exploring the signaling capabilities of classical cadherin cleavage fragments. *Communicative and Integrative Biology*, 2(2):155–157.

Mellacheruvu, D., Wright, Z., Couzens, A. L., Lambert, J. P., St-Denis, N. A., Li, T., Miteva, Y. V., Hauri, S., Sardiu, M. E., Low, T. Y., Halim, V. A., Bagshaw, R. D., Hubner, N. C., Al-Hakim, A., Bouchard, A., Faubert, D., Fermin, D., Dunham, W. H., Goudreault, M., Lin, Z. Y., Badillo, B. G., Pawson, T., Durocher, D., Coulombe, B., Aebersold, R., Superti-Furga, G., Colinge, J., Heck, A. J., Choi, H., Gstaiger, M., Mohammed, S., Cristea, I. M., Bennett, K. L., Washburn, M. P., Raught, B., Ewing, R. M., Gingras, A. C., and Nesvizhskii, A. I. (2013). The CRAPome: A contaminant repository for affinity purification-mass spectrometry data. *Nature Methods*, 10(8):730–736.

Menassa, D. A. and Gomez-Nicola, D. (2018). Microglial dynamics during human brain development. *Frontiers in Immunology*, 9(MAY).

Mincheva-Tasheva, S., Nieto Guil, A. F., Homan, C. C., Gecz, J., and Thomas, P. Q. (2021). Disrupted Excitatory Synaptic Contacts and Altered Neuronal Network Activity Underpins the Neurological Phenotype in PCDH19-Clustering Epilepsy (PCDH19-CE). *Molecular Neurobiology*.

Miura, Y., Li, M. Y., Birey, F., Ikeda, K., Revah, O., Thete, M. V., Park, J. Y., Puno, A., Lee, S. H., Porteus, M. H., and Pașca, S. P. (2020). Generation of human striatal organoids and cortico-striatal assembloids from human pluripotent stem cells. *Nature Biotechnology*, 38(12):1421–1430.

Miyata, T., Kawaguchi, A., Okano, H., and Ogawa, M. (2001). Asymmetric inheritance of radial glial fibers by cortical neurons. *Neuron*, 31(5):727–741.

Miyata, T., Kawaguchi, A., Saito, K., Kawano, M., Muto, T., and Ogawa, M. (2004). Asymmetric production of surface-dividing and non-surface-dividing cortical progenitor cells. *Development*, 131(13):3133–3145.

Molyneaux, B. J., Arlotta, P., Menezes, J. R., and Macklis, J. D. (2007). Neuronal subtype specification in the cerebral cortex. *Nature Reviews Neuroscience*, 8(6):427–437.

Morrow, E. M., Yoo, S.-Y., Flavell, S. W., Kim, T.-K., Lin, Y., Hill, R. S., Mukaddes, N. M., Balkhy, S., Gascon, G., Hashmi, A., Al-Saad, S., Ware, J., Josheph, R. M., Greenblatt, R., Gleason, D., Ertelt, J. A., Apse, K. A., Bodell, A., Partlow, J. N., Barry, B., Yao, H., Markianos, K., Ferland, R. J., Greenberg, M. E., and Walsh, C. A. (2008). Identifying Autism Loci and Genes by Tracing Recent Shared Ancestry. *Science*, 321.

Munji, R. N., Choe, Y., Li, G., Siegenthaler, J. A., and Pleasure, S. J. (2011). Wnt signaling regulates neuronal differentiation of cortical intermediate progenitors. *Journal of Neuroscience*, 31(5):1676–1687.

Musardo, S., Marcello, E., Gardoni, F., and Di Luca, M. (2014). ADAM10 in synaptic physiology and pathology. *Neurodegenerative Diseases*, 13(2-3):72–74.

Nadarajah, B., Brunstrom, J. E., Grutzendler, J., Wong, R. O., and Pearlman, A. L. (2001). Two modes of radial migration in early development of the cerebral cortex. *Nature Neuroscience*, 4(2):143–150.

Nagappan-Chettiar, S., Johnson-Venkatesh, E. M., and Umemori, H. (2017). Activity-dependent proteolytic cleavage of cell adhesion molecules regulates excitatory synaptic development and function. *Neuroscience Research*, 116:60–69.

Nakao, S., Platek, A., Hirano, S., and Takeichi, M. (2008). Contact-dependent promotion of cell migration by the OL-protocadherin-Nap1 interaction. *Journal of Cell Biology*, 182(2):395–410.

Nguyen Ba, A. N., Pogoutse, A., Provart, N., and Moses, A. M. (2009). NLStradamus: A simple Hidden Markov Model for nuclear localization signal prediction. *BMC Bioinformatics*, 10:1–11.

Nieman, M. T., Kim, J. B., Johnson, K. R., and Wheelock, M. J. (1999). Mechanism of extracellular domain-deleted dominant negative cadherins. *Journal of Cell Science*, 112(10):1621–1632.

Nieto, M., Monuki, E. S., Tang, H., Imitola, J., Haubst, N., Khoury, S. J., Cunningham, J., Gotz, M., and Walsh, C. A. (2004). Expression of Cux-1 and Cux-2 in the subventricular zone and upper layers II-IV of the cerebral cortex. *Journal of Comparative Neurology*, 479(2):168–180.

Niimura, M., Isoo, N., Takasugi, N., Tsuruoka, M., Ui-Tei, K., Saigo, K., Morohashi, Y., Tomita, T., and Iwatsubo, T. (2005). Aph-1 contributes to the stabilization and trafficking of the γ-secretase complex through mechanisms involving intermolecular and intramolecular interactions. *Journal of Biological Chemistry*, 280(13):12967–12975.

Noctor, S. C., Flint, A. C., Weissman, T. A., Dammerman, R. S., and Kriegstein, A. R. (2001). Neurons derived from radial glial cells establish radial units in neocortex. *Nature*, 409:714–720.

Noctor, S. C., Martinez-Cerdeño, V., Ivic, L., and Kriegstein, A. R. (2004). Cortical neurons arise in symmetric and asymmetric division zones and migrate through specific phases. *Nature Neuroscience*, 7(2):136–144.

Nose, A., Tsuji, K., and Takeichi, M. (1990). Localization of specificity determining sites in cadherin cell adhesion molecules. *Cell*, 61(1):147–155.

Ogawa, M., Miyata, T., Nakajimat, K., Yagyu, K., Seike, M., Ikenaka, K., Yamamoto, H., and Mikoshibat, K. (1995). The reeler gene-associated antigen on cajal-retzius neurons is a crucial molecule for laminar organization of cortical neurons. *Neuron*, 14(5):899–912.

Ogawa, Y., Miyamoto, Y., Oka, M., and Yoneda, Y. (2012). The Interaction Between Importin-α and Nup153 Promotes Importin-α/β-Mediated Nuclear Import. *Traffic*, 13(7):934–946.

Pan, D. and Gerald M, R. (1997). Kuzbanian controls proteolytic processing of Notch and mediates lateral inhibition during Drosophila and vertebrate neurogenesis. *Cell*, 90(2):271–280.

Pan, F. and Gan, W. B. (2008). Two-photon imaging of dendritic spine development in the mouse cortex. *Developmental Neurobiology*, 68(6):771–778.

Panning, B. (2008). X-chromosome inactivation: The molecular basis of silencing. *Journal of Biology*, 7(8).

Paradies, N. E. and Grunwald, G. B. (1993). Purification and characterization of NCAD90, a Soluble endogenous form of N-cadherin, which is generated by proteolysis during retinal development and retains adhesive and neurite-promoting function. *Journal of Neuroscience Research*, 36(1):33–45.

Pederick, D. T., Homan, C. C., Jaehne, E. J., Piltz, S. G., Haines, B. P., Baune, B. T., Jolly, L. A., Hughes, J. N., Gecz, J., and Thomas, P. Q. (2016). Pcdh19 Loss-of-Function Increases Neuronal Migration In Vitro but is Dispensable for Brain Development in Mice. *Scientific Reports*, 6(February):1–10.

Pederick, D. T., Richards, K. L., Piltz, S. G., Kumar, R., Mincheva-Tasheva, S., Mandelstam, S. A., Dale, R. C., Scheffer, I. E., Gecz, J., Petrou, S., Hughes, J. N., and Thomas, P. Q. (2018). Abnormal Cell Sorting Underlies the Unique X-Linked Inheritance of PCDH19 Epilepsy. *Neuron*, 97(1):59–66.e5.

Peek, S. L., Mah, K. M., and Weiner, J. A. (2017). Regulation of neural circuit formation by protocadherins. *Cellular and Molecular Life Sciences*, 74(22):4133–4157.

Perez-Pinera, P., Ousterout, D. G., Brown, M. T., and Gersbach, C. A. (2012). Gene targeting to the ROSA26 locus directed by engineered zinc finger nucleases. *Nucleic Acids Research*, 40(8):3741–3752.

Pham, D. H., Tan, C. C., Homan, C. C., Kolc, K. L., Corbett, M. A., McAninch, D., Fox, A. H., Thomas, P. Q., Kumar, R., and Gecz, J. (2017). Protocadherin 19 (PCDH19) interacts

with paraspeckle protein NONO to co-regulate gene expression with estrogen receptor alpha (ERα). *Human Molecular Genetics*, 26(11):2042–2052.

Piper, M., Dwivedy, A., Leung, L., Bradley, R. S., and Holt, C. E. (2008). NF-protocadherin and TAF1 regulate retinal axon initiation and elongation in vivo. *Journal of Neuroscience*, 28(1):100–105.

Polleux, F., Giger, R. J., Ginty, D. D., Kolodkin, A. L., and Ghosh, A. (1998). Patterning of cortical efferent projections by semaphorin-neuropilin interactions. *Science*, 282(5395):1904–1906.

Polleux, F., Morrow, T., and Ghosh, A. (2000). Semaphorin 3A is a chemoattractant for cortical apical dendrites. *Nature*, 404(6778):567–573.

Popp, M. and Maquat, L. (2016). Leveraging Rules of Nonsense-Mediated mRNA Decay for Genome Engineering and Personalized Medicine Maximilian. *Cell*, 165(6):100–106.

Porlan, E., Martí-Prado, B., Morante-Redolat, J. M., Consiglio, A., Delgado, A. C., Kypta, R., López-Otín, C., Kirstein, M., and Fariñas, I. (2014). MT5-MMP regulates adult neural stem cell functional quiescence through the cleavage of N-cadherin. *Nature Cell Biology*, 16(7):629–638.

Prox, J., Bernreuther, C., Altmeppen, H., Grendel, J., Glatzel, M., Hooge, R. D., Stroobants, S., Ahmed, T., Balschun, D., Willem, M., Lammich, S., Isbrandt, D., Schweizer, M., Horre, K., Strooper, B. D., and Saftig, P. (2013). Postnatal Disruption of the Disintegrin / Metalloproteinase ADAM10 in Brain Causes Epileptic Seizures , Learning Deficits , Altered Spine Morphology , and Defective Synaptic Functions. *The Journal of Neuroscience*, 33(32):12915–12928.

Qi, H., Rand, M. D., Wu, X., Sestan, N., Wang, W., Rakic, P., Xu, T., and Artavanis-Tsakonas, S. (1999). Processing of the Notch ligand Delta by the metalloprotease Kuzbanian. *Science*, 283(5398):91–94.

Raefski, A. S. and O'Neill, M. J. (2005). Identification of a cluster of X-linked imprinted genes in mice. *Nature Genetics*, 37(6):620–624.

Rakic, P. (1972). Mode of cell migration to the superficial layers of fetal monkey neocortex. *Journal of Comparative Neurology*, 145(1):61–83.

Rakic, P. (1974). Neurons in Rhesus Monkey Visual Cortex: Systematic Relation between Time of Origin and Eventual Disposition. *Science*, 183(4123):425–427.

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*, 8(11):2281–2308.

Reddy, D. S. (2014). Neurosteroids and their role in sex-specific epilepsies. *Neurobiology of Disease*, 72(PB):198–209.

Redies, C., Vanhalst, K., and Van Roy, F. (2005). δ-Protocadherins: Unique structures and functions. *Cellular and Molecular Life Sciences*, 62(23):2840–2852.

Reiss, K., Maretzky, T., Haas, I. G., Schulte, M., Ludwig, A., Frank, M., and Saftig, P. (2006). Regulated ADAM10-dependent ectodomain shedding of γ-protocadherin C3 modulates cell-cell adhesion. *Journal of Biological Chemistry*.

Reiss, K., Maretzky, T., Ludwig, A., Tousseyn, T., De Strooper, B., Hartmann, D., and Saftig, P. (2005). ADAM10 cleavage of N-cadherin and regulation of cell-cell adhesion and β-catenin nuclear signalling. *EMBO Journal*, 24(4):742–752.

Restituito, S., Khatri, L., Ninan, I., Mathews, P. M., Liu, X., Weinberg, R. J., and Ziff, E. B. (2011). Synaptic autoregulation by metalloproteases and γ-secretase. *Journal of Neuroscience*, 31(34):12083–12093.

Reyes, A., Anders, S., Weatheritt, R. J., Gibson, T. J., Steinmetz, L. M., and Huber, W. (2013). Drift and conservation of differential exon usage across tissues in primate species. *Proceedings of the National Academy of Sciences of the United States of America*, 110(38):15377–15382.

Romasko, E. J., DeChene, E. T., Balciuniene, J., Akgumus, G. T., Helbig, I., Tarpinian, J. M., Keena, B. A., Vogiatzi, M. G., Zackai, E. H., Izumi, K., Massey, S. L., and Tayoun, A. N. (2018). PCDH19-related epilepsy in a male with Klinefelter syndrome: Additional evidence supporting PCDH19 cellular interference disease mechanism. *Epilepsy Research*, 145(January):89–92.

Roux, K. J., Kim, D. I., Burke, B., and May, D. G. (2018). BioID: A Screen for Protein-Protein Interactions. *Current Protocols in Protein Science*, 91(1):19.23.1–19.23.15.

Rubinstein, R., Thu, C. A., Goodman, K. M., Wolcott, H. N., Bahna, F., Mannepalli, S., Ahlsen, G., Chevee, M., Halim, A., Clausen, H., Maniatis, T., Shapiro, L., and Honig, B. (2015). Molecular Logic of Neuronal Self-Recognition through Protocadherin Domain Interactions. *Cell*, 163(3):629–642.

Sachan, N., Mishra, A. K., Mutsuddi, M., and Mukherjee, A. (2013). The Drosophila Importin-α3 Is Required for Nuclear Import of Notch In Vivo and It Displays Synergistic Effects with Notch Receptor on Cell Proliferation. *PLoS ONE*, 8(7).

Sachse, S. M., Lievens, S., Ribeiro, L. F., Dascenco, D., Masschaele, D., Horré, K., Misbaer, A., Vanderroost, N., De Smet, A. S., Salta, E., Erfurth, M., Kise, Y., Nebel, S., Van Delm, W., Plaisance, S., Tavernier, J., De Strooper, B., De Wit, J., and Schmucker, D. (2019). Nuclear import of the DSCAM-cytoplasmic domain drives signaling capable of inhibiting synapse formation. *The EMBO Journal*, 38(6):e99669.

Saint Marie, R. L. and Peters, A. (1985). The morphology and synaptic connections of spiny stellate neurons in monkey visual cortex (area 17): A golgi-electron microscopic study. *Journal of Comparative Neurology*, 233(2):213–235.

Sano, K., Tanihara, H., Heimark, R. L., Obata, S., Davidson, M., St. John, T., Taketani, S., and Suzuki, S. (1993). Protocadherins: A large family of cadherin-related molecules in central nervous system. *EMBO Journal*, 12(6):2249–2256.

Sardi, S. P., Murtie, J., Koirala, S., Patten, B. A., and Corfas, G. (2006). Presenilin-Dependent

ErbB4 Nuclear Signaling Regulates the Timing of Astrogenesis in the Developing Brain. *Cell*, 127(1):185–197.

Sato, T., Diehl, T. S., Narayanan, S., Funamoto, S., Ihara, Y., De Strooper, B., Steiner, H., Haass, C., and Wolfe, M. S. (2007). Active γ-secretase complexes contain only one of each component. *Journal of Biological Chemistry*, 282(47):33985–33993.

Sauer, F. C. (1935). Mitosis in the neural tube. *Journal of Comparative Neurology*, 62(2):377–405.

Scala, F., Kobak, D., Shan, S., Bernaerts, Y., Laturnus, S., Cadwell, C. R., Hartmanis, L., Froudarakis, E., Castro, J. R., Tan, Z. H., Papadopoulos, S., Patel, S. S., Sandberg, R., Berens, P., Jiang, X., and Tolias, A. S. (2019). Layer 4 of mouse neocortex differs in cell types and circuit organization between sensory areas. *Nature Communications*, 10(1):1–12.

Schaarschuch, A. and Hertel, N. (2018). Expression profile of N-cadherin and protocadherin-19 in postnatal mouse limbic structures. *Journal of Comparative Neurology*, 526(4):663–680.

Schaefer, A. W., Kabir, N., and Forscher, P. (2002). Filopodia and actin arcs guide the assembly and transport of two populations of microtubules with unique dynamic parameters in neuronal growth cones. *Journal of Cell Biology*, 158(1):139–152.

Schedin-Weiss, S., Caesar, I., Winblad, B., Blom, H., and Tjernberg, L. O. (2016). Super-resolution microscopy reveals γ-secretase at both sides of the neuronal synapse. *Acta neuropathologica communications*, 4:29.

Schmucker, D., Clemens, J. C., Shu, H., Worby, C. A., Xiao, J., Muda, M., Dixon, J. E., and Zipursky, L. (2000). Drosophila Dscam Is an Axon Guidance Receptor Exhibiting Extraordinary Molecular Diversity. *Cell*, 101:671–684.

Schoch, H., Kreibich, A. S., Ferri, S. L., White, R. S., Bohorquez, D., Banerjee, A., Port, R. G., Dow, H. C., Cordero, L., Pallathra, A. A., Kim, H., Li, H., Bilker, W. B., Hirano, S., Schultz, R. T., Borgmann-Winter, K., Hahn, C. G., Feldmeyer, D., Carlson, G. C., Abel, T., and Brodkin,

E. S. (2017). Sociability Deficits and Altered Amygdala Circuits in Mice Lacking Pcdh10, an Autism Associated Gene. *Biological Psychiatry*.

Schreiner, D. and Weiner, J. A. (2010). Combinatorial homophilic interaction between γ-protocadherin multimers greatly expands the molecular diversity of cell adhesion. *Proceedings of the National Academy of Sciences of the United States of America*, 107(33):14893–14898.

Schroeter, E. H., Kisslinger, J. A., and Kopan, R. (1998). Notch-1 signalling requires ligand-induced proteolytic release of intracellular domain. *Nature*, 393(6683):382–386.

Schwartz, M. L., Rakic, P., and Goldman-Rakic, P. S. (1991). Early phenotype expression of cortical neurons: Evidence that a subclass of migrating neurons have callosal axons. *Proceedings of the National Academy of Sciences of the United States of America*, 88(4):1354–1358.

Serafini, T., Colamarino, S., Leonardo, D., Wang, H., Beddington, R., Skarnes, W., and Marc, T.-L. (1996). Netrin-1 Is Required for Commissural Axon Guidance in the Developing Vertebrate Nervous System. *Cell*, 87(1):1001–1014.

Serratto, G. M., Pizzi, E., Murru, L., Mazzoleni, S., Pelucchi, S., Marcello, E., Mazzanti, M., Passafaro, M., and Bassani, S. (2020). The Epilepsy-Related Protein PCDH19 Regulates Tonic Inhibition, GABAAR Kinetics, and the Intrinsic Excitability of Hippocampal Neurons. *Molecular neurobiology*.

Shapiro, L., Fannon, A. M., Kwong, P. D., Thompson, A., Lehmann, M. S., Grübel, G., Legrand, J. F., Als-Nielson, J., Colma, D. R., and Hendricksonn, W. A. (1995). Structural basis of cell-cell adhesion by cadherins. *Nature*, 374:327–337.

Shi, Y., Kirwan, P., and Livesey, F. J. (2012). Directed differentiation of human pluripotent stem cells to cerebral cortex neurons and neural networks. *Nature Protocols*, 7(10):1836–1846.

Shoval, I., Ludwig, A., and Kalcheim, C. (2007). Antagonistic roles of full-length N-cadherin and its soluble BMP cleavage product in neural crest delamination. *Development*, 134(3):491–501.

Silva, C. G., Peyre, E., Adhikari, M. H., Tielens, S., Tanco, S., Van Damme, P., Magno, L., Krusy, N., Agirman, G., Magiera, M. M., Kessaris, N., Malgrange, B., Andrieux, A., Janke, C., and Nguyen, L. (2018). Cell-Intrinsic Control of Interneuron Migration Drives Cortical Morphogenesis. *Cell*, 172(5):1063–1078.e19.

Silva, C. G., Peyre, E., and Nguyen, L. (2019). Cell migration promotes dynamic cellular interactions to control cerebral cortex morphogenesis. *Nature Reviews Neuroscience*, 20(6):318–329.

Sopko, R., Huang, D., Preston, N., Chua, G., Papp, B., Kafadar, K., Snyder, M., Oliver, S. G., Cyert, M., Hughes, T. R., Boone, C., and Andrews, B. (2006). Mapping pathways and phenotypes by systematic gene overexpression. *Molecular Cell*, 21(3):319–330.

Sotomayor, M., Gaudet, R., and Corey, D. P. (2014). Sorting out a promiscuous superfamily: Towards cadherin connectomics. *Trends in Cell Biology*, 24(9):524–536.

Soudy, M., Anwar, A. M., Ahmed, E. A., Osama, A., Ezzeldin, S., Mahgoub, S., and Magdeldin, S. (2020). UniprotR: Retrieving and visualizing protein sequence and functional information from Universal Protein Resource (UniProt knowledgebase). *Journal of Proteomics*, 213(November 2019):103613.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., and Mesirov, J. P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America*, 102(43):15545–15550.

Subramanian, L., Bershteyn, M., Paredes, M. F., and Kriegstein, A. R. (2017). Dynamic behaviour of human neuroepithelial cells in the developing forebrain. *Nature Communications*, 8.

Suo, L., Lu, H., Ying, G., Capecchi, M. R., and Wu, Q. (2012). Protocadherin clusters and cell adhesion kinase regulate dendrite complexity through Rho GTPase. *Journal of Molecular Cell Biology*, 4(6):362–376.

Suzuki, K., Hayashi, Y., Nakahara, S., Kumazaki, H., Prox, J., Horiuchi, K., Zeng, M., Tanimura, S., Nishiyama, Y., Osawa, S., Sehara-Fujisawa, A., Saftig, P., Yokoshima, S., Fukuyama, T., Matsuki, N., Koyama, R., Tomita, T., and Iwatsubo, T. (2012). Activity-Dependent Proteolytic Cleavage of Neuroligin-1. *Neuron*, 76(2):410–422.

Suzuki, S. C. and Takeichi, M. (2008). Cadherins in neuronal morphogenesis and function. *Development Growth and Differentiation*, 50:S119–S130.

Tai, K., Kubota, M., Shiono, K., Tokutsu, H., and Suzuki, S. T. (2010). Adhesion properties and retinofugal expression of chicken protocadherin-19. *Brain Research*, 1344(1994):13–24.

Takeichi, M. (1977). Functional correlation between cell adhesive properties and some cell surface proteins. *The Journal of Cell Biology*, 75:464–474.

Takeichi, M. (2007). The cadherin superfamily in neuronal connections and interactions. *Nature Reviews Neuroscience*, 8(1):11–20.

Takeichi, M. (2014). Dynamic contacts: Rearranging adherens junctions to drive epithelial remodelling. *Nature Reviews Molecular Cell Biology*, 15(6):397–410.

Takiguchi-Hayashi, K., Sekiguchi, M., Ashigaki, S., Takamatsu, M., Hasegawa, H., Suzuki-Migishima, R., Yokoyama, M., Nakanishi, S., and Tanabe, Y. (2004). Generation of Reelin-Positive Marginal Zone Cells from the Caudomedial Wall of Telencephalic Vesicles. *Journal of Neuroscience*, 24(9):2286–2295.

Tan, C., Shard, C., Ranieri, E., Hynes, K., Pham, D. H., Leach, D., Buchanan, G., Corbett, M., Shoubridge, C., Kumar, R., Douglas, E., Nguyen, L. S., Mcmahon, J., Sadleir, L., Specchio, N., Marini, C., Guerrini, R., Moller, R. S., Depienne, C., Haan, E., Thomas, P. Q., Berkovic, S. F., Scheffer, I. E., and Gecz, J. (2015). Mutations of protocadherin 19 in female epilepsy (PCDH19-FE) lead to allopregnanolone deficiency. *Human Molecular Genetics*, 24(18):5250–5259.

Tasic, B., Yao, Z., Graybuck, L. T., Smith, K. A., Nguyen, T. N., Bertagnolli, D., Goldy, J., Garren, E., Economo, M. N., Viswanathan, S., Penn, O., Bakken, T., Menon, V., Miller, J., Fong, O., Hirokawa, K. E., Lathia, K., Rimorin, C., Tieu, M., Larsen, R., Casper, T., Hirschstein, D., Pendergraft, J., Barkan, E., Kroll, M., Parry, S., Nadiya, V., Sullivan, H. A., Kim, T. K., Szafer, A., Dee, N., Groblewski, P., Wickersham, I., Cetin, A., Harris, J. A., Sunkin, S. M., Madisen, L., Daigle, T. L., Looger, L., Bernard, A., Phillips, J., Lein, E., Hawrylycz, M., Svoboda, K., Jones, A. R., Koch, C., and Zeng, H. (2018). Shared and distinct transcriptomic cell types across neocortical areas. *Nature*, 563:72–78.

Theka, I., Sottile, F., Cammisa, M., Bonnin, S., Sanchez-Delgado, M., Di Vicino, U., Neguembor, M. V., Arumugam, K., Aulicino, F., Monk, D., Riccio, A., and Cosma, M. P. (2019). Wnt/β-catenin signaling pathway safeguards epigenetic stability and homeostasis of mouse embryonic stem cells. *Scientific Reports*, 9(1):1–19.

Thompson, K. R., Otis, K. O., Chen, D. Y., Zhao, Y., O'Dell, T. J., and Martin, K. C. (2004). Synapse to Nucleus Signaling during Long-Term Synaptic Plasticity. *Neuron*, 44(6):997–1009.

Thu, C. A., Chen, W. V., Rubinstein, R., Chevee, M., Wolcott, H. N., Felsovalyi, K. O., Tapia, J. C., Shapiro, L., Honig, B., and Maniatis, T. (2014). Generation of single cell identity by homophilic interactions between combinations of α, β and γ protocadherins Chan. *Cell*, 158(5):1045–1059.

Tian, L., Stefanidakis, M., Ning, L., Van Lint, P., Nyman-Huttunen, H., Libert, C., Itohara, S., Mishina, M., Rauvala, H., and Gahmberg, C. G. (2007). Activation of NMDA receptors promotes dendritic spine development through MMP-mediated ICAM-5 cleavage. *Journal of Cell Biology*, 178(4):687–700.

Togashi, H., Abe, K., Mizoguchi, A., Takaoka, K., Chisaka, O., and Takeichi, M. (2002). Cadherin regulates dendritic spine morphogenesis. *Neuron*, 35(1):77–89.

Tomita, T., Tanaka, S., Morohashi, Y., and Iwatsubo, T. (2006). Presenilin-dependent intramembrane cleavage of ephrin-BI. *Molecular Neurodegeneration*, 1(1):1–9.

Trivisano, M., Lucchi, C., Rustichelli, C., Terracciano, A., Cusmai, R., Ubertini, G. M., Giannone, G., Bertini, E. S., Vigevano, F., Gecz, J., Biagini, G., and Specchio, N. (2017). Reduced steroidogenesis in patients with PCDH19-female limited epilepsy. *Epilepsia*, 58(6):91–95.

Trovato, F., Parra, R., Pracucci, E., Landi, S., Cozzolino, O., Nardi, G., Cruciani, F., Pillai, V., Mosti, L., Cwetsch, A. W., Cancedda, L., Gritti, L., Sala, C., Verpelli, C., Maset, A., Lodovichi, C., and Ratto, G. M. (2020). Modelling genetic mosaicism of neurodevelopmental disorders in vivo by a Cre-amplifying fluorescent reporter. *Nature Communications*, 11(1):1–13.

Tucker, K. L., Meyer, M., and Barde, Y. A. (2001). Neurotrophins are required for nerve growth during development. *Nature Neuroscience*, 4(1):29–32.

Uemura, K., Kihara, T., Kuzuya, A., Okawa, K., Nishimoto, T., Bito, H., Ninomiya, H., Sugimoto, H., Kinoshita, A., and Shimohama, S. (2006a). Activity-dependent regulation of β-catenin via ϵ-cleavage of N-cadherin. *Biochemical and Biophysical Research Communications*, 345(3):951–958.

Uemura, K., Kihara, T., Kuzuya, A., Okawa, K., Nishimoto, T., Ninomiya, H., Sugimoto, H., Kinoshita, A., and Shimohama, S. (2006b). Characterization of sequential N-cadherin cleavage by ADAM10 and PS1. *Neuroscience Letters*, 402(3):278–283.

Uemura, M., Nakao, S., Suzuki, S. T., Takeichi, M., and Hirano, S. (2007). OL-protocadherin is essential for growth of striatal axons and thalamocortical projections. *Nature Neuroscience*, 10(9):1151–1159.

Utton, M. A., Eickholt, B., Howell, F. V., Wallis, J., and Doherty, P. (2001). Soluble N-cadherin stimulates fibroblast growth factor receptor dependent neurite outgrowth and N-cadherin and the fibroblast growth factor receptor co-cluster in cells. *Journal of Neurochemistry*, 76(5):1421–1430.

van Tetering, G., van Diest, P., Verlaan, I., van der Wall, E., Kopan, R., and Vooijs, M. (2009). Metalloprotease ADAM10 is required for Notch1 site 2 cleavage. *Journal of Biological Chemistry*, 284(45):31018–31027.

Vanderklish, P., Saido, T. C., Gal, C., Arai, A., and Lynch, G. (1995). Proteolysis of spectrin by calpain accompanies theta-burst stimulation in cultured hippocampal slices. *Molecular Brain Research*, 32(1):25–35.

Vanhalst, K., Kools, P., Staes, K., Van Roy, F., and Redies, C. (2005). δ-Protocadherins: A gene family expressed differentially in the mouse brain. *Cellular and Molecular Life Sciences*, 62(11):1247–1259.

Vassar, R., Bennett, B. D., Babu-Khan, S., Kahn, S., Mendiaz, E. A., Denis, P., Teplow, D. B., Ross, S., Amarante, P., Loeloff, R., Luo, Y., Fisher, S., Fuller, J., Edenson, S., Lile, J., Jarosinski, M. A., Biere, A. L., Curran, E., Burgess, T., Louis, J. C., Collins, F., Treanor, J., Rogers, G., and Citron, M. (1999). β-Secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE. *Science*, 286(5440):735–741.

Vernes, S. C., Oliver, P. L., Spiteri, E., Lockstone, H. E., Puliyadi, R., Taylor, J. M., Ho, J., Mombereau, C., Brewer, A., Lowy, E., Nicod, J., Groszer, M., Baban, D., Sahgal, N., Cazier, J. B., Ragoussis, J., Davies, K. E., Geschwind, D. H., and Fisher, S. E. (2011). FOXP2 regulates gene networks implicated in neurite outgrowth in the developing brain. *PLoS Genetics*, 7(7).

Vicario, A., Kisiswa, L., Tann, J. Y., Kelly, C. E., and Ibáñez, C. F. (2015). Neuron-type-specific signaling by the p75NTR death receptor is regulated by differential proteolytic cleavage. *Journal of Cell Science*, 128(8):1507–1517.

Vlaskamp, D. R., Bassett, A. S., Sullivan, J. E., Robblee, J., Sadleir, L. G., Scheffer, I. E., and Andrade, D. M. (2019). Schizophrenia is a later-onset feature of PCDH19 Girls Clustering Epilepsy. *Epilepsia*, 60(3):429–440.

Whitford, K. L., Dijkhuizen, P., Polleux, F., and Ghosh, A. (2002). Molecular control of cortical dendrite development. *Annual Review of Neuroscience*, 25:127–149.

Wichterle, H., Lieberam, I., Porter, J. A., and Jessell, T. M. (2002). Directed Differentiation of Embryonic Stem Cells into Motor Neurons. *Cell*, 110:385–397.

Wieland, I., Jakubiczka, S., Muschke, P., Cohen, M., Thiele, H., Gerlach, K. L., Adams, R. H., and Wieacker, P. (2004). Mutations of the ephrin-B1 gene cause craniofrontonasal syndrome. *American Journal of Human Genetics*, 74(6):1209–1215.

Williams, C. C., Allison, J. G., Vidal, G. A., Burow, M. E., Beckman, B. S., Marrero, L., and Jones, F. E. (2004). The ERBB4/HER4 receptor tyrosine kinase regulates gene expression by functioning as a STAT5A nuclear chaperone. *Journal of Cell Biology*, 167(3):469–478.

Williams, R. L., Hilton, D. J., Pease, S., Willson, T. A., Stewart, C. L., Gearing, D. P., Wagner, E. F., Metcalf, D., Nicola, N. A., and Gough, N. M. (1988). Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature*, 336(6200):684–687.

Wolfe, M. S., Xia, W., Ostaszewski, B. L., Diehl, T. S., Kimberly, W. T., and Selkoe, D. J. (1999). Two transmembrane aspartates in presenilin-1 required for presenilin endoproteolysis and γ-secretase activity. *Nature*, 398(6727):513–517.

Wolverton, T. and Lalande, M. (2001). Identification and characterization of three members of a novel subclass of protocadherins. *Genomics*, 76(1-3):66–72.

Wu, Q. and Maniatis, T. (1999). A Striking Organization of a Large Family of Human Neural Cadherin-like Cell Adhesion Genes. *Cell*, 97:779–790.

Wu, Q., Zhang, T., Cheng, J. F., Kim, Y., Grimwood, J., Schmutz, J., Dickson, M., Noonan, J. P., Zhang, M. Q., Myers, R. M., and Maniatis, T. (2001). Comparative DNA sequence analysis of mouse and human protocadherin gene clusters. *Genome Research*, 11(3):389–404.

Xiao, X., Zheng, F., Chang, H., Ma, Y., Yao, Y. G., Luo, X. J., and Li, M. (2018). The Gene Encoding Protocadherin 9 (PCDH9), a Novel Risk Factor for Major Depressive Disorder. *Neuropsychopharmacology*, 43(5):1128–1137.

Yamakawa, K., Huo, Y. K., Haendel, M. A., Hubert, R., Chen, X. N., Lyons, G. E., and Korenberg, J. R. (1998). DSCAM: A novel member of the immunoglobulin superfamily maps in a Down

syndrome region and is involved in the development of the nervous system. *Human Molecular Genetics*, 7(2):227–237.

Yang, J., Bogerd, H. P., Wang, P. J., Page, D. C., and Cullen, B. R. (2001). Two closely related human nuclear export factors utilize entirely distinct export pathways. *Molecular Cell*, 8(2):397–406.

Yang, Y. T., Wang, C. L., and Van Aelst, L. (2012). DOCK7 interacts with TACC3 to regulate interkinetic nuclear migration and cortical neurogenesis. *Nature Neuroscience*, 15(9):1201–1210.

Ying, Q. L. and Smith, A. (2017). The Art of Capturing Pluripotency: Creating the Right Culture. *Stem Cell Reports*, 8(6):1457–1464.

Yoon, K., Nery, S., Rutlin, M. L., Radtke, F., Fishell, G., and Gaiano, N. (2004). Fibroblast growth factor receptor signaling promotes radial glial identity and interacts with Notch1 signaling in telencephalic progenitors. *Journal of Neuroscience*, 24(43):9497–9506.

Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). ClusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS A Journal of Integrative Biology*, 16(5):284–287.

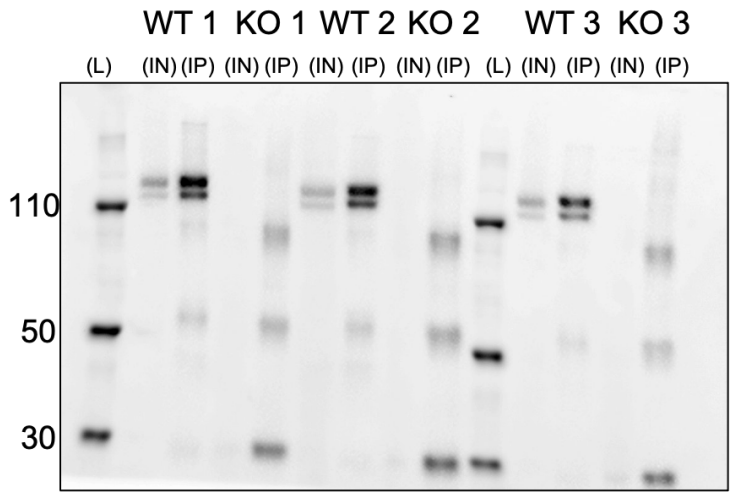
Yu, G., Wang, L. G., Yan, G. R., and He, Q. Y. (2015). DOSE: An R/Bioconductor package for disease ontology semantic and enrichment analysis. *Bioinformatics*, 31(4):608–609.

Chapter 8

Appendix

8.1 PCDH19 immunoprecipitation

(a)



(b)

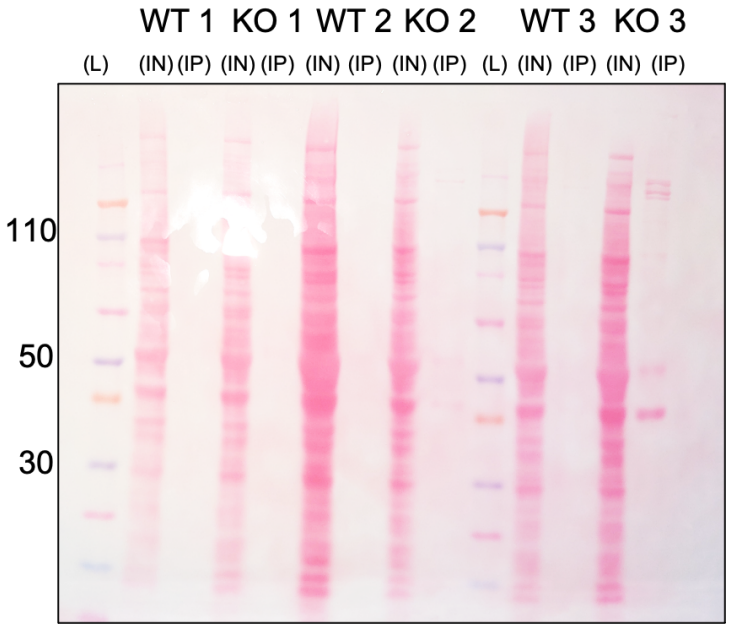
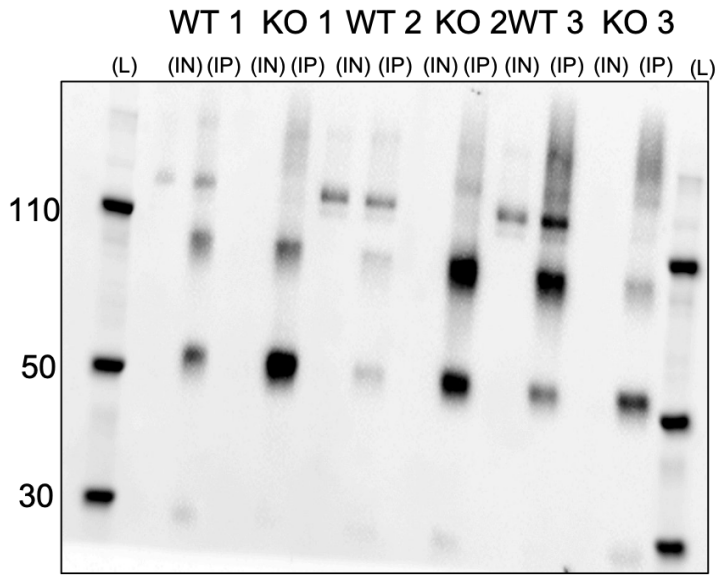


Figure 8.1: PCDH19 IP for LC-MS on cortex samples, uncropped. **(a)** Anti-PCDH19 western blot after anti-PCDH19 immunoprecipitation. **(b)** Ponceau stain.

(a)



(b)

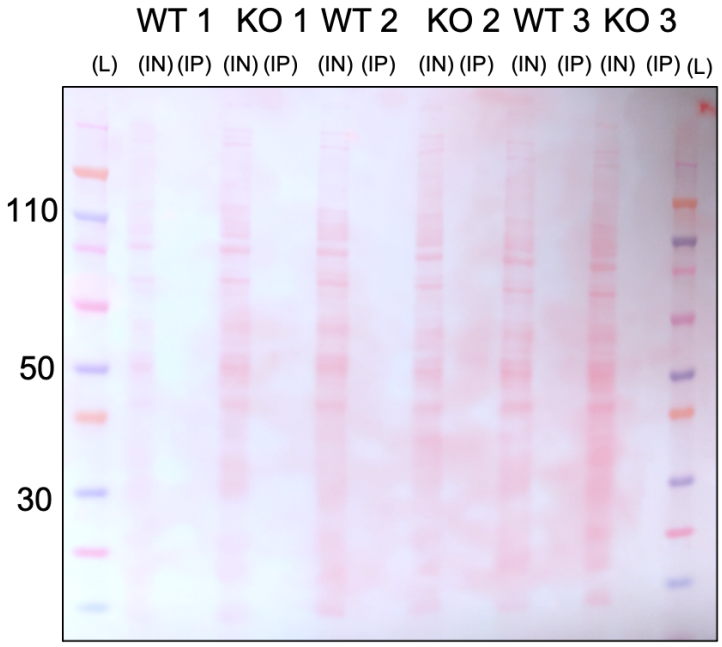


Figure 8.2: PCDH19 IP for LC-MS on E11.5 samples, uncropped. (a) Anti-PCDH19 western blot after anti-PCDH19 immunoprecipitation. **(b)** Ponceau stain.

8.6 DEGs with p. adjusted cut-off < 0.01

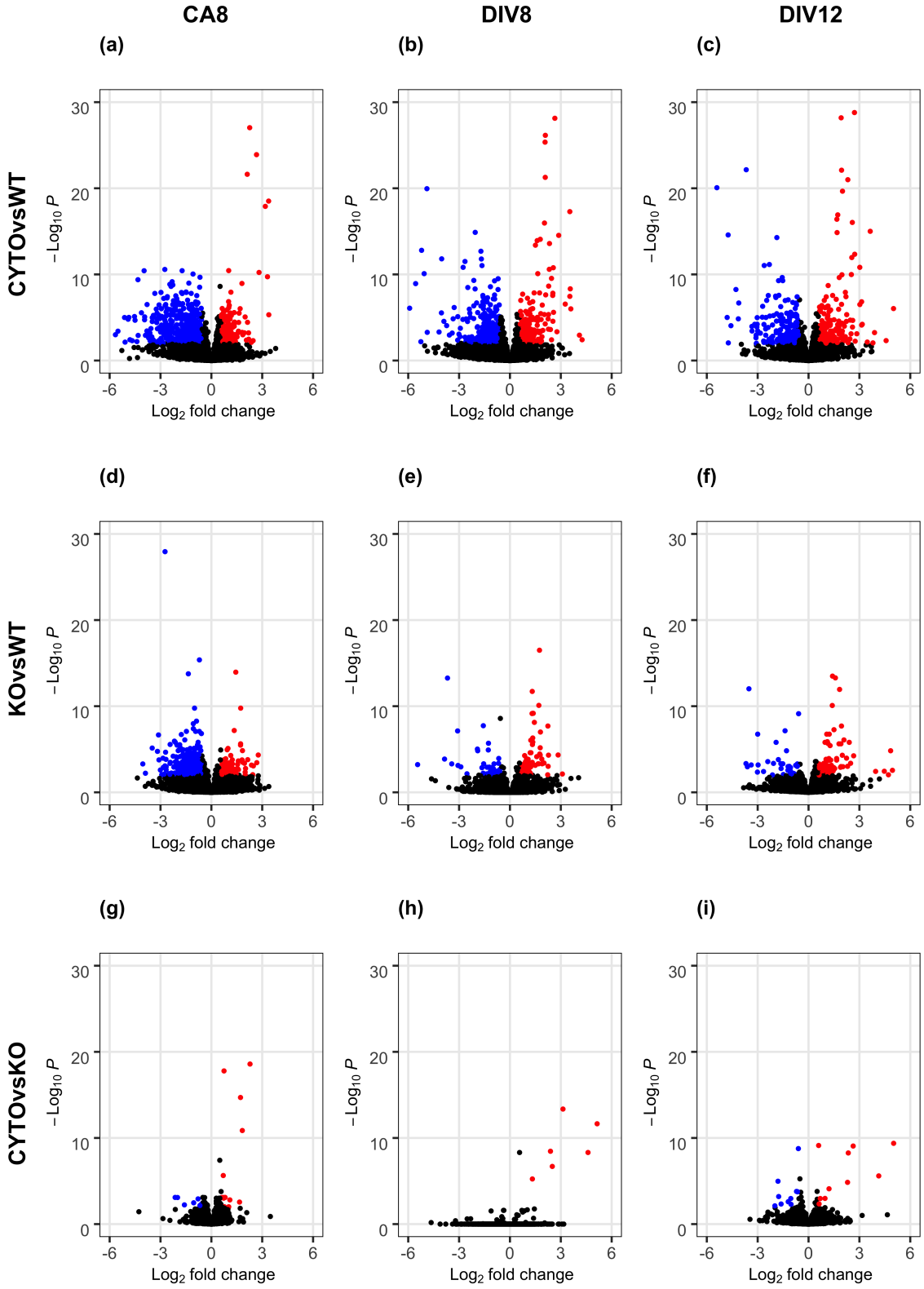


Figure 8.3: Vulcano plots showing DEGs for CYTO *vs* WT, KO *vs* WT and CYTO *vs* KO at CA8, DIV8 and DIV12 at p.adjusted < 0.01. Downregulated genes are shown in blue, upregulated genes are shown in red. Threshold for significance is p.adjusted < 0.01 and $\log_2$ cutoff > 0.58.

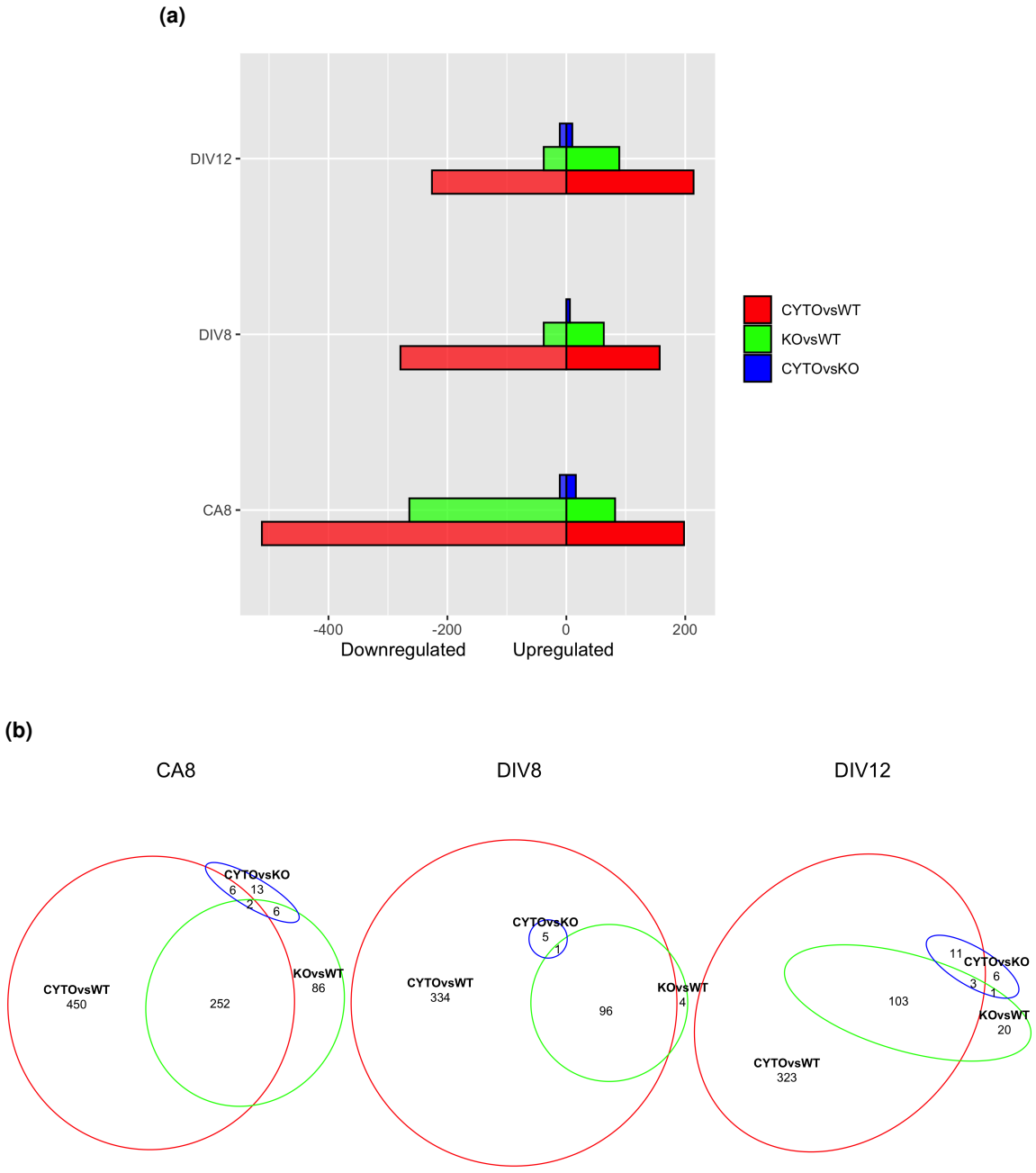


Figure 8.4: Number of DEGs at p.adj $<$ 0.01. (a) Total upregulated and downregulated number of genes at each timepoint for CYTO *vs* WT, KO *vs* WT and CYTO *vs* KO. **(b)** Venn diagram showing common differentially expressed genes between CYTO, KO and WT cells at timepoints CA8, DIV8 and DIV12. Threshold for significance is p.adjusted $<$ 0.01 and $\log_2$ cutoff $>$ 0.58.

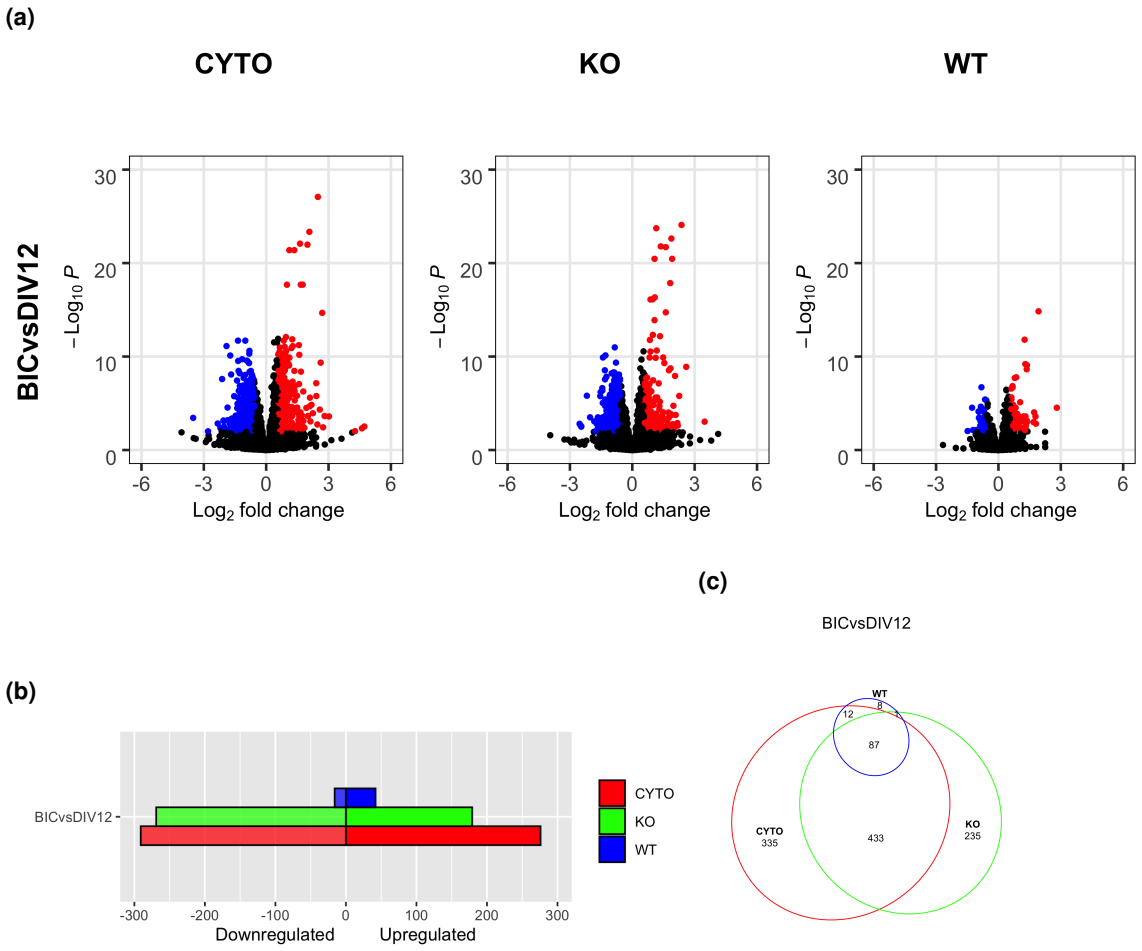


Figure 8.5: DEGs for bicuculline treated (DIV12 + BIC) *vs* untreated (DIV12) neurons for each cell line (CYTO, KO, WT) with p.adjusted cut-off $<$0.01. **(a)** Downregulated genes are shown in blue, upregulated genes are shown in red. Threshold for p.adj value is 0.01, $\log_2$ fold change cut off $>$ 0.58. **(b)** Total upregulated and downregulated number of genes for each comparison. **(c)** Overlapping genes up or downregulated in response to bicuculline in WT, CYTO and KO neurons at DIV12. Threshold for significance is p.adjusted $<$ 0.01 and $\log_2$ cutoff $>$ 0.58.

8.7 DEGs with p. adjusted cut-off < 0.05

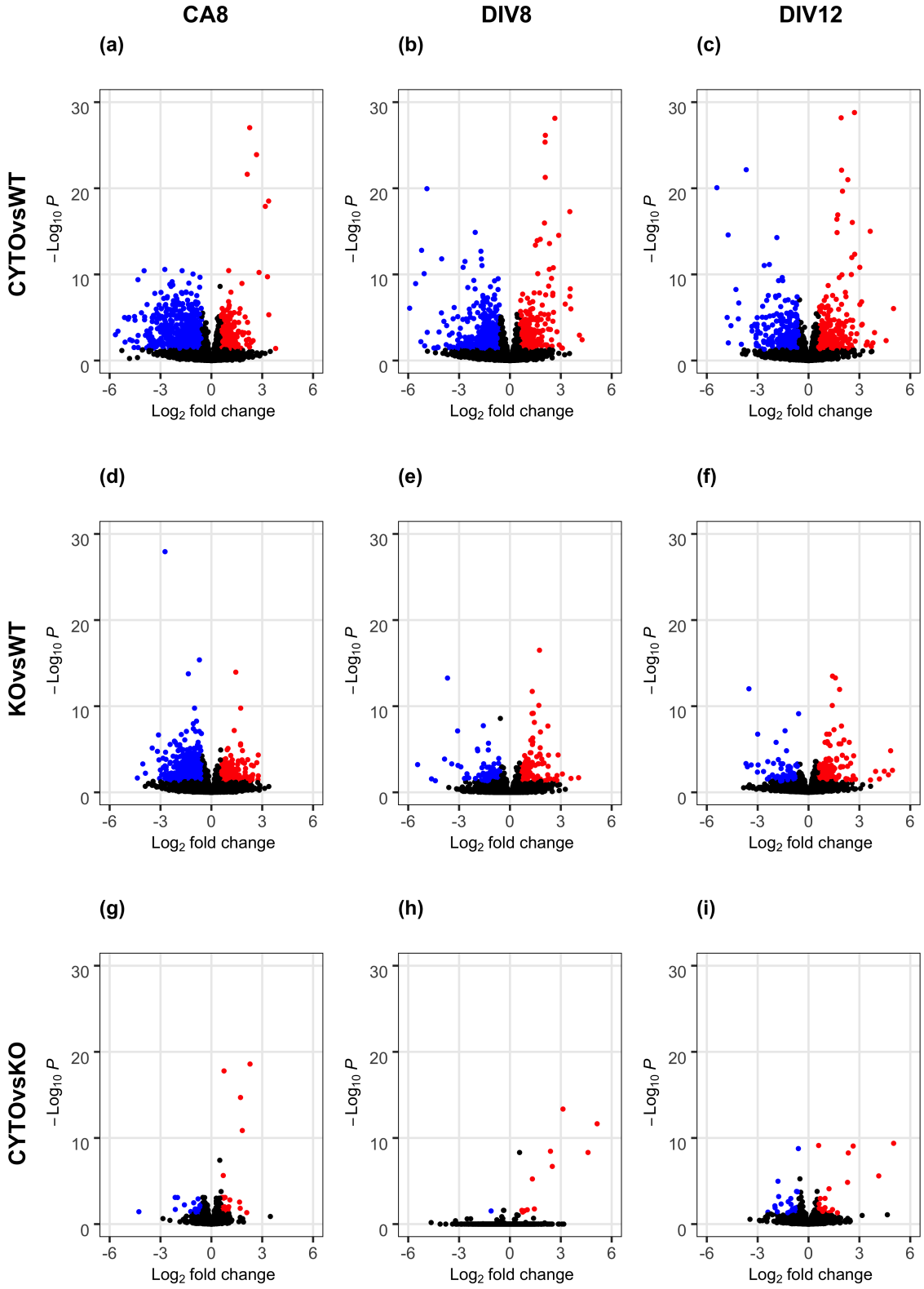


Figure 8.6: Vulcano plots showing DEGs for CYTO *vs* WT, KO *vs* WT and CYTO *vs* KO at CA8, DIV8 and DIV12 at p.adj < 0.05. Downregulated genes are shown in blue, upregulated genes are shown in red. Threshold for significance was p.adj < 0.05 and $\log_2$ cutoff > 0.58.

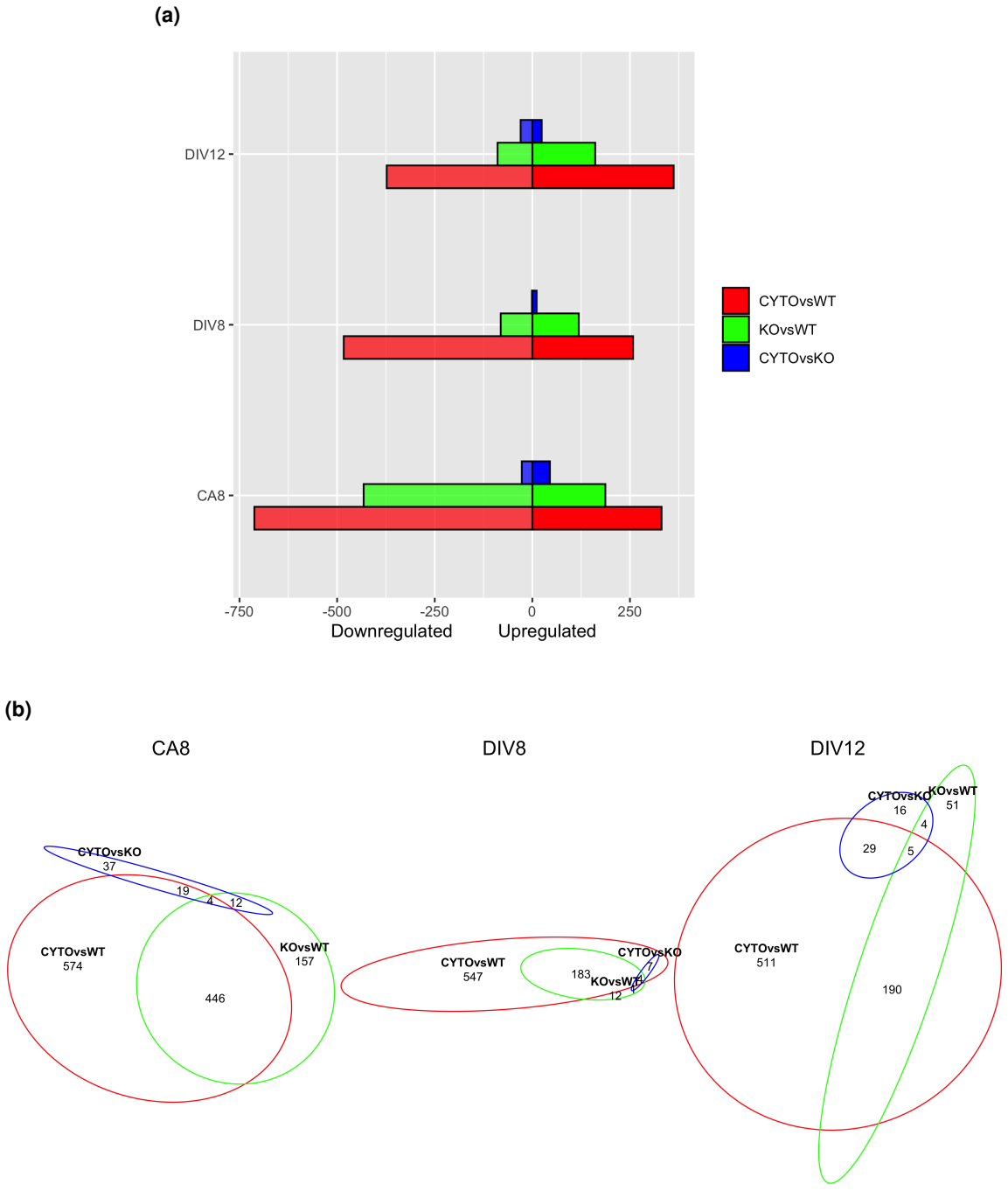


Figure 8.7: Number of DEGs at p.adj $<$ 0.05. (a) Total upregulated and downregulated number of genes at each timepoint for CYTO *vs* WT, KO *vs* WT and CYTO *vs* KO. **(b)** Venn diagram showing common differentially expressed genes between CYTO, KO and WT cells at timepoints CA8, DIV8 and DIV12. Threshold for significance was p.adj $<$ 0.05 and $\log_2$ cutoff $>$ 0.58.

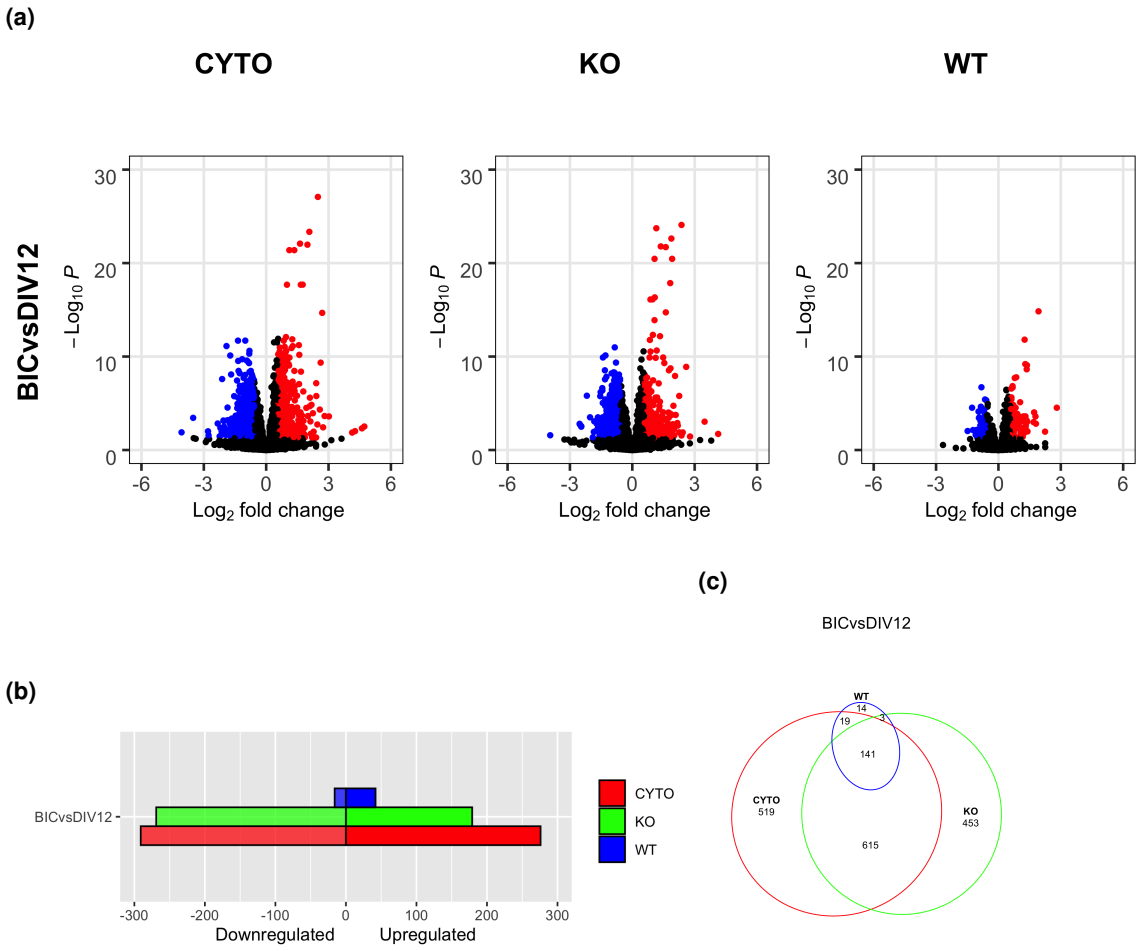


Figure 8.8: DEGs for bicuculline treated (DIV12 + BIC) *vs* untreated (DIV12) neurons for each cell line (CYTO, KO, WT) with p.adjusted cut-off $<$0.05. (a) Downregulated genes are shown in blue, upregulated genes are shown in red. Threshold for p.adj value is < 0.05, $\log_2$ fold change cut off > 0.58. **(b)** Total upregulated and downregulated number of genes for each comparison. **(c)** Overlapping genes up or downregulated in response to bicuculline in WT, CYTO and KO neurons at DIV12. Threshold for significance was p.adj < 0.05 and $\log_2$ cutoff > 0.58.