Synaptic gene *DLG2* contributes to cortical interneuron development

By

Asmaa Ghazwani

A thesis submitted to Cardiff University for the degree of Doctor of Philosophy

June 2022

Abstract

Several genetic as well as environmental factors have been identified concerning schizophrenia (SCZ). For example, the uncommon deletion of the gene, *DLG2,* has been observed in numerous individuals suffering from SCZ. This gene encodes a structural scaffolding protein found in fully developed synapses. For exploring DLG2's proposed function concerning early inhibitory cortical development and establishing this gene's possible relevance to pathological manifestations, wild-type (WT) and *DLG2*$^{-/-}$ human embryonic stem cell (hESC) lines underwent differentiation into cortical inhibitory interneurons. Their properties were observed at numerous time points using various methods. The present study implies that DLG2 protein is expressed earlier than currently believed and before synaptic development in cortical interneurons. Further, DLG2's novel functions during cortical interneuron development were identified, such as the neural progenitor cell gene expression governance and the extracellular matrix's (ECM) adherence. Also, some altered biological processes identified through single-cell RNA sequencing (scRNAseq) concerned the *regulation of translational initiation*, *DNA transcription* and *initiation, RNA splicing* and *regulation of chromosome organization.* The *DLG2*$^{-/-}$ gene set was not enriched for SCZ common risk variants. Nonetheless, analysing WT DEGs for a link with SCZ common risk variants showed that the developing neuronal cell subtype was associated with SCZ regardless of DLG2's contribution. A greater frequency of immature neuronal cells was induced in *DLG2*$^{-/-}$ cultures, typified by aberrations in migration, morphology, and protein expression, such as for the transcription factors, NKX2.1 and OLIG2. Embryonic mice brains were used to determine the murine interneurons' properties that were homozygous or heterozygous for *Dlg2.* Heterozygous brains showed patterns of diminishing proportions of CGE COUPTF2^{+} cells, whereas homozygous brains showed decreasing frequencies of NKX2.1 and COUPTF2 positive cells. Combining all *in vitro* and *in vivo* observations, *DLG2* significantly contributes to interneuron development/birth, maturation, differentiation, and migratory characteristics.

Acknowledgments

First, I praise Allah, Almighty God, for offering me a healthy environment to complete this project. I wish to extend my thanks and voice my sincere appreciation to my supervisors Dr Jenny Shin, Dr Andrew Pocklington, and Dr Daniel D'Andrea; without their generous and valuable support, I would not have become the scientist I am today. Their outstanding patience and input enabled me to successfully finish my PhD, and their tutorials were always inspiring and held in an extremely friendly atmosphere. I am immensely grateful to Jenny for giving me the opportunity to undertake this doctorate. I also need to warmly thank Dr Daniel for his help in tackling the bioinformatics components in this research and for giving me such useful and clear explanations. Finally, I want to thank my country, Kingdom of Saudi Arabia, for funding me; without their generous funding, this research would not have been possible.

I wish to thank all the colleagues and friends I worked with in the Neuroscience and Mental Health Research Institution (NMHRI) (particularly Dr Bret Sanders and Dr Karima Azzouni) for enriching my time in the lab and throughout the four years of my PhD. The NMHRI technical team's help and support proved to be invaluable. I also wish to acknowledge Angela and Joanne for helping me with scRNA sequencing. I must single out that every PhD has its challenges, especially during the COVID-19 pandemic, and my third and fourth years were tough.

My husband was my proverbial rock on this journey, which I could not have finished without him. My grandfather and my sisters provided me with the confidence and encouragement to finish this research, particularly when I began to flag during the exhausting write-up stage — thank you both for your prayers and love. Finally, I am dedicating this thesis to my dear parents who have always dreamed that I would become a doctor one day.

List of abbreviation

5HT3A	Inotropic serotoninergic receptor expression
BMP	Bone morphogenic protein
CCK	Cholecystokinin
CGE	Caudal ganglionic eminence
cIN(s)	Inhibitory cortical interneuron(s)
CNV	Copy number variants
DE	Differential expression
DEGs	Differentially expressed genes
Dev N	Developing neuron
Div RG	Dividing radial glia
DLG	Discs large homologue
dLGE	Dorsal lateral ganglionic eminence
DLPFC	Dorsolateral prefrontal cortex
DMEM/F-12	Dulbecco's modified Eagle medium: nutrient mixture F-12
ECM	Extracellular matrix
FGF	Fibroblast growth factor
GABA	γ-aminobutyric acid
GE	Ganglionic eminence
GK	Guanylate kinase
GO	Gene ontology
GWAS	Genome wide association studies
hESCs	Human embryonic stem cells
hPSCs	Human pluripotent stem cells
IPC	Intermediate precursor cell
IPC DN	Intermediate precursor cell-derived neuron
iPSCs	Induced pluripotent stem cells
KO	Knockout
LGE	Lateral ganglionic eminence
MAGMA	Multi-marker Analysis of GenoMic Annotation
MAGUKs	Membrane associated guanylate kinases
MGE	Medial ganglionic eminence
MZ	Marginal zone
NCS	Neural stem cell/neuroepithelial stem cell
Newb N	Newborn neuron
NMDAR	N-methyl-d-aspartate receptor

NOS	Nitric oxide synthase
NPCs	Neural precursor cells
NPY	Neuropeptide Y
OR	Odds ratio
PBS	Phosphate buffered saline
PCA	Principal component analysis
PCW	Post conception week
PFA	Paraformaldehyde
PFC	Prefrontal cortex
POA	Preoptic area
PSD	Postsynaptic density
PV	Parvalbumin
QC	Quality control
RA	Retinoic acid
RG	Radial glia
scRNAseq	Single cell RNA sequencing
SCZ	Schizophrenia
SH3	SRC homology 3
SHH	Sonic hedgehog pathway
SNP	single nucleotide polymorphisms
SNV	Single nucleotide variants
SST	Somatostatin
SVZ	Subventricular zones
t-SNE	t-distributed stochastic neighbour embedding
TF/ TFs	Transcription factor/ Transcription factors
Tran RG	Transitioning radial glia
UMI	Unique molecular identifier
VIP	Vasointestinal peptide
VZ	Ventricular zone
WPC	Weeks posts conception
WT	Wild type

List of contents

List of figures

List of tables

1. Introduction

1.1. Cortical interneuron development

1.1.1. Neurons in cortex

The neocortex is the brain area where the processing of higher cognitive functions occurs. Such activity is accomplished via equilibrium of neuronal excitation and inhibition within cellular cortical circuits (Wonders and Anderson, 2006). There are two neuronal types in the cortex, the projection neurons and the interneurons, that form a complex neural network and cooperate to ensure efficient cortical function. The pyramidal projection neurons (also known as glutamatergic/excitatory neurons) are predominantly glutamatergic, functioning through the axonal secretion of the excitatory neurotransmitter, glutamate, which acts on non-cortical and cortical targets (Tremblay et al., 2016). In contrast to the excitatory neurons, cortical interneurons (cINs) are inhibitory and typically modify the excitatory activity of the projection glutamatergic neurons by the release of neurotransmitter, γ-aminobutyric acid (GABA) (Rudy et al., 2011; Tremblay et al., 2016). This occurs at the both individual neuronal level, resulting in reduced neuronal firing, and network level wherein the oscillatory activity is coordinated and circuit over-excitation is prohibited, which are vital to enable normal cortical performance (Rossignol, 2011).

1.1.2. Origin of cortical interneurons

GABAergic interneurons exhibit considerable heterogeneity; the differences between mature cortical interneuron types occur gradually during development. These adult interneurons are derived from a handful of neuronal stem and progenitor cell types (Kelsom and Lu, 2013). During embryonic development, interneurons are mainly produced within the ganglionic eminence (GE), a temporary anatomical structure present throughout embryogenesis and found within the ventral region of the telencephalon (Figure 1.1) (Corbin and Butt, 2011; Kelsom and Lu, 2013; Sultan et al., 2013). There are three identifiable GE structures that are described in rodents, which are medial (MGE), caudal (CGE) and lateral (LGE) (Figure 1.1). Their names reflect their rostral-caudal and medio-lateral position within the telencephalon; however, they are also differentiated according to the cellular types they generate and molecular markers expressed (Kelsom and Lu, 2013; Sultan et al., 2013). During embryogenesis, the GE triad further develop and ultimately fuse until they no longer exist within the adult brain. Despite being subject of controversy, it is largely agreed that the LGE does not produce cortical interneurons, but rather, it is the origin of a few olfactory bulb interneurons and striatal projection neurons (Batista-Brito et al., 2008; Deacon et al., 1994; Rossignol, 2011; Waclaw et al., 2009). In general, the cortical interneurons are principally derived

from the embryonic MGE and CGE in the developing brain (Corbin and Butt, 2011; Rossignol, 2011; Wonders and Anderson, 2006).

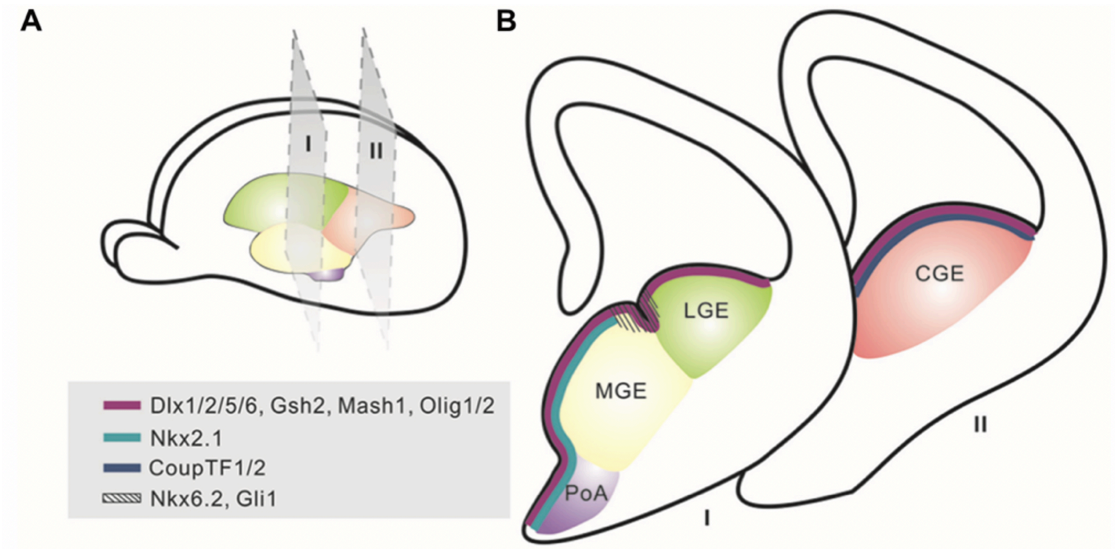


Figure 1.1 Origins of neocortical interneurons. A) Shows two coronal locations of slices in the embryonic mouse brain. **B)** Precursor cells are located in the ventral telencephalon proliferative regions, particularly within the MGE and CGE, the major sources of neocortical interneurons, with a small number produced in the preoptic area (POA). Throughout the subpallial germinal zones, a number of transcription factors are expressed in different patterns; Dlx1/2/5/6, Mash1, Olig1/2, and Gsh2 (magenta) are expressed in ventricular zone (VZ) of the entire GE region, whereas Nkx2.1 (light blue) transcription factor expressed within MGE/ POA, and CoupTF1/2 (dark blue) transcription factor expressed within the CGE. However, Nkx6.2 and Gli1 (hashed lines) are restrictedly expressed in the sulcus area between MGE and LGE. From (Sultan et al., 2013).

1.1.3. Key morphogens and transcription factors in GEs

Within the telencephalon, there are proliferative and non-proliferative regions where neurons migrate. Discrete progenitor domains are defined by morphogen-regulated transcription factors (TFs) (Cohen et al., 2013), which in part determine the neuroepithelial cells' neuronal subtype identity. Morphogens that pattern the telencephalon include fibroblast growth factor FGF, sonic hedgehog pathway (SHH), and Wnt signalling (Hebert and Fishell, 2008; Sousa and Fishell, 2010). To direct neural tube patterning throughout morphogenic development, Wnt signalling and the Shh pathway directly oppose each other. Shh is known to have a ventralising function (Dessaud et al., 2008), which is antagonized by Wnt signalling (Ulloa and Marti, 2010). Furthermore, a Shh downstream signalling molecule, *Gli3*, appears to contribute to this process. *Gli3* critically functions in the dorsal telencephalon where it opposes the effects of Shh and suppresses Shh signalling (Rallu et al., 2002). In this process, Wnt provokes the expression of *Gli3* (Aviles et al., 2013; Ulloa and Marti, 2010).

The early-acting TFs include *Pax6*, *Foxg1*, *Gli1*/*2*/*3*, *Gsx2*, *Ascl1* (*Mash1*), *Nkx2.1* and *Six3* (Hebert and Fishell, 2008; Kessaris et al., 2014; Long et al., 2009; Sousa and Fishell, 2010). These TFs, expressed prior to the generation of cortical interneurons,

influence cortical interneuronal development by limiting precursor differentiation potential (Kessaris et al., 2014).

The *Dlx* group of homeobox genes, are essential TFs, particularly *Dlx1, 2, 5* and *6*, which in combination with *Mash1*, are expressed during neurogenesis in the majority of subpallial neural precursor cells (Figure 1.1). Absent *Dlx1/2* function results in a failure of *Dlx5/6* expression (Long et al., 2009) and a consequent failure to specify the identity of dorsal CGE progenitors and a reduction of the interneuron migration from the MGE precursors. Furthermore, it has been reported that knockouts (KO) of *Mash1*, exhibited a decrease in S-phase cell cycle precursor cells in GE ventricular zone (VZ), significant loss of neuronal precursors particularly in the MGE subventricular zones (SVZ) and premature differentiation of VZ cells in the LGE (at E12.5 and E15.5) together with a clear reduction in the numbers of cortical interneurons in the basal ganglia and cortex at E18.5. However, the number of striatal neurons were unaffected (Casarosa et al., 1999; Long et al., 2009). Despite shared expression of *Dlx* and *Mash1* (Casarosa et al., 1999), each GE displays genetic diversity with precise gene expression pattern (Flames et al., 2007).

1.1.4. MGE

For many years, it has been believed that the primary origin of GABAergic interneurons is the MGE; mice studies have shown that around 60% of cortical interneurons arise from the MGE (Butt et al., 2005; Kelsom and Lu, 2013; Kessaris et al., 2014; Wonders and Anderson, 2006). Regional identity of MGE is recognised by homeobox TF's NKX2.1 expression (Xu et al., 2010a) which is activated by SHH signalling in the ventral forebrain; NKX2.1 expression levels vary between the dorsal and ventro-medial areas of the MGE. The role of *Nkx2.1* was demonstrated by KO studies, which led to a 50% decrease of cortical GABAergic cells (Sussel et al., 1999). The timing of *Nkx2.1* gene expression varies across the different zones, with early expression (E10.5) observed in the VZ and SVZ zones and later expression (E13.5) in the preoptic area (POA). In the MGE *Nkx2.1* regulates proliferation and differentiation of progenitors, development of newborn neurons and their further differentiation into the major classes of GABAergic interneurons. Continued *Nkx2.1* expression in neurons results in striatal interneuron development, while discontinued expression results in the neurons following the cortical interneuron fate (Butt et al., 2008; Nobrega-Pereira et al., 2008). It is noteworthy that cholinergic neurons in the ventral telencephalon are derived from MGE, and $NKX2.1^{+}$ progenitors also give rise to cholinergic neurons. Cholinergic neurons are identified by their expression of choline acetyltransferase (ChAT) (Marin et al., 2000).

Lhx6 is downstream of *Nkx2.1* and plays an essential role in specifying the MGE-originating interneurons (Figure 1.2) (Du et al., 2008; Zhao et al., 2003). Interneurons lacking *Lhx6* do not have the capacity to integrate effectively to their corresponding cortical layers (Zhao et al., 2008). Downstream of *Lhx6* and *Nkx2.1* is *Sox6*, an additional important TF in MGE, which functions in combination with either *Nkx2.1* or *Lhx6* (Azim et al., 2009; Batista-Brito et al., 2009). As with LHX6, SOX6 is expressed in MGE-originated GABAergic interneurons from the early post-mitotic phase to maturity. During the mouse forebrain developmental phase, *Sox6* is strongly expressed in MGE-derived cINs, and its KO results in a significant reduction in the major interneurons subtypes. Additionally, defects in transition from tangential to radial migration were observed in *Sox6* mutant MGE-derived interneurons (Batista-Brito et al., 2009).

Expression profiling, together with other functional studies, has identified further transcription factor encoding genes within cortical interneurons, e.g. *Cux2*, *and MafB* (*bMaf*) (Cobos et al., 2006; Zhao et al., 2008). The expression of some of these genes remains to be validated; consequently, their specific roles in the maturation process of cINs have not been fully elucidated (Kessaris et al., 2014).

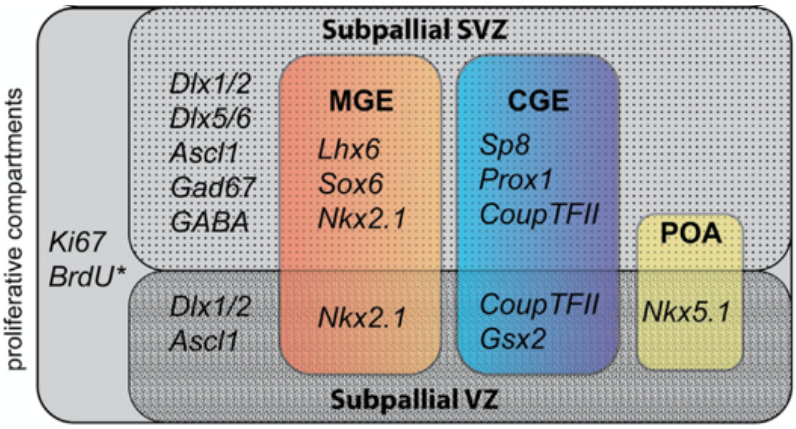


Figure 1.2 Key transcription factors expressed during early development and differentiation of mouse interneurons. Boxes illustrate the common and specific markers in the SVZ (medium grey) and VZ (dark grey) of the MGE (orange) and the CGE (blue). Each region generates specific cortical interneurons. From (Laclef and Metin, 2018).

1.1.5. CGE

The second largest source of interneuron precursors is the CGE, which generate around 30% of the total number of cINs (Kelsom and Lu, 2013; Lee et al., 2010; Miyoshi et al., 2010; Nery et al., 2002). The CGE expresses characteristics common to both the MGE and LGE, and gene expression profiles show a substantial overlap between the CGE and LGE, indicating that the CGE arises as a caudal extension of the LGE (Flames et al., 2007). For example, CGE expresses *Gsx2* (also called *Gsh2*) which is downstream target of SHH and regulates the development of the striatum (Yun et al., 2003), is also a

TF required for LGE patterning (Waclaw et al., 2009). While there is much to be explored to define the molecular architecture that determine CGE-derived interneuron fate the genes *Gsx2*, *CoupTF2* and *Prox1* are implicated in fate determination of CGE derivatives (Figure 1.2) (Kanatani et al., 2008; Miyoshi et al., 2015; Xu et al., 2010a). Of these, *Prox1* is maintained in the mature cortex and has been shown to be essential to the acquisition of CGE-derived interneuron characteristics (Miyoshi et al., 2015).

In the mouse *CoupTF2* (*Nr2f2*), an orphan nuclear receptor, is critical for gene expression regulation at specific time during development (called temporal specification) of neural precursor cells derived from CGE (Naka et al., 2008). It is also a prerequisite for the differentiation and tangential migration of CGE-interneurons (Kanatani et al., 2008). In the subpallium of the human embryonic telencephalon, a gradient of COUPTF2 expression is observed i.e. rostral (low) to caudal (high) and dorsal (low) to ventral (high) of positive cells (Alzu'bi et al., 2017a). While COUPTF2 positive cells in ventral CGE (vCGE) are progenitors, those from the dorsal CGE (dCGE) are post-mitotic. These post-mitotic cells which co-express SP8, potentially originate from the vCGE and undergo cortical migration (Hansen et al., 2013; Ma et al., 2013). SP8, TF, labels CGE-derived progenitors and interneurons (Ma et al., 2012). SP8 has been reported to be enriched in the SVZ of dCGE (Cai et al., 2013; Ma et al., 2012). The expression of COUPTF2 falls dramatically at the pallium-subpallium border (Alzu'bi et al., 2017a).

1.1.6. LGE and POA

There is a general consensus that in the embryonic rodent nervous system, cINs are mainly derived from the MGE and CGE. However, whether they also originate from the LGE remains a subject of controversy; outcomes from a number of studies have suggested that the LGE is likely to only play a minor role in giving rise to interneurons (Anderson et al., 2001; Wichterle et al., 1999; Wichterle et al., 2001). However, it remains feasible that a tiny proportion of the LGE (mainly dorsal LGE) may generate some cINs in both human and rodent species (Anderson et al., 1997; Wichterle et al., 1999; Wichterle et al., 2001). It is notable that the above mentioned SP8 TF is expressed in the dorsal LGE (dLGE) (in mice) with enrichment in SVZ, where it contributes to arranging olfactory bulb cytoarchitecture, development of striatal neurons, survival and migration of the GABAergic interneurons of the olfactory bulb (Cai et al., 2013; Waclaw et al., 2006). *Gsx2* located upstream of *Sp8* specifies the development of dorsal dLGE progenitors, and it is found enriched in the LGE (Yun et al., 2003).

The foetal POA has been identified as a potential origin of cortical GABAergic interneurons (Gelman et al., 2011). The POA is considered to be part of hypothalamus,

and studies have shown that the POA may generate up to 10% of the GABAergic interneurons in the murine cerebral cortex (Gelman et al., 2011). As in the MGE, *Nkx2.1* expression is present in the embryonic POA; however, *Lhx6* is not expressed in the cells of POA (Flames et al., 2007), nor is *Olig2*, which is strongly expressed in the MGE (Flames et al., 2007; Gelman et al., 2009); however, the POA has been shown to express additional markers not found in the MGE, e.g. in the VZ *Dbx1* and in SVZ *Nkx5.1* and *Nkx6.2.* These findings were based on utilising a CRE line, expressed from the TF *Nkx5.1* (Gelman et al., 2011).

1.1.7. Differences in origin and gene expression of cortical interneurons in rodents, non-human primates and human

The majority of information relating to the development of cINs and differences in their functional roles has been obtained predominantly from mice studies. Although there are many similarities between humans and mice when considering interneuron diversity and cellular organization, notable differences in origin and gene expression have been identified (Boldog et al., 2018).

In rodents, ~60%, ~30% and ~10% of cINs are derived from the MGE, CGE and POA, respectively (Kelsom and Lu, 2013; Kessaris et al., 2014). Similarly, in primates the ventral GE (MGE, CGE and POA) is the main source of cINs. However, the degree to which the cortical VZ and SVZ take part in cortical interneuron generation remains controversial; several primate studies (including human) have indicated that the VZ and SVZ of the dorsal forebrain are essential source of cINs (Al-Jaberi et al., 2015; Jakovcevski et al., 2011; Letinic et al., 2002; Petanjek et al., 2009b). A retroviral labelling technique in organotypic slice cultures has been employed to show the presence of two autonomous lineages of human foetal forebrain cINs (Letinic et al., 2002). In contrast to the previous work detailing that the GE provided the majority of rodent interneurons, 65% of neocortical interneurons (express *Dlx1/2* and *Mash1* TFs) were derived from MASH1 positive precursors of the neocortical VZ/SVZ of the dorsal forebrain of the human foetus. And only 35% of the human interneurons, (identified as DLX1/2^{+} and MASH1-negative) derived from the GE of the ventral forebrain (Letinic et al., 2002). Indeed, MASH1^{+} cells in the VZ/SVZ cortex are indicative of both interneurons and excitatory precursors. It is proposed that these cells are derived from CGE, not from the cortex (Hansen et al., 2013; Ma et al., 2013). However, two studies have examined a small number of human foetuses of less than 24 weeks gestation, and the results have indicated that, as anticipated, the GE is the most significant source, and that the VZ/SVZ of cortex is a minor contributor of cINs during the first two trimesters (Hansen et al., 2013; Ma et al., 2013). Cortical area selection, experimental reagents, protocols and techniques used in

these studies may explain the inconsistencies observed. Thus, it is crucial to determine the population of cIN precursors in both VZ and SVZ compared to the ganglionic-emence in a larger sample with a robust and standardised method (Arshad et al., 2016).

Recently, single nucleus RNA-sequencing analysis has been utilised to comprehend the different cell types within the human cortical middle temporal gyrus. The single nucleus RNA-sequencing data revealed that the human middle temporal gyrus contained equivalent portions of MGE and CGE interneurons, i.e. 44% *LHX6*$^{+}$ and 50% *ADARB2*$^{+}$, respectively (Hodge et al., 2019). This is as opposed to data from earlier work which described mouse cortical interneurons proportions as being ~60% and ~30% from MGE and CGE, respectively (Butt et al., 2005; Kelsom and Lu, 2013; Kessaris et al., 2014; Lee et al., 2010; Miyoshi et al., 2010; Nery et al., 2002). In order to investigate these discrepancies, the ratios of *ADARB2*$^{+}$ and *LHX6*$^{+}$ interneurons were measured in the human middle temporal gyrus and the mouse temporal association area. Results confirmed that equivalent numbers of CGE and MGE interneurons were present in the human, i.e. mean ± standard deviation of 50.2 ± 2.3%, and 44.2 ± 2.4%, respectively, compared with 30.8 ± 1.2% and 67.8 ± 0.9%, respectively, in the mice. In humans, the highest reduction in MGE interneurons were observed in layer 4–layer 6, and the area of layer 4 exhibited the higher relative ratio of CGE interneurons (Hodge et al., 2019).

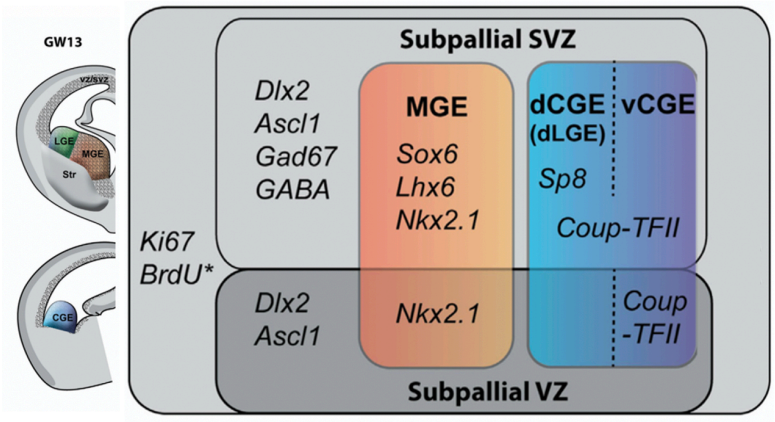


Figure 1.3 A summary of cIN subtypes origin in human. Cortical interneurons are generated from different proliferative zones in the MGE (orange) and the CGE (blue). The CGE ventral and dorsal parts are considered individually. VZ, dark grey; SVZ, medium grey. From (Laclef and Metin, 2018).

It is essential to note that TF expression in cINs and their progenitors in rodents and primates are not exactly same (see Figure1.3 and Figure1.2). Studies designed to further analyse the process relating to human interneurons are ongoing. One research group evaluated NKX2.1, DLX1/2, LHX6 and MASH1 expression in human foetal forebrains during the first 20 weeks of gestation (Zecevic et al., 2011). Expression of these TFs was identified in the VZ and SVZ of MGE and CGE and maintained throughout the 20 weeks of gestation. The results indicate that different progenitor groups for the various cINs are present within several areas,

ventrally, as illustrated in rodents, and dorsally (pallium), in the VZ and SVZ in humans (Zecevic et al., 2011). This implies that the diverse nature of the mature cortical interneurons may arise from their diverse sites of origin (Jakovcevski et al., 2011). Characterizing cIN progenitor populations appears more complex in humans and undoubtedly reflects their higher cognitive function compared with other species (Kelsom and Lu, 2013).

1.1.8. Migration to cortex

In order to migrate to the cortex, interneurons arising from GE areas utilise fundamentally similar cellular processes (Figure 1.4) (Marin and Rubenstein, 2001; Marin et al., 2006). Some neuronal-migration population-dedicated genes have been identified, including, *Robo 1* and *2*, *Cdh8*, *Sema5a*, *Plxnd1*, and *Dab1* (the reelin receptor) (Peyre et al., 2015). On their journey, regardless of their identity, cINs circumvent the striatum and additional areas of the subpallium. With respect to cINs originating from the MGE, the molecular pathways which govern this activity have been delineated. Receptors for chemorepulsive elements of the semaphoring class, e.g. Sema3F and Sema3A, which are plentiful in the striatum, are expressed by cINs migrating to the cortex (Marin and Rubenstein, 2001). Of note, Nrp2 and Nrp1 receptor expression is suppressed by the transcription factor, *Nkx2.1*. Thus, for postmitotic MGE-derived cINs, its downregulation (*Nkx2.1*^{OFF}) dictates migration to the cortex, and its expression (*Nkx2.1*^{ON}) directs cells towards the striatum (Nobrega-Pereira et al., 2008).

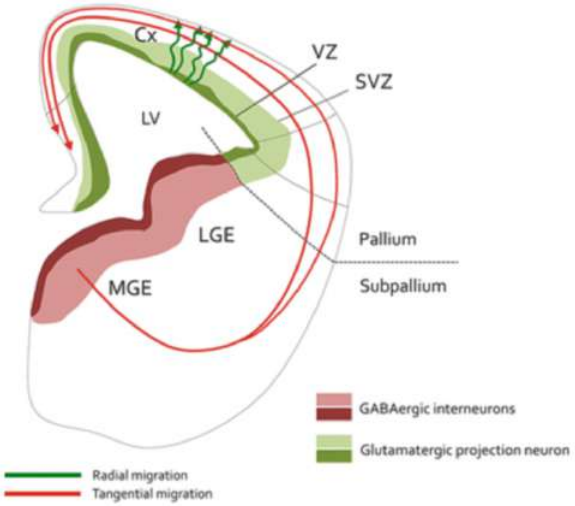


Figure 1.4 Migration of interneurons in the developing cortex. A hemisection showing an E14.5 embryonic mouse forebrain. GABAergic interneurons arise from the subpallial MGE and CGE. GABAergic interneurons are born from stem and precursor cells located in the SVZ (light red) and VZ (dark red) zones. These GABAergic interneurons undergo tangential migration along various paths (red) to integrate the developing wall of the cortex. From (Godin and Nguyen, 2014).

Chemoattractive signals, e.g. Neuregulin1 (NRG1), form a corridor of rising permissiveness in order to encourage interneuronal migration towards the cortex (Flames et al., 2004). The tyrosine kinase receptor, ERBB4, which facilitates NRG1

activity, is the principal neuregulin receptor associated with cINs. The lack of *ErbB4* prevents the arrival of numerous cINs to the cortex (Flames et al., 2004). Within the mature cortex, only certain forms of cINs exhibit ERBB4 expression (Fazzari et al., 2010; Vullhorst et al., 2009). However, a high proportion of perinatal cINs express this receptor (Bean et al., 2014; Lim et al., 2018).

Two sizeable migratory torrents lead to the dissemination of cINs across the growing cortex. One stream passes on the surface via the marginal zone MZ (also called MZ stream); the other takes a deeper course that imbricates the SVZ of the pallium (Lavdas et al., 1999; Wichterle et al., 2001), while a smaller number travel through the subplate. Throughout this developmental stage, cINs uniformly settle within the entire cortex, encompassing the piriform cortex, neocortex and hippocampus (Lim et al., 2018; Marin and Rubenstein, 2001). During this event, it is necessary that tangentially migrating cINs proactively evade passing into the cortical plate, which they ultimately colonise. The obliquely migrating cINs achieve evade breaching of the cortical plate by synchronising with projection neurons which are undergoing radial migration to generate the cortical strata. Evasion of the cortical plate does not require signals to repulse the cINs to be produced by the projection neurons, but rather the process is guided by a permissive passage along the MZ and intermediate zone created via chemokine signalling (Lopez-Bendito et al., 2008; Tiveron et al., 2006). Expression of the chemokine CXCL12, occurring in cells situated along each pathway, contributes to the constraint of cINs within this triad of pathways (Li et al., 2008; Lopez-Bendito et al., 2008; Tiveron et al., 2006). CXCR4 and CXCR7 are two interneuronal receptors for CXCL12, which present cues through varying downstream pathways (Lopez-Bendito et al., 2008; Tiveron et al., 2006; Wang et al., 2011). In pre-mature MGE cells, CXCR7 behaves as a powerful MAP kinase signalling trigger, which is necessary for phosphorylation of ERK1/2. Cortical plate expression of *Cxcr7* arises along an incline in the dorsoventral orientation, which is contrary to the *Cxcl12* gradation alignment in the SVZ. One theory is that cortical plate CXCL12 levels may be diminished by CXCR7, creating a gradient which extends from the MZ and SVZ to the cortical plate. A CXCL12 gradation could be a significant governing factor for cortical penetration (Wang et al., 2011). In fact, disrupting of either *Cxcr7* or *Cxcr4* activity promotes early departure of cINs from their allocated migratory cohorts and disrupts their neocortical stratified and territorial dissemination (Lopez-Bendito et al., 2008; Peyre et al., 2015; Tanaka et al., 2010; Wang et al., 2011).

Human findings are based solely on immunohistology of post-mortem tissue which are bound to include artifacts. However, it is useful to report the drawbacks of human studies. In the human embryo, the initial GABAergic cells are identified on the surface of the

anlage of the neocortex at approximately 6.5-7 weeks of gestation (Meyer et al., 2000; Zecevic and Milosevic, 1997). At this point, the cortical wall is simply composed of proliferative neuroepithelium (VZ) and a top layer of post-mitotic cells (Meyer et al., 2000). During this time, the GABAergic cells are allocated along a lateromedial gradient, the majority of which exhibiting either a bi- or unipolar appearance, typical of cells undergoing migration. There is frequently tangential extension to the cerebral surface with a single elongated process culminating in a growth cone (Rakic and Zecevic, 2003b). As noted in mice studies, it is possible that immature $DLX2^+$ cells migrate to the cortex from the GE in humans (Rakic and Zecevic, 2003b). As growth continues, the population of GABAergic cells increases within the human cerebral wall and becomes distributed in the developing cortical plate. By 11 weeks of gestation, cell density is uniformly observed across all areas (Zecevic and Milosevic, 1997). Mature neuron markers TUJ1 and MAP2 can be identified on the GABAergic cells in the cortical SVZ and VZ and in their expansion of elongated processes that run parallel to the ventricular surface, indicating their post-mitotic and young nature (Laclef and Metin, 2018).

These anatomical findings relating to human GABAergic cell patterns are in keeping with those seen in non-human primates and rodents. Between weeks 14 and 20 of gestation, GABAergic cell density rises in several areas, most notably in the SVZ/intermediate zone (Yan et al., 1992; Zecevic et al., 2011). Similarly, an analogous pattern of growth has been reported for cells that express GAD65 in the cynomolgus monkey (Petanjek et al., 2009a). At E55, a point in their gestation equivalent to week 13 or 14 in humans (Laclef and Metin, 2018), a stream of tangentially oriented GAD65 expressing cells spread across the lower intermediated zone and upper SVZ throughout the developing cortex. At a time equivalent to midterm-gestation in humans, the number of GAD45 expressing cells increase within the VZ and SVZ (Petanjek et al., 2009a; Zecevic et al., 2011). Immunological studies in humans have demonstrated that maximum GABA immunoreactivity shifts towards the surface of the developing cortical plate during the last 20 weeks of gestation, and consequently, in this area, the ratio of multipolar to bipolar GABAergic cells rises (Laclef and Metin, 2018; Yan et al., 1992).

1.1.9. Differences of interneuron migration in rodents and human

When GABAergic cortical interneurons have been created and specified within their various sites of the ventral telencephalon, they then migrate to their cortical target areas (Wonders and Anderson, 2006). In mouse and rat (at E12.5 and E14 respectively), interneurons migrate at the preplate level, correlating with the initial phases of neurogenesis. A further and more sizeable population travel through the intermediate zone (E13-15 mouse and E15-17 rat) (Anderson et al., 2001; Metin et al., 2006).

In humans, the process is slightly different, with the majority of $Sp8^+$ and/or $COUPTF2^+$ neurons arriving at the superficial cortex in humans before birth (Ma et al., 2013). However, a population can still be identified within the SVZ of the cortex, implying that some cells continue to undergo cortical migration during infancy (Paredes et al., 2016). Using post-mortem data, one study showed a large group of immature neurons to be in the process of migration in the infant human frontal lobe at birth (Paredes et al., 2016). This cell population, which expresses DCX and GABA but lacks progenitor and pallial markers, integrates itself within the frontal area during the first few months after birth. The high proportion of subpallial marker expression in these cells indicates that they are derived from the CGE and MGE, and perhaps, also the LGE (Paredes et al., 2016).

In the mouse, *Dlx1/2* and *Mash1* expression play a role in interneuron migration. In KO mouse models for these genes (Long et al., 2009), the progenitors accumulate within the ventral forebrain, and a 75% decrease in GABAergic cells can be observed within the cortex (Anderson et al., 1997). This may be a consequence of migration deficit or cell specification. That is, less are made and migrated, so less found in cortex. The complex details of interneuron migration are poorly delineated; there are numerous factors yet to be identified that play a role in this phenomenon (Faux et al., 2012). However, one key observation is that interneurons do not follow the trajectory of corticofugal fibres when travelling (Nery et al., 2002). Although tangentially migrating cells are not in contact with corticofugal fibres they do touch the cortical VZ once they reach cortex prior to radial migration (Nadarajah et al., 2002). It is shown that the various subgroups of cortical interneurons appear to have heterogeneous target areas, with respect to gross cortical areas as well as to specific cortical layers (Arber and Li, 2013; Nery et al., 2002).

The above research has verified essential variations between the development of humans and rodents. *Nkx2.1* expression is absent in migrating rodent interneurons; while in humans precursor cells, NKX2.1 expression is preserved in the cortex (Rakic and Zecevic, 2003a; Zecevic et al., 2011). Thus, murine data cannot be fully transferred to human development, and any findings arising from studies in rodents must be carefully considered in relation to other species (Arber and Li, 2013; Hodge et al., 2019).

1.2. Cortical interneuron subtype

Within the adult cortex, GABAergic cINs exhibit many overlapping characteristics, which encompass local projections and aspiny dendrites within one specific cortical column, or which extend to adjacent columns. These cells form a highly diverse neuronal group, with the heterogeneous subtypes categorised according to their target specificity,

morphology, place (location), and also by their electrophysical and molecular properties (Markram et al., 2004; Rudy et al., 2011).

In rodent, nearly 60% of neocortical interneurons arise from the MGE. These encompass the parvalbumin (PV^+) fast-spiking and the somatostatin (SST^+) neurons. PV interneurons comprise ~40% of the total number of neocortical interneurons, whereas SST interneurons represent ~30% (Kelsom and Lu, 2013; Miyoshi et al., 2010; Rossignol, 2011; Rudy et al., 2011; Sultan and Shi, 2018; Xu et al., 2008), with the remaining ~30% originating from the CGE. These cortical interneurons are more diverse, but they are unified by expression of the 5HT3A inotropic serotoninergic receptor. These receptors confer extreme sensitivity to the neuro-modulatory actions of serotonin (Férézou et al., 2002; Lee et al., 2010; Miyoshi et al., 2010; Nery et al., 2002; Rudy et al., 2011; Sultan and Shi, 2018). This third group is classified into three populations, reelin-positive multipolar interneurons, which incorporate late-spiking neurogliaform cells; the vasointestinal peptide (VIP^+) cells, which encompass a calretinin (CR^+) cohort; and the VIP^+, CR-negative (CR^-) bipolar interneuron group (Figure 1.5) (Rossignol, 2011).

In terms of their creation date, the birthdates of MGE PV and SST interneurons are somewhat similar and follow the inside-out generation process (Butt et al., 2005; Fairen et al., 1986; Miller, 1985). Within layers 5 and 6, the proportion of PV to SST local circuit neurons is approximately 1.5:1. But, in layers 2 and 3, this increases to 3:1 (Xu et al., 2010a), as a higher percentage of the total SST local circuit neurons are born earlier in the neurogenesis phase than the percentage of the total PV local circuit neurons (Chu and Anderson, 2015; Xu et al., 2010b). However, CGE-derived interneurons are born at later stage of neurogenesis, thus failing to comply with the ‘inside-out’ relationship described above for the MGE-derived local circuit neurons (Figure 1.5) (Butt et al., 2005).

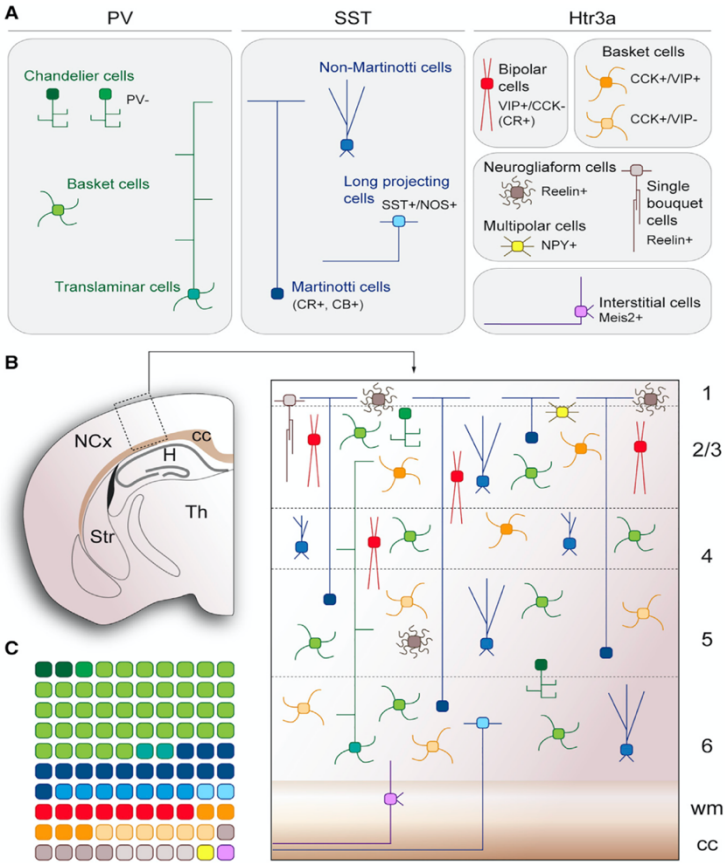


Figure 1.5 Three main types and subtypes of cortical interneurons, their location and relative abundance in the cortex. A) Shows different subtypes of PV, SST and Htr3a interneurons. **B)** Depicts the general laminar distribution of different types of interneurons in the neocortex. **C)** The diagram shows the estimated relative frequency of every interneuron type, color-coded as in (A). It must be noted that there are no direct estimates of the chandelier cells' frequency and many other interneuron classes. Furthermore, the interneurons' relative proportion varies likely across different cortical areas. wm, white matter; cc, corpus callosum. From (Lim et al., 2018).

1.2.1. PV interneuron

PV interneurons comprise of three types, perisomatically targeting basket cells, the smaller group of axon-initial segment-targeting chandelier cells, and trans-laminar interneurons (all classified as PV^+ interneurons). These can be further categorised in relation to their electrophysiological and morphological properties, for example, their somatic diameter, and degree of dendritic and axonal arborisation (Goldberg et al., 2008; Helmstaedter et al., 2009; Markram et al., 2004; Rossignol, 2011; Rudy et al., 2011). PV^+ basket cells have numerous properties which contribute to their recognition as the fastest and most powerful inhibitory interneuron type within the cortex. In the neocortex, PV^+ basket cells are distributed between layers 2 to 6 (Lim et al., 2018). They offer low input resistance, brief action potentials with sizeable after-hyperpolarisation and a low spike modification; they can also maintain elevated firing rate speeds (Butt et al., 2005; Goldberg et al., 2008; Kawaguchi and Kubota, 1997; Markram et al., 2004; Petilla Interneuron Nomenclature Group et al., 2008; Rossignol, 2011). These kinetic properties

are predominantly a consequence of Kv3 voltage-gated potassium channels (Erisir et al., 1999; Goldberg et al., 2008), which facilitate rapid repolarisation and action potential termination. It is also thought that PV^+ basket cells are efficacious in buffering calcium due to elevated expression of calcium ion binding proteins, such as parvalbumin and calbindin. This buffering capacity may render these cells protected against the impact of calcium-provoked excitotoxicity that can be a consequence of rapid firing speeds (Rossignol, 2011).

PV^+ basket cells become configured into complex nests of synaptic contacts on adjacent pyramidal cell soma, thus enabling them to quickly govern the excitation properties of their targeted pyramidal cells. Therefore, these interneurons are appropriately situated to offer robust and rapid feedforward inhibition to their neighbouring pyramidal cells, and thereby efficiently modulate excitation inputs and spike generation in the pyramidal cell populations. This phenomenon sharpens the responsiveness of the cortex and safeguards against runaway excitation after thalamocortical stimulation (Pinto et al., 2000; Pouille et al., 2009; Rossignol, 2011).

There are extensive chemical and electrical synapse connections between PV^+ basket cells, termed gap junctions, giving rise to a vast network of highly coordinated interneuronal activity (Tamas et al., 2000). This network of inhibitory interneurons stimulates and sustains high frequency gamma oscillations within cortical excitatory cell populations (Cardin et al., 2009; Rossignol, 2011; Tamas et al., 2000). Accordingly, if connexin36, a component which constitutes intercellular gap junctions between the PV^+ interneurons, is absent then there is a degree of reduction in task-induced gamma oscillations (Buhl et al., 2003). The gamma oscillations are essential for maintaining attention, working memory and the refinement of executive roles in both rodents and humans (Rossignol, 2011; Sohal et al., 2009; Tallon-Baudry et al., 1998).

In a similar manner to PV^+ basket cells, PV^+ chandelier cells are also characterised by brief non-adapting triggered sequences of action potentials, they can also maintain elevated rates of firing frequencies (Gonzalez-Burgos et al., 2005; Kawaguchi and Kubota, 1996). They are abundant in layer 6 and through layers 1 and 2 (Taniguchi et al., 2013). Morphologically, this cellular cohort is characterised by their vertically arranged cartridges with axonal arbours (DeFelipe et al., 1989; Fairen and Valverde, 1980; Jones, 1975), resembling candlesticks, and they extend to the axon initial segment of excitatory cells to form synapses (Fairen and Valverde, 1980; Somogyi, 1977). The effect of the chandelier cell's synaptic transmission depends on the excitatory cell's resting membrane potential (Szabadics et al., 2006; Woodruff et al., 2011). Therefore

upon activation, that is when the chandelier cells fire/release GABA to pyramidal cells, target pyramidal cellular depolarisation is initiated in the dentate gyrus and cortex (Szabadics et al., 2006; Woodruff et al., 2011). This may be a consequence of increased chloride concentrations and raised GABA reversal potential at the axon initial segment, which arises because of efficacious chloride ion import by the NKCC1 transporter when the KCC2 transporter does not exist (Khirug et al., 2008; Szabadics et al., 2006). In other contexts, hyperpolarising responses are induced by chandelier cells in alternative networks, such as in the hippocampal CA1 zone (Glickfeld et al., 2009). The reason why some cells depolarised and others hyperpolarised, in response to the chandelier cell firing, is a consequence of differences in initial resting membrane potentials of the target cells. Ultimately, the influence of chandelier cells may reflect regional circuitry properties and activities, and as such, the specific ion channel configuration of pyramidal cells within different brain areas is likely to play a significant role. It has been suggested that *in vivo*, chandelier cells are implicated in the initiation of certain oscillatory activities, since they fire just prior to excitatory cells within the hippocampus in sharp-wave-associated ripples (high frequency field oscillations) (Klausberger et al., 2003; Rossignol, 2011). However, much less is known about the chandelier cells and this group's functions have yet to be elucidated (Kelsom and Lu, 2013)

Lastly, compared with the other subgroups, PV-positive translaminar interneurons are few in number, but they are notably populous in layers 5 and 6 (Bortone et al., 2014; Buchanan et al., 2012). Their axons extend across the cortical thickness and so they are able to target pyramidal cells within multiple cortical layers (Bortone et al., 2014; Lim et al., 2018).

1.2.2. SST interneurons

The second subtype is SST^+ interneurons, which encompass Martinotti and non-Martinotti cells, and are immunohistochemically and morphologically diverse, that is they variable co-labelling with calretinin and calbindin and they may be multipolar, bipolar or unipolar, respectively. This subtype exhibits axonal projections that target dendrites of layer 1 pyramidal cell or operate *in situ* within their own cortical layer, and possess different intrinsic electrophysiological characteristics (Halabisky et al., 2006; Ma et al., 2006; McGarry et al., 2010; Miyoshi et al., 2007; Rudy et al., 2011). However, there are a number of common physiological traits in most SST^+ cells, which are also features of the Martinotti cells, such as a low spike threshold, which is notable after-hyperpolarisation and spike rate adaptation. Conversely, they are not similar with respect to their spiking pattern, whether this is regular or in bursts, and this is particularly obvious when triggered from hyperpolarised step currents (Halabisky et al., 2006; Kawaguchi and Kubota, 1996; Ma et al., 2006). In contrast to the fast-spiking basket cells, SST^+

interneurons exhibit greater excitability, with a lower spike threshold and an elevated membrane potential at rest (Fanselow et al., 2008). A subgroup of non-Martinotti cells, however, do not conform to this description, as they display a higher threshold and speed of firing, a diminished spike halfwidth, and lower input resistance (Ma et al., 2006; Miyoshi et al., 2007). This group has mostly been observed within cortical layer 4, and favourably labelled in the X94 (specific GAD67 promoter) GAD67-GFP transgenic lineage (Ma et al., 2006; Rossignol, 2011; Rudy et al., 2011).

The SST^+ Martinotti cells have been detected within layers 2-4 of the cortex, but are most plentiful in layer 5. Their projections extend in a vertical manner in the direction of layer 1. Here, they connect with pyramidal cell dendrites and numerous elongated axonal collaterals reach out to adjoining cortical columns (Kawaguchi and Kubota, 1997; Uematsu et al., 2008; Wang et al., 2004). This cellular cohort modulates the excitability of pyramidal cells by governing the dendritic summation and integration of synaptic inputs. Because their connectivity is jointly convergent, divergent and recurrent, they mediate disynaptic inhibition, which is the effect from one cell to another via two synapses, between connected pyramidal cells, and recurrent feedback inhibition onto presynaptic excitatory cells (Berger et al., 2009; Silberberg and Markram, 2007). SST^+ Martinotti cells are consequently well suited to safeguard against overexcitation and recurrent stimulation in the cortical circuitry (Fanselow and Connors, 2010; Fanselow et al., 2008; Rossignol, 2011).

Approximately 40% of SST cells exhibit intrinsic bursting capacity, and thus it is postulated that they may have a pacemaker function, stimulating certain cortical oscillations (Le Bon-Jego and Yuste, 2007). This may arise because of low threshold, voltage-gated calcium channel expression and sustained sodium ion currents. The SST-positive cells, which are extensively linked through gap junctions, tend to sustain oscillation once triggered (Fanselow et al., 2008; Rossignol, 2011).

As well as in the SST^+ interneurons, SST is expressed in long-range cortical GABAergic projection neurons. The highest density of these cells is observed within the deep cortical layers where these cells project to remote neocortical regions. Their characteristics include adaptive and irregular spiking firing pattern. Neuropeptide Y (NPY), chondrolectin (Chodl) and nitric oxide synthase (NOS) are often expressed in combination from these cells (He et al., 2016), and they have been shown to be effective throughout sleep (Dittrich et al., 2012; Lim et al., 2018).

1.2.3. Other interneurons

An extremely diverse and third category of cINs is recognised by the expression of 5HT3aR, the serotonin receptor (Lee et al., 2010; Rudy et al., 2011). The most frequently found cell forms within this group are interneurons exhibiting vasoactive intestinal peptide (VIP) expression. Their function is mainly disinhibitory, and they generally target PV-positive and SST-positive cortical interneurons (Jiang et al., 2015). Stereotypical VIP-positive cINs have a bipolar morphology, and their axons are vertically aligned. Populations of this cell subtype are particularly dense in layers 2 and 3 of the cortex, wherein continuous adapting firing characteristics can be detected (Pronneke et al., 2015), and calretinin, is found to be co-expressed (Lim et al., 2018).

From a transcription perspective, multipolar VIP^+ cells are highly dissimilar to bipolar interneurons. As basket cells which coexpress neuropeptide cholecystokinin (CCK) they bear more resemblance to other CCK-positive basket cells that are VIP-negative (He et al., 2016; Tasic et al., 2016). Basket cells that express both CCK and VIP have a tiny soma and are enriched within the cortical supragranular layers. In contrast, layers 5 and 6 contain numerous CCK^+, VIP^- basket cells. (He et al., 2016). 5HT3aR (VIP) is expressed by both CCK^+ and CCK^- interneurons (Lim et al., 2018; Rudy et al., 2011). CCK-positive basket cells form synapses on the pyramidal cells' soma and on other cINs; they display either burst or regular spiking patterns of firing (Kawaguchi and Kubota, 1998; Lim et al., 2018). However, the exact role of VIP^+ interneurons within cortical circuitries has yet to be fully elucidated (Rossignol, 2011).

The most plentiful interneurons within layer 1 of the cortex are neurogliaform cells, and their presence within the other cortical layers is scant. They are recognised by a highly dense and typical axonal arbour, and their late spiking firing characteristics. They express 5HT3aR, frequently in combination with NPY and reelin (Lee et al., 2010). Neurogliaform cells are proposed to arbitrate GABA volumetric transmission, that is, nonsynaptic transmission (Lim et al., 2018; Olah et al., 2007), as the neurogliaform cells from both rat and human provoked slow inhibitory postsynaptic potentials and reached GABAA and GABAB receptors on numerous types of interneuron. However, less in known about neurogliaform cells' biophysical properties and further research is required (Overstreet-Wadiche and McBain, 2015).

A further cell type that is found in quantity in layer 1, and which, in transcription terms is almost analogous to neurogliaform cells, is the single bouquet cell (Tasic et al., 2018), which express both 5HT3aR and reelin. In contrast to the neurogliaform cells, their axons spread deep into the cortical layers (Jiang et al., 2015). Two further interneuron subgroups that exhibit 5HT3aR expression include NPY-positive multipolar cells and

Meis2-positive interstitial cINs. The former cohort display irregular spiking patterns and tends to be observed predominantly at the interface between layers 1 and 2 of the cortex (Gelman et al., 2009; Lim et al., 2018; Miyoshi et al., 2010), while Mies2^{+} interneurons are generally found within the white matter and their axons reach the deep cortical strata and the striatum (Frazer et al., 2017; Lim et al., 2018; von Engelhardt et al., 2011).

1.2.4. Interneuron subtypes in human

Many of the previous observations have been formed from rodent models. However, more recent observations in humans have supported the conserved principle of interneuron subtypes in human (Fertuzinhos et al., 2009). Furthermore, a detailed examination of cell types within the human cortical middle temporal gyrus was conducted utilising single nucleus RNA-sequencing, and the researchers found notable variations between the homologous mouse and human cell types, encompassing laminar distributions, morphology, gene expression and clear alterations in proportions; however, it was surprising that both exhibit a well-conserved cellular architecture (Hodge et al., 2019). Given that numerous characteristics were markedly different between the two species, these results highlight the necessity to perform research on human brain directly. In a number of areas their findings were unexpected, such as the absence of the canonical mouse CGE interneuron marker HT3RA expression in human CGE type. The human identified subclass *LAMP5 PAX6* included 6 types enriched mostly in layer1 and layer2. The inhibitory layer 1-layer 6 *LAMP5* were analogous to rosehip cells located within layer 1 in the mouse, but also evident throughout the layers of the cortex. Amongst types of the *LAMP5 PAX6*, *LHX6* was only expressed by inhibitory layer 2-layer 6 *LAMP5*, indicating that, LAMP5 expression is indicative of MGE origin, and this type is similar to *Lamp5*. Furthermore, *SST* was expressed within layer 1 cell types originating from the CGE; this has previously been noted in human, but not in mouse layer 1 interneurons (Hodge et al., 2019).

Also, eleven types of cells were noted within the *SST* subgroup. These demonstrated limited spatial description and encompassed the well-recognised layer 5-layer 6 SST-TH and layer 3-layer 6 (in mouse cortex) SST NPY in layer 5-layer 6 (in human cortex). In-situ-hybridization revealed only modest expression of TH (Tyrosine Hydroxylase) in L5-L6 within human middle temporal gyrus and the mouse temporal association area region. This gene therefore appears to identify highly similar cell types in both mice and humans. In contrast, NPY expression was notably less common in human, thus suggesting that this marker was expressed in a discriminatory manner in the two species (Raghanti et al., 2013; Xu et al., 2010b). Additionally, single nucleus RNA-sequencing (6.1% of GAD1^{+} cells) and cell counts (5.6±0.3% of GAD1^{+} cells) confirmed an elevated ratio of

interneurons that expressed both *ADARB2* and *LHX6* in the human compared with the mouse (1.4±0.2% of $GAD1^+$ cells), an effect specifically noted in layer 6 (Hodge et al., 2019).

Interneurons are a highly heterogeneous population, with numerous differences being ascribed to their individual developmental origins, their diverse response to morphogens and fate-determining gene expression. These results can be utilised by stem cell biologists when designing analogous development models *in vitro* (Arber and Li, 2013). Mouse research has been extensively performed as a model for delineating the processes underlying human cortical neurogenesis, since experiments on primates present multiple challenges. From an evolutionary perspective, however, the mouse is remote from the human, and its physiology frequently dissimilar. Thus, there is some controversy with respect to the pertinence of the use of the mouse for studying developmental processes in the human cortex (Hodge et al., 2019). Consequently, models that are capable of simulating the development of the human cortex are needed.

1.3. Modelling cortical interneuron development *in vitro*

1.3.1. hPSCs biology

Human pluripotent stem cells (hPSCs) comprise human embryonic stem cells (hESCs) and induced pluripotent stem cells (iPSCs). Their properties are such that it is possible for them to produce any cell type and they have a self-renewal capacity when maintained as PSCs. The human embryonic inner cell mass at the blastocyst phase gives rise to hESCs (Thomson et al., 1998), whereas iPSCs are generated *in vitro* utilising forced TF expression, that is, expression of KLF4 (Kruppel Like Factor 4), OCT3/4 (Octamer-Binding TF 3 and 4), SOX2 (SRY-Box TF 2) and C-MYC (MYC Proto-Oncogene, bHLH TF) (Takahashi et al., 2008; Takahashi and Yamanaka, 2006). The label hPSCs applied herein relates to hESCs, and may be recognised through expression of pluripotency-associated TFs in the lines that have been used in the current research, i.e. SOX2, OCT4, and NANOG (Kashyap et al., 2009). Currently, working with hPSCs became physically straightforward and less onerous (Chambers et al., 2009).

1.3.2. Neuronal development modelling

Essentially, there are two methods for the differentiation of stem cells into neurons. The first is the construction of 3-dimensional (3D) entities via organoid generation (Chandrasekaran et al., 2017); the second is based on a 2-dimensional (2D) culture system. The former process is technically challenging in relation to 3D neural induction, that is, it is technically complex to achieve consistent cell type composition and size within individual 3D spheres, and govern the dimensions and morphology of each

sphere. However, the 2D method is a directed differentiation using exogenous factors such as morphogens and small molecules and is a more controllable method (Chandrasekaran et al., 2017). In this thesis, the focus is on the 2D method for cell differentiation via addition of exogenous molecules.

One cornerstone study, which utilised monolayer conditions, reported that endoderm or mesoderm inhibition, will cause neural induction of hPSCs ectoderm, achievable by inhibition of transforming growth factor β (TGFβ) and bone morphogenetic protein (BMP) signalling (Chambers et al., 2009). Thus creating a highly flexible system which could be utilised to investigate numerous conditions for enhancing human cell differentiation. Moreover, the activin/TGFβ/nodal branch of the TGFβ superfamily, an entity which signals by triggering the SMAD2/3 transducer, was proven to be activated during hPSC self-renewal. Consequently, the observed SMAD2/3 activity diminished following initial differentiation. Chambers *et al.* designed a base strategy whereby SB431542, a low molecular weight SMAD inhibitor, was also added to noggin (or LDN193189-BMP inhibitor) (Chambers et al., 2009; Chambers et al., 2012). This successful process became recognised as the dual-SMAD inhibition protocol (Chambers et al., 2009).

Morphogen is responsible for specific neuronal subtype development *in vivo* (see Section 1.1.3), and so the potential of neural stem cells (NSCs), produced by the dual-SMAD inhibition protocol (Chambers et al., 2009), to be influenced by additional patterning factors was investigated. The addition of ascorbic acid, retinoic acid (RA), brain-derived neurotrophic factor (BDNF), and SHH at day 5 yielded motoneurons (Chambers et al., 2009). The neuronal differentiation and maturation are promoted by RA signalling (Maden, 2007). However, the generation of dopaminergic neurons was accomplished following addition of ascorbic acid, SHH, BDNF and FGF8 to neural stem cells on day 9 to 12 (Chambers et al., 2009). These outcomes were useful because it permits the use of the desired mixture of patterning factors and the culture media used are chemically delineated enabling consistent experimental reproduction.

1.3.3. Current protocols

To date, there are a few protocols that enable induction of cIN fate (*in vitro* cultures) (Table 1.1) (Kim et al., 2014; Maroof et al., 2013; Nicholas et al., 2013). These protocols are founded on an admixture of inhibition or stimulation of essential signalling pathways designed to mimic the conditions of cell signalling within the MGE of the ventral telencephalon, where SHH activation (using Wnt inhibitors to initiate a cell fate towards a rostro-ventral state) and Wnt inhibition occurs to enhance ventralising activity (Kim et al., 2014; Maroof et al., 2013).

In neurodevelopment, signalling of SHH exert time-sensitive effects; the concentration of SHH varies over time, which influences the type of cINs created. Whilst low concentrations of SHH stimulate the differentiation of telencephalic progenitors to follow the LGE fate, elevated signals of SHH cause the progenitors to differentiate into the MGE and CGE destiny (ventral telencephalic identity is stimulated by the upregulation of the SHH signalling pathway) (Danjo et al., 2011; Goulburn et al., 2012; Rallu et al., 2002; Xu et al., 2010a). In response to the Shh concentration gradient along the dorsal–ventral axis of the neural tube (during development), different TFs are expressed (Danjo et al., 2011; Goulburn et al., 2012; Rallu et al., 2002; Xu et al., 2010a). In the absence of Shh, Shh receptor Patched 1 (Ptc1) inhibits the activity of the Smoothened Shh-coreceptor (Smo). However, the binding of Shh to Ptch1, removes Smo inhibition. This subsequently stimulates the Gli family of TFs, which are the primary downstream effectors of the Shh signalling pathway (Belgacem et al., 2016). *Gli1* and *Ptch1* are upregulated by purmorphamine (SHH agonist), verifying the agonistic effect upon the Shh pathway exerted by this small molecule. Through the direct binding of purmorphamine to Smo receptor in an agonistic manner, this Shh agonist promotes SHH signalling (Sinha and Chen, 2006; Stanton and Peng, 2010). Therefore, direct Smo activation by purmorphamine (small molecule SHH agonist) will activate SHH signalling (Sinha and Chen, 2006).

Paper	Cell source	Molecules added	Gene expression	Protein expression
(Nicholas et al., 2013)	Induced pluripotent cells, embryonic stem cell	SB431542, BMPRIA, WNT (DKK1), PM, BDNF, DAPT	*ASCL1, DLX1/5/6, LHX6/8, GAD1, SLC32A1, SLC6A1, CB, CR, SST*	OLIG2, ASCL1, GABA, VGAT, CXCR4, NKX2.1, LHX6, CB, CR, NPY, PV
(Kim et al., 2014)	Embryonic stem cell, Induced pluripotent cells	LDN-193189, SB431542, FGF8, SAG, IWP2, GDNF, DAPT	*NKX2.1*	Nkx2.1, GSX2, DLX2, Lhx6, GAD65/67, GABA, CB, PV, SST
(Maroof et al., 2013)	Embryonic stem cells	LDN-193189, SB431542, XAV939, SHH	*OLIG2, ASCL1, NKX6.2, NKX2.1*	CB, GABA, LHX6, DLX2, ASCL1

Table 1.1 Summary of key stem cell-derived cINs assessments. DAPT inhibitor of g-secretase, PM Purmorphamine**,** SB inhibitor of the TGFb1 activin receptor-like kinases, BDNF brain-derived neurotrophic factor, BMPRIA bone morphogenetic protein receptor 1a Fc chimera; DKK1 Dickkopf homolog 1, LDN BMP inhibitor Noggin, SAG Smoothened Agonist, IWP2 a chemical inhibitor of Wnt signaling, SHH sonic hedgehog, GDNF Glial cell line-derived neurotrophic factor, XAV tankyrase inhibitor, FGF8 Fibroblast Growth Factor 8. Adapted from (Fitzgerald et al., 2020).

In 2014, the embryoid body strategy was applied using a differentiation protocol which involved deployment of IWP2 (a Wnt inhibitor) (Kim et al., 2014), which promoted neuroectodermal rostralisation and inhibited neuroectodermal dorsalisation (Gunhaga et

al., 2003), and SAG, an SHH agonist, the dose of which was initially determined via neuroectodermal phenotypic specification (Kim et al., 2014). The extent of SHH activation governs the sub-lineage identity of the LGE as opposed to the MGE in the ventral telencephalon (Figure1.7) (Danjo et al., 2011). The new protocol included the use of the rostralising factor, FGF8 (Kim et al., 2014). The researchers' theory was that addition of extrinsic FGF8 would enhance anterior character, raising NKX2.1 expression by 80%, and diminish COUPTF2 expression, a CGE marker. Within their cell population they identified 0.87%, 1.53% and 22% of PV^+, SST^+ and calbindin cells, respectively after 6 weeks (day 42) of differentiation. However, to date, this is the only group that has eradicated COUPTF2 expression, by ventral forebrain progenitor exposure to FGF8 (Kim et al., 2014).

In the second study, NKX2.1-GFP reporter hESCs were differentiated using SB431542, BMPRIA (bone morphogenetic protein receptor) and DKK1-Wnt inhibitor, together with purmorphamine-induced stimulation of the SHH pathway from day zero (Figure1.6) (Nicholas et al., 2013). Using this resulted in the production of 74.9% of NKX2.1-GFP^+ cells, the majority of which also expressed FOXG1 (81%). GFP positive cells were sorted by day 35 and cocultured onto murine glial cells. Following 30 weeks of differentiation, 1.5%, 40% and 77% of PV^+, SST^+ and CR^+ cells, were obtained, respectively (Nicholas et al., 2013). Nicholas *et al.* demonstrated that an elongated timeline in mimicking human development is necessary for hESC differentiation into functional interneurons (Figure1.6).

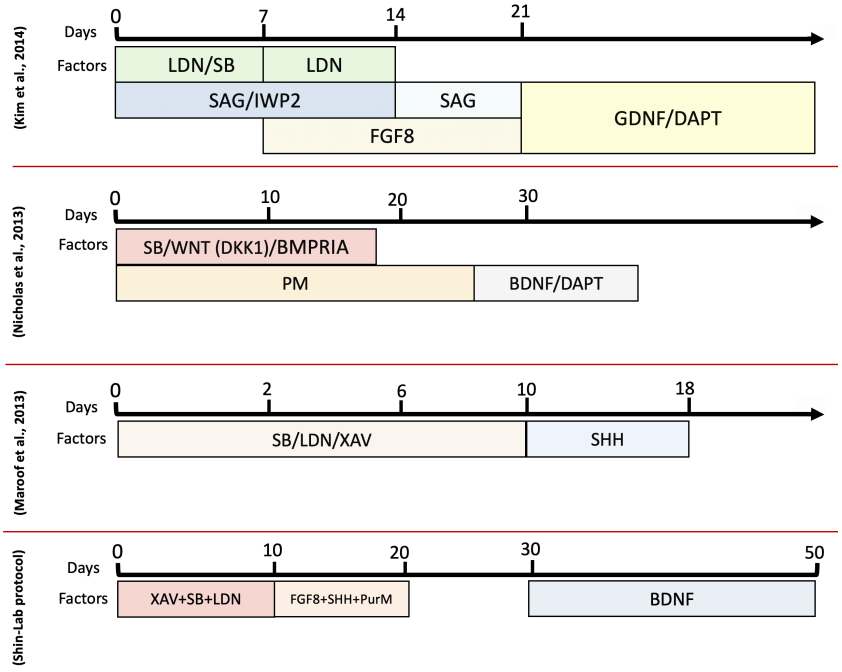


Figure 1.6 Summary of cINs generation protocols. Showing four different protocols and the factors added during differentiation period with days. BDNF brain-derived neurotrophic factor, PM or PurM Purmorphamine**,** DAPT inhibitor of g-secretase, BMPRIA bone morphogenetic protein receptor 1a Fc chimera; DKK1 Dickkopf homolog 1, LDN BMP inhibitor Noggin, SB inhibitor of

the TGFb1 activin receptor-like kinases, SAG Smoothened Agonist, IWP2 a chemical inhibitor of Wnt signaling, GDNF Glial cell line-derived neurotrophic factor, SHH sonic hedgehog, XAV tankyrase inhibitor, FGF8 Fibroblast Growth Factor 8. From (Kim et al., 2014), (Maroof et al., 2013) and (Nicholas et al., 2013).

Lastly, Maroof *et al.* used a monolayer culture, once again deploying NKX2.1-GFP reporter cells (Figure1.6) (Maroof et al., 2013). A dual-SMAD inhibition protocol was utilised, adding SB431542, LDN-193189 and the novel factor XAV939 (a tankyrase inhibitor), a combination labelled XLSB. The XAV939 (Huang et al., 2009) was employed to enhance antero-ventral character during the initial days of differentiation (Maroof et al., 2013). This group demonstrated that, compared to other WNT inhibitors, XAV939 was cheaper (cost effective) and efficient. These authors noted that, for inducing ventralisation, initial SHH activation was not required. Starting the exposure to SHH on days 2-18 produced hypothalamic precursor equivalents. SHH treatment on an intermediate timescale favoured cholinergic (50%) interneuron production as opposed to GABAergic interneurons (15%). However, if SHH and purmorphamine were added following neurulation on day 10 onwards, robust $NKX2.1^{+}$ expression was seen in over 80% of the cells; of these, 90% were observed to co-express FOXG1. Finally, of the GFP^{+} group, 5% and 40% were identified as PV^{+} and SST^{+}, respectively, at day 62 of differentiation (Maroof et al., 2013).

A large fraction of cIN progenitors was induced with all three described protocols, thus in this thesis, through modulating the dorsoventral and rostrocaudal developmental cues a modified version of the protocols of (Kim et al., 2014) and (Maroof et al., 2013) together with the dual SMAD inhibition protocol (adjusted from (Chambers et al., 2009)) were applied to differentiated hPSCs (hESC) into GABA cortical interneurons.

1.3.4. Shin Lab protocol

Our lab has taken a rigorous approach to fine tune the concentration of exogenous factors by testing 12 different combinations of conditions. The aim was to identify the optimal concentrations of extrinsic factors to drive hESCs differentiation towards the desired cell types of the dorsoventral and rostrocaudal position in the forebrain. Three protocols including (Kim et al., 2014), (Maroof et al., 2013) and (Chambers et al., 2009) form the basis of our protocol. This protocol was validated by Dr Karima Azzouni.

These three papers used different concentrations of the small molecules to promote differentiation into neuroectodermal progenitors. Kim *et al.* used media containing LDN193189 and SB431542, and administered IWP-2, together with SAG, from the initial day of differentiation. They detected a small increase in MGE specification at 10 nM of SAG, however at 100 nM, MGE phenotype induction was notably enhanced. By day 7 of

differentiation, PAX6 expression was detected in the majority of cells, and thereafter they compared the efficacy of adjusting Wnt or SHH signalling to determine if greater advantage was gained during or after neuroectodermal development. MGE specification was augmented when factors were administered on day 0 of differentiation, as opposed to from day 7. When only one molecule was administered, only modest elevation in MGE specification was observed; however, concomitant application of both a Wnt blocker and SHH activator produced a mutually potentiating increase in MGE-specification cells. They further investigated whether additional treatment with SHH promoted MGE induction, and their results revealed that NKX2.1 expression showed a modest but significant rise when cells were exposed to SAG during week three of differentiation (Kim et al., 2014). Nicholas *et al*. contributed by discovering that 1-2 μM of purmorphamine, enhance activation of the early SHH pathway, and together with the addition of B-27 to hESCs, produced efficacious and reproducible ventral forebrain-equivalent differentiation. When appraised between days 20 and 30 following differentiation, mean NKX2.1-GP^{+} efficacy was 74.9%.

However, Maroof *et al.* used XAV939, with a specific time window of SHH signalling, that is, from day 10 to day 18 to achieve ventral forebrain characteristics. They also reported that in cells treated with XAV939 from day 0 to day 10, canonical Wnt signalling was inhibited, consequently increasing FOXG1 expression in approximately 90% of LDN-SB-treated cultures. Moreover, the effect of XAV939 on FOXG1 induction was uniform across numerous autonomous hESC and human iPSC lines. Thus, rapid, reliable induction of forebrain fates within hESCs can be induced with the triad of XAV939, SB431542 and LDN (Maroof et al., 2013).

From these above protocols, our lab introduced a novel protocol (validated by Dr Karima Azzoni). First, XAV and SHH was tested at different concentrations. XAV was tested at 0, 1, 2μM concentrations with SB at 10μM and LDN at 100nM from day 0 to day 10 of differentiation. Also, different concentrations of SHH were tested, specifically 0, 50, 100, 200 ng/ml with FGF8 at 100ng/ml and PurM at 1μM from day 10 to day 20 of differentiation, producing a total of 12 different conditions. Thereafter, it was concluded that the optimal condition for production of an efficient quantity of GABAergic cortical interneuron progenitors was as follows: XAV at 1μM and SHH at 200 ng/ml in combination with FGF8 at 100ng/ml and purmorphamine at 1μM. This has produced around 76% of NKX2.1 expressing cells at day 20.

1.4. Schizophrenia

1.4.1. General information of schizophrenia

Schizophrenia (SCZ) affects approximately 1% of the world's population (Escudero and Johnstone, 2014). Typically, SCZ presents in the latter teenage or early adult years. However, there is good evidence that perturbation of neurodevelopment has an effect (Clifton et al., 2019; Walker et al., 2019), but also good evidence for disruption of mature brain function – particularly synaptic signalling – playing pivotal roles in the pathogenesis of the condition (Kirov et al., 2012; Pocklington et al., 2015; Trubetskoy et al., 2022). SCZ arises from a mix of developmental and mature deficits, which can be hard to disentangle as many genes appear to play a role in both (Harrison, 1997, 1999; Sanders et al., 2022). Facts supporting perturbation of development include the link between pregnancy complications (Gupta and Kulhara, 2010) and the risk of SCZ and; the expression of disease-linked genes during foetal development (Clifton et al., 2019). Furthermore, numerous SCZ genome-wide association studies' (GWAS) show that risk loci impact the placenta directly and are predictive indicators of complications during pregnancy (Ursini et al., 2018).

There are three typical groups of symptoms: positive (or psychotic symptoms such as delusions and hallucinations) symptoms, cognitive and negative symptoms (deficits in behaviour and emotional processes) (Patel et al., 2014). The cognitive deficiencies encompass a variety of functional difficulties including working memory, attention, executive abilities (problem solving) and verbal learning (Joyce and Roiser, 2007; Patel et al., 2014). Identification of these cognitive impairments can be utilised as potential markers for SCZ as they often precede diagnostic positive symptoms (Davidson et al., 1999; Hafner et al., 1992; Rund, 1998). SCZ cognitive and negative symptoms are frequently chronic and are associated with lowered life expectancy (Joyce and Roiser, 2007). Furthermore, the symptoms impact on day-to-day activities, reduce an individual's independence, disrupt social interactions, and frequently prevent the patient from undertaking and maintaining paid work. Even when psychotic episodes are absent or in remission, these life-altering symptoms often remain (Solanki et al., 2008).

SCZ is considered a multi-factorial disorder; complicated interactions between biological, social and environmental elements have been postulated as contributing to its onset (Mueser and McGurk, 2004; Patel et al., 2014). Proposed events that increase the risk of SCZ include maternal infection (Khandaker et al., 2013), difficulties or abuse in childhood (Vassos et al., 2012), epilepsy (Clancy et al., 2014) and migration (Cantor-Graae and Selten, 2005). Besides these environmental risk factors, twin studies have

permitted the estimation of SCZ heritability, were it comes out as ~80% (Cardno and Gottesman, 2000) and the SCZ pair-wise concordance rates are ~30-40% in monozygotic twins (Kringlen, 2000). Heritability is the proportion of variability in a trait that is due to genetic variation (Gejman et al., 2010). Therefore, SCZ is highly heritable and genetic factors play a major role in determining an individual's risk of developing the disorder. However, given its complexity, it has proven difficult to implicate specific genes/variants in disease. Some variants, particularly some copy number variants (CNVs), can be reliably implicated, and increasingly individual genes can be too, for example *NRXN1,* from CNVs, also 10 genes identified from The Schizophrenia Exome Sequencing Meta-Analysis (SCHEMA) Consortium rare variant study, specifically *CACNA1G*, *GRIN2A*, *GRIA3*, *TRIO*, *SP4*, *RB1CC1*, and *SETD1A*, *XPO7*, *CUL1*, *HERC1* (Singh et al., 2022), as well as high-priority genes identified, for example, *GRIN2A*, *SP4*, *STAG1*, *FAM120A*, which were identified in the Psychiatric Genomics Consortium3 GWAS study (Trubetskoy et al., 2022). These identified genes are highly expressed in neurons of central nervous system and have various molecular roles including synapse structure, function and formation.

1.4.2. Common and rare genetic risk variants (SNPs, CNVs, SNVs)

A number of methods have been employed to investigate the contribution of genetic variants to the risk of developing SCZ. Common risk variants exist at varied population frequencies, and individually they have a relatively small impact; however, it is the cumulative impact of many such single-nucleotide-polymorphisms (SNPs - defined as a single base alteration in the sequence of DNA) that makes a significant contribution to total illness risk (Pardinas et al., 2018; Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014), with SNPs contributing up to half (50%) of the liability for SCZ (Purcell et al., 2014). Case-control comparison research is now being conducted which enables the study of thousands of samples to investigate SNPs. In 2018, GWAS revealed SNPs associated with SCZ at 145 distinct genetic loci (Pardinas et al., 2018). In addition, copy number variants (CNVs) – duplications or deletions of large genomic regions, the size of which varies from a single kilobase to numerous megabases, have been recognised to bestow an elevated risk of SCZ (Kirov et al., 2012; Kushima et al., 2022), presenting a notably higher odds ratios of 2.7 to >26 than the common variants. However, alone, these do not predict the onset of SCZ, with only 5% to 30% of CNV carriers develop SCZ, an outcome that is dependent on the variant (CNV) (Forsingdal et al., 2019; Mowry and Gratten, 2013). CNVs can arise spontaneously (*de novo*) (Kirov et al., 2012; Warland et al., 2020) as well as hereditarily (Harrison, 2015). CNVs may encompass several genes (Bassett et al., 2010) or impact a sole gene, e.g. *NRXN1* (Rujescu et al., 2009). Two independent *de novo* single gene deletions in *DLG2* have

been identified (Kirov et al., 2012). There are also SNVs (single nucleotide variants) which are consider as rare variants for SCZ and are distributed across multiple genes (Fromer et al., 2014; Purcell et al., 2014), the prevalence of these is <1 in 10000 (Purcell et al., 2014).

One large study that evaluated the available GWAS samples relating to SCZ identified 108 loci. These data suggested involvement of the dopaminergic system, via D2, a dopamine receptor, voltage-dependent calcium channels, synaptic plasticity, and the glutamatergic system (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014). This work has subsequently been extended, and the inclusion of additional samples has led to the identification of 145 loci (Pardinas et al., 2018). A more recent study reported associations in common variants at 270 distinct loci. Analysis of the GWAS data indicated enrichment for SCZ common variants in both excitatory and inhibitory neurons (Trubetskoy et al., 2022). *SP4* and *GRIN2A* were identified as strong candidates within the GWAS GWS (genome-wide significant) loci, and these two genes were also GWS in the most recent rare variant study (Singh et al., 2022).

1.4.3. Biological pathways implicated in schizophrenia

In order to find new effective treatments for SCZ, it is imperative to delineate the altered biological pathways underpinning the disease (Network and Pathway Analysis Subgroup of Psychiatric Genomics, 2015). Genetic pathway analysis, specifically testing sets of genes for enrichment in genetic variants associated with disease, can be performed for all classes of variants, and work in this area on schizophrenia, incorporating genetic and experimental research has revealed a number of pathways that appear to be disturbed.

In keeping with the results of the GWAS and the *de novo* CNV studies (Fromer et al., 2014; Kirov et al., 2012; Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014), mutations influencing constituents of post-synaptic density (PSD) were plentiful, together with genes in the glutamatergic pathway, such as activity-regulated cytoskeleton-associated protein, synaptic organisation, trans-synaptic signalling (Fromer et al., 2014; Pardinas et al., 2018; Trubetskoy et al., 2022) and histone modification (Network and Pathway Analysis Subgroup of Psychiatric Genomics, 2015). These results were obtained from small sample size, however, recently, the Psychiatric Genomics Consortium GWAS, used older studies (Lam et al., 2019; Ripke et al., 2013; Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014) and added extra samples which has led to the largest GWAS performed to date (Trubetskoy et al., 2022). This large psychiatric genomics consortium (PGC) study included 236,642 controls and 69,369 individuals with SCZ, and ranked 130 genes, of which 114 were protein coding genes, based on their expression. These proteins map to 17 unique

hierarchical biological terms, with certain themes emerging. Numerous genes encode receptors, e.g. the metabotropic receptors for GABA and glutamate, and the ligand-gated NMDA receptor subunit (*GRIN2A*), and ion channels, such as the voltage-gated chloride (*CLCN3*) and calcium (*CACNA1C*) channels (Trubetskoy et al., 2022). Other genes are associated with proteins that participate in synaptic governance and differentiation, such as *IL1RAPL1, ADGRB3*, *GPM6A* and *LRRC4B*, the trans-synaptic signalling complex including *SNAP91* and *IL1RAPL1*, and chemical transmission regulation, for example, *DCC, CLCN3* and *MAPK3*. A number of identified genes, specifically *STAG1*, *GRIN2A*, *FAM120A* and *SP4*, where in junction of common and rare variant associations which supports the disorders' pathogenic relevance. The associations noted were large in the genes expressed in excitatory and inhibitory neurons, and the processes implicated were related to synaptic transmission, organisation, and differentiation. However, it is unclear whether these proteins work in concert at certain types of synapse or if the diversity indicates that there are several types in distinct brain areas (Trubetskoy et al., 2022).

In support of the above study, a recent study has conducted whole-exome meta-analyses of 97,322 controls and 24,248 SCZ individuals. They identified 10 individual genes with their annotated function which varied and involved neuronal growth and migration (*TRIO*), nuclear transport (*XPO7*), transcriptional regulation (*SETD1A*, *SP4* and *RB1CC1*), ubiquitin ligation (*CUL1*, *HERC1*) and ion transport (*GRIA3*, *CACNA1G* and *GRIN2A*) (Singh et al., 2022). Also, pathway analysis revealed significant enrichment in 33 gene sets of rare protein-coding variants that recapped biological processes and cellular compartments, involving chromatin modification, postsynaptic density definition, axon guidance, voltage-gated cation channel activity, synaptic transmission and ion transmembrane transport regulation (Singh et al., 2022).

In addition, similar findings to the above genetic studies were originated from work on iPSC. Alterations in activity-induced expression of genes were diminished in neurons originating from an admixture of glutamatergic and GABAergic nerve cells from human schizophrenia iPSCs (n=4) in comparison to those from healthy controls (n=4) (Roussos et al., 2016). Binding of calcium ion, cell adhesion and extracellular matrix organization were the most enriched categories, with enrichment in regulation of transcription *ERBB4* and *NRG* signalling biological processes. Also, one other group studied a combined population of glutamatergic and GABAergic neurons from 8 patients with SCZ and 7 controls. Those from iPSCs derived neural precursor cells (NPCs), collected day 14 of differentiation, in the patient group displayed altered gene expression, particularly in relation to apoptosis, MAPK signalling, survival and cell cycle pathways, when compared to the controls (Lin et al., 2016).

Data from neuronal differentiation of patient/control iPSC lines is an emerging field, and substantially larger sample sizes are required for the production of reliable insight into disease mechanisms. The results so far regarding the identification of affected pathways in SCZ do indicated the convergence of excitatory and inhibitory neuronal signalling and development for the molecular mechanisms that underlie the etiology of SCZ. Thus, determining the influence of SCZ-associated genetic modifications on biological function will provide a comprehensive understanding of the molecular mechanism of SCZ.

1.4.4. Cell types involved in schizophrenia

A range of dysfunctional cellular activities in patients with SCZ have been highlighted in current genetic and experimental research. These studies have also assessed whether the identified genomic loci could be correlated with particular types of brain cells.

Inhibitory and glutamatergic neuronal cell types have been shown to be involved in SCZ aetiology (Polioudakis et al., 2019). One research group investigated the expression of genes in a number of human brain cell types, encompassing oligodendrocytes, foetal astrocytes, neurons, mature astrocytes and macrophage / microglia cells. The cells were isolated and RNA sequencing was used to evaluate gene expression. Cell type-specific expression analysis revealed that anticipated gene expression was notably greater in neurons as opposed to oligodendrocytes and microglia ($p < 0.05$). Within these genes, there was a significantly higher prevalence of synaptic transmission-related genes. These data further reinforce the theory that synaptic transmission is adversely impacted in schizophrenia (Ma et al., 2018).

Another study endeavoured to associate data from the human genome relating to schizophrenia to certain types of cells within the brain, as characterised by their single nucleus RNA-sequencing profiles. They investigated the cell types that exhibit enriched gene expression when harbouring common variants for SCZ. The study provided evidence for SCZ GWAS enrichment in genes with relatively high expression in excitatory and inhibitory (e.g. medium spiny neurons) neurons compared to other cell types. MAGMA analysis for glutamatergic neurons, that is, neurons from the somatosensory cortex and CA1 area of the hippocampus, striatal medium spiny neurons, and cINs demonstrated a greater association with schizophrenia when compared to other neuronal or glial cell types (Skene et al., 2018). Additionally, human gene expression profiling demonstrated that there was some degree of enrichment of oligodendrocyte progenitors; this was not the case in the mouse.

The above cell-type specific data sets (Skene et al., 2018) - plus many more – were tested in a larger, more powerful GWAS. These included fifteen human cell types, derived from the cortical and hippocampal areas, and 24 cell types obtained from the mouse cortex, striatum, midbrain, hypothalamus and hippocampus, as well as oligodendrocytes and cortical PV interneurons. The most robust enrichments included genes expressed by glutamatergic neurons situated in the cortical deep layers, hippocampus and amygdala, although broader enrichments were observed in groups of excitatory and inhibitory neurons (Trubetskoy et al., 2022).

1.4.5. Cortical interneurons in schizophrenia

Data from experimental and clinical studies implies that impaired GABAergic inhibition and the resulting loss of cerebral cortical equilibrium of excitation and inhibition contributes to the pathophysiological mechanisms of many neuropsychiatric diseases (Levitt et al., 2004), including schizophrenia (Guidotti et al., 2005; Lewis et al., 2005). However, not all interneuron subgroups are influenced in the same way by the GABAergic deficiencies (Benes et al., 1991). To date, the use of magnetic resonance spectroscopy has yielded limited and mixed results with respect to the *in vivo* measurement of cortical GABA levels in schizophrenia patients (Ongur et al., 2010; Yoon et al., 2010). In patients with schizophrenia, reduced levels of GABA in the visual area of the cortex have been associated with a decrease in a specific GABA neurotransmission-dependent behavioural parameter of visual inhibition (Yoon et al., 2010). Additionally, GABA levels identified in the frontal lobe have shown a degree of correlation to working memory performance in patients with early-stage schizophrenia (Goto et al., 2009). These two observations add weight to the conclusions that impaired GABA production in schizophrenia underlies the cognitive dysfunction.

Two homologous isoforms of GAD, which vary according to their protein size, specifically, 67 and 65 kDa, are expressed by GABAergic neurons. These are encoded by two distinct genes, *GAD1* and *GAD2* for GAD67 and GAD65, respectively. Within the mouse brain, GAD67 is thought to produce between 80% and 90% of the total GABA (Asada et al., 1997; Condie et al., 1997). Deletion of the mouse gene which encodes for GAD67 results in a 90% decrease in GABA levels and embryos do not survive (Asada et al., 1997). In contrast, deletion of the mouse GAD65 gene is linked with only a 20% decrease of brain GABA levels and the affected mice achieve a normal life expectancy (Asada et al., 1996). According to genetic research, in the schizophrenic prefrontal cortex (PFC), reduced *GAD1* mRNA expression is linked with a SNP in the 5′ untranslated region of *GAD1* (Straub et al., 2007). Several studies have shown that *GAD67* mRNA (Gonzalez-Burgos et al., 2010) and protein levels (Curley et al., 2011; Guidotti et al.,

2000) are reproducibly less in the dorsolateral PFC of patients with SCZ than in controls; all these GABAergic deficiencies are heterogeneously distributed across the subtypes of interneurons (Benes et al., 1991). The most commonly found and most reproducible abnormal finding in schizophrenia is decreased GAD67-mRNA expression within the dorsolateral PFC (Akbarian and Huang, 2006; Gonzalez-Burgos et al., 2010; Nakazawa et al., 2012).

In the brains of patients with SCZ, diminished quantities of interneuron proteins, such as PV, SST, reelin and GABA transporter-I have been described (Guidotti et al., 2000; Nakazawa et al., 2012). A lack of *GAD67* is especially notable in PV^+ interneurons. A quantitative meta-analysis based on PV cell density in the frontal cortex suggested that PV interneurons are reduced in density in schizophrenia individuals (Kaar et al., 2019). Considering another subtype of interneurons, post-mortem analysis has revealed that SST levels are notably diminished across multiple regions of the brain, including the dorsolateral PFC, the hippocampus and orbitofrontal cortex, in patients with SCZ (Alherz et al., 2017). Also, mRNA studies, again performed on post-mortem specimens, indicated a significant reduction in SST in individuals with schizophrenia (Alherz et al., 2017). Furthermore, post-mortem dorsolateral PFC SCZ samples, assessed in a microarray study, identified decreased expression in mRNAs associated with GABA-neurons, encompassing neuropeptide mRNAs for SST, NPY, and CCK (Hashimoto et al., 2008). Thus, in patients with SCZ, it appears that several interneuron cell types exhibit abnormalities when compared with healthy controls.
ErbB4 has been noted as a potential candidate gene that confers vulnerability to schizophrenia (Buonanno, 2010; Rico and Marin, 2011). *ErbB4* codes for a receptor tyrosine kinase, and its expression is favoured by SST^+ and PV^+ cINs (Neddens et al., 2011). The interplay between *Nrg1* and *ErbB4* has been shown to have a major influence on many facets of neuronal maturation, such as axon guidance migration and synapse plasticity and formation (Mei and Xiong, 2008; Rico and Marin, 2011). One study eradicated the expression of *ErbB4* from mouse local circuit neurons originating from the MGE immediately after cell cycle exit, utilising Lhx6-Cre (Del Pino et al., 2013). While the SST expression group was not affected, PV^+ cINs showed a significant decrease of excitatory inputs, and the mice exhibited behavioural abnormalities resulting from phenocopy of some features of cognitive impairment as observed in schizophrenia. Interestingly, interneuron-selective absence of *ErbB4* resulted in a notable decrease in dendritic spines on pyramidal neurons.

Vasoactive intestinal peptide (VIP) expressing interneurons are known to be involved in cortical function governance, and a related study has described decreases in the levels

of SST and VIP mRNAs in the PFC and orbitofrontal cortex in 31 and 35 individuals with bipolar disorder and schizophrenia, respectively, compared to 34 healthy controls (Fung et al., 2014). Elevated mRNA levels for calbindin were also seen in the schizophrenic subjects (Fung et al., 2014).

It is important to note that post-mortem studies have the potential for confounding via reverse causation, as changes found in post-mortem samples may be caused by disease or medication rather than the changes which directly cause the disease. However, the current published data support the association of cell type-specific changes in markers of GABAergic neurotransmission in schizophrenia.

1.5. DLG2

1.5.1. Postsynaptic density

The postsynaptic density (PSD) is a protein-dense specialisation positioned at the postsynaptic membrane of excitatory synapses (Figure 1.7) (Zhu et al., 2016). Electron microscopy has shown that the PSD is a morphologically discrete area. It is characterised by the association of several proteins and a cytoskeleton arranged in a systematic configuration. Vital for synaptic transmission, proteins are thus enabled to be present as part of functional microstructures (Colgan and Yasuda, 2014; Gray, 1959; Palay, 1956). As glutamate is secreted from the presynaptic terminal, receptors, adhesion molecules and postsynaptic scaffolding proteins are located at the PSD in order to facilitate an effective synaptic response (Figure 1.7) (Colgan and Yasuda, 2014). The scaffolding proteins are copious and comprise many protein domains that are specific for protein-protein interactions (PPIs). They operate as a frame that holds the components of the PSD together via binding the postsynaptic receptors, adhesion molecules and signalling proteins from the cytoplasm, for example, protein kinases, phosphatases and GTPases (Kim et al., 1998; Zhu et al., 2016).

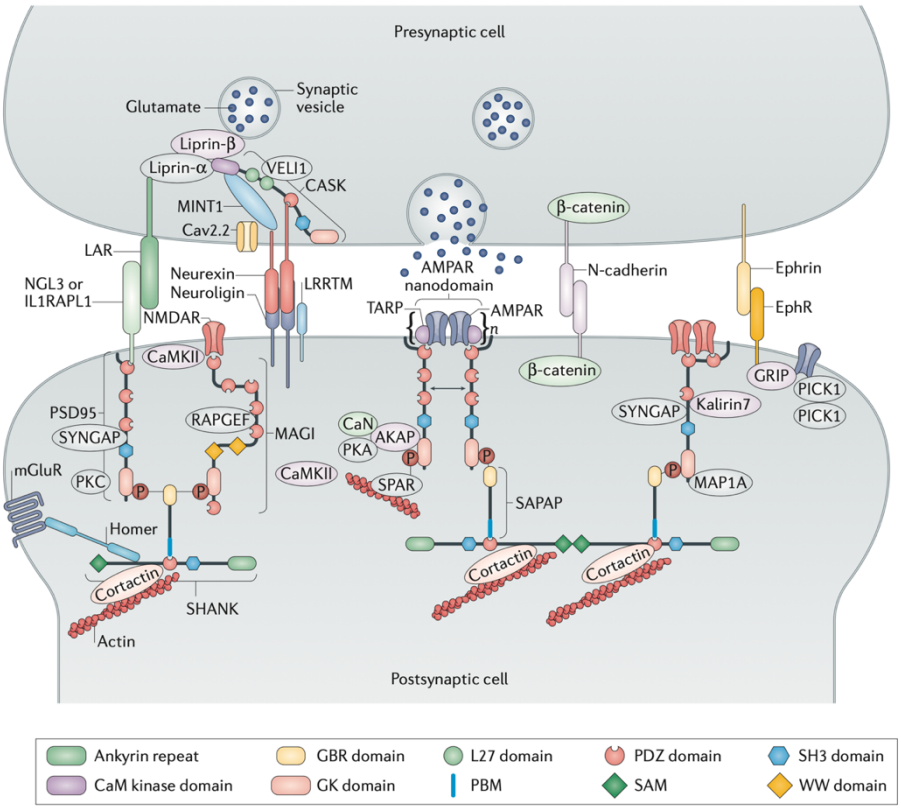


Figure 1.7 The organization in excitatory synapses of scaffold protein-mediated protein complex. Membrane-associated guanylate kinases (MAGUKs) as scaffold proteins are central to the interface of channels/receptors of the synaptic membrane with cytoskeletal proteins and signalling enzymes deep in the postsynaptic density (PSD). The figure shows DLG PSD95-associated protein (SAPAP) with other MAGUKs proteins which contains one SRC homology 3 (SH3), three PDZ domains and one GK domain and serves as a connector between downstream signalling molecules and receptors of cell surface. Key: CaN, calcineurin; AKAP, A-kinase anchor protein; Cav2.2, voltage-gated calcium channel subunit-α Cav2.2; CaMKII, calcium/calmodulin-dependent protein kinase II; EphR, ephrin receptor; GRIP, glutamate receptor-interacting protein; MAP1A, microtubule- associated protein 1A; MAGI, membrane-associated guanylate kinase inverted; mGluR, metabotropic glutamate receptor; TARP, transmembrane AMPAR regulatory protein; SAM, sterile α-motif; PKC, protein kinase C; NGL-3, netrin-G3 ligand; PICK1, protein interacting with C kinase 1; NMDAR, NMDA- type glutamate receptor; VELI1, vertebrate lin-7 homologue 1; PBM, PDZ-binding motif; PDZ, PSD95–DLG1–Zonula occludens 1; LRRTM, leucine-rich repeat transmembrane protein; PKA, protein kinase A. GBR, GK-binding region; L27, LIN2–LIN7; LAR, leukocyte common antigen related; GK, guanylate kinase-like. From (Zhu et al., 2016).

Information with respect to synaptic function and governance has largely been derived from receptor and channel research, and the membrane-associated guanylate kinases, commonly referred to as MAGUK. Their main function is binding and maintaining the integrity of synaptic proteins (Funke et al., 2005). The existence of PDZ, SRC homology 3 (SH3) and guanylate kinase (GK) (catalytically inactive) domains, are characteristic of MAGUKs (Funke et al., 2005; Zhu et al., 2016). This family plays a major part in the PSD, anchoring and maintaining the attachment of glutamatergic receptors (α-amino-3-hydroxy-5-methylisoxazole-4-propionic acid AMPARs and N-methyl-d-aspartate NMDAR) to the cell membrane, as well as coupling their activity to signalling cascades within the cell (Feng and Zhang, 2009). Thus, the functions of the MAGUKs enable them

to govern the strength of synaptic activity (Gardoni et al., 2009). In addition, the accepted role of MAGUKs is that of regulators of adhesion, where they anchor adhesion proteins at tight junctions in order to sustain robust cell to cell interaction (Funke et al., 2005). This function results in MAGUKs being the target of proteases that cleave them former to cell separation (Ivanova et al., 2011). Through interaction with the Scribble cell polarity component, MAGUKs may additionally play important roles in numerous signaling pathways responsible for regulating cell to cell adhesion, cell polarity, migration, differentiation and proliferation (Stephens et al., 2018).

NMDAR are ligand-gated and voltage-dependant ion channels that exhibit calcium permeability. They are thought to compel concomitant glutamate attachment and membrane depolarisation (Dingledine et al., 1999). The exact localization of NMDA-glutamatergic receptors and AMPA receptors at neuronal synapses and dendrites is governed by the discs large (DLG) subfamily proteins. The proteins of DLG regulate extracellular signal transmission to downstream signalling molecules of the PSD. The DLG family of proteins comprises DLG1 (SAP97), DLG2 (PSD93/Chapsyn-110), DLG3 (SAP102), and DLG4 (PSD95/SAP90) (Won et al., 2017; Zhu et al., 2016). They have the capacity to perform postsynaptic membrane glutamate receptor stabilisation together with concomitant binding of multiple signalling molecules, thus earning their place as the superficial layer of scaffolding protein. They also control trafficking, grouping of ion channels and presynaptic terminal excitability, which impacts the quantity of released neurotransmitter. The DLG family members display variable patterns of spatial and temporal distribution (Al-Hallaq et al., 2001; Oliva et al., 2012). As Dlg binds Scribble in invertebrates (Scrib), it is thought that DLG1 binds the scribbled homologue (SCRIB) in humans (Humbert et al., 2008; Stephens et al., 2018).

All four DLG proteins exist with three PDZ domains present in the N-terminus, a SH3 domain and a GK domain in proximity to the C-terminus (Figure 1.8). This set up enables them to engage with their diverse range of binding roles in the PSD, e.g. glutamatergic receptors and cytoplasmic scaffolding proteins (Oliva et al., 2012; Zhu et al., 2016). The role of the PDZ domains is to attach to the C-terminal of their opposite number, including receptors, ion channels and enzymes (Zhu et al., 2016). DLG is present as numerous isoforms which arise from alternative transcription commencement positions and alternative splicing. Alternative splicing is a common mechanism for changing protein function, localization and composition (Funke et al., 2005; Schluter et al., 2006; Zheng et al., 2011). It is possible that the individual isoforms have varied roles, e.g. governing neuronal trafficking, control the strength of AMPA receptor mediated transmission or activity-dependent synaptic transmission (Kruger et al., 2013).

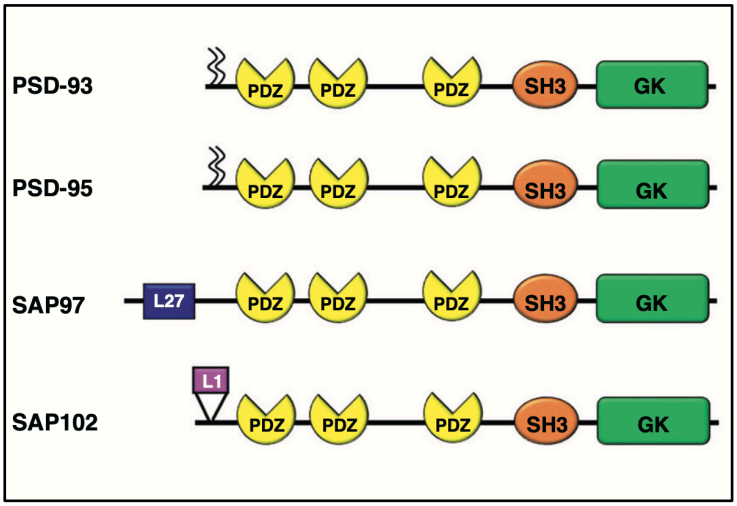


Figure 1.8 Members of MAGUK family. In their N-terminus, PSD-95 and PSD-93 have two palmitoylation motifs. L27: Lin2-Lin7; PDZ: PSD-95/Discs large/Zona occludens-1; SH3: Src-homology-3; GK: Guanylate Kinase. SAP102 has a splice variant exclusive of the L1 region in the N-terminus (Won et al., 2017).

1.5.2. DLG protein family and their role

The gene for *DLG1* has been mapped onto chromosome 3q29. It encodes for synapse-associated protein 97 (SAP97), which is implicated in synapse development (Poglia et al., 2011) and glutamatergic receptor trafficking during neurogenesis (Howard et al., 2010). In *Drosophila*, *dlg* is a lone gene relating to the DLG-MAGUK subfamily and their functions are thought to be inherited into *DLG1*. There is neuromuscular synaptic expression of DLG1 that has been shown to be necessary for normal synaptic structure (Budnik et al., 1996). Work in *Drosophila* has shown that *dlg* is a requisite for synaptic maturation, specifically cell polarity and adhesion (Franco and Carmena, 2019), disorganization of axonal branching in neuronal differentiation, synaptic function and structure (Budnik et al., 1996; Mendoza et al., 2003).

Chromosome 11q14.1 hosts the *DLG2* gene, which encodes a 110 kDa protein (PSD-93, DLG2). A range of neural cell types express DLG2, including cerebellar Purkinje neurones, and the hippocampal pyramidal and granular cells. The protein product of the DLG2 gene has been implicated in synaptic plasticity. It interacts with Fyn, a tyrosine kinase, which contributes to the phosphorylation-based regulation of NMDA receptors that are necessary to induce NMDA-receptor-dependent long-term potentiation (Sato et al., 2008).

Located on chromosome Xq13.1, the *DLG3* gene and its corresponding protein is the initial protein to be associated with intellectual disability (Tarpey et al., 2004). This has been supported by subsequent studies, which have reproduced the link between DLG3 and intellectual impairment (Philips et al., 2014; Zanni et al., 2010), thus implying that DLG3 may influence cognition.

The *DLG4* gene is located on chromosome 17p13.1. It encodes for the protein PSD-95 (DLG4), which contributes to synapse maturation and the NMDA receptor signalling pathway. It is also a key player in the PSD, influencing the grouping and trafficking of NMDA and AMPA receptors (Chen et al., 2000; Oliva et al., 2012; Schnell et al., 2002; Soler et al., 2018). In mice, the deletion of *DLG4* produces molecular and behavioural abnormalities related to autism spectrum disorders (Feyder et al., 2010).

1.5.3. *DLG2*

Expressed DLG2 has the N-terminal with the greatest level of divergence, six variants have been described, which are, one L27-containing domain isoform (β), two palymitoylated isoforms (α) and three further forms (ε, δ, γ) (Kruger et al., 2013; Reggiani et al., 2017). The existence of many isoforms may indicate that each one has an individual function, and this theory was reinforced by the discovery that some of the isoforms appear to be linked with neurodevelopmental disorders (Reggiani et al., 2017).

DLG2 function has been investigated in several studies, including *in vitro* analysis utilising genetic transient knockdown and knockout cell culture models. In cultures of cortical neurones, disruption of *Dlg2* decreased NMDAR-triggered neurotoxicity, and *Dlg2* KO decreased calcium ion influx through the NMDAR (Zhang et al., 2010). It is believed that hypofunction of NMDAR and its interacting PSD components contribute to the SCZ's pathogenesis (Fromer et al., 2014; Kirov et al., 2012; Kristiansen et al., 2006; Purcell et al., 2014). The absence of *Dlg2* (PSD-93) was assessed in *Dlg2* KO cultures, specifically to evaluate whether the localisation or numbers of NMDAR were impacted. The synaptic expression of NR2A and NR2B (NMDA receptor subunits) was reduced which is indicative of altered distribution of neuronal NMDAR (Zhang et al., 2010). However, when the populations of NMDAR and their localisation were appraised in *Dlg2*$^{-/-}$ mice cerebellum, no changes in the actual count or in the synaptosomal fractions were seen (McGee et al., 2001). The perceived difference in the results of the *in vivo* (McGee et al., 2001) and *in vitro* (Zhang et al., 2010) studies are likely due to the different cell types that were investigated. The *in vivo* study looked at the cerebellum, however the *in vitro* looked at cortical neurons (from the striatum and hippocampus). Despite the fact that the specific neurobiological processes following synaptic disruptions remain unclear, it is possible to suspect the involvement of particular biological pathways (Won et al., 2017). One such example is the possibility that DLG2 might influence the tyrosine phosphorylation-based regulation of NMDA receptors, consequently altering their signalling processes (Sato et al., 2008). One related study used multi-electrode array which permits the recording of the spontaneously generated action potentials pattern. They detected disrupted neuronal network activity as a consequence of reduced *Dlg2*

expression following a temporary 7-day small-interfering-RNA knockdown in cultured mouse hippocampus neurons (MacLaren et al., 2011). Elevated bursting rate, synaptic depolarisation duration and synaptic transmission frequency were reported, suggesting that *Dlg2* is involved in regulation of normal neuronal circuitry (MacLaren et al., 2011).

Another study used mouse mutant models for *Dlg2* (homozygous and heterozygous) and *Dlg4* (heterozygous) (Winkler et al., 2018). Hypersocial behaviour was noted in both the $Dlg2^{-/-}$ mice and the $Dlg4^{+/-}$ mutants, but not in the $Dlg2^{+/-}$ mice. This led the authors to postulate that DLG2 is involved in social behaviour, but to a lesser extent than DLG4. In addition, the *Dlg2* mutants demonstrated no anxiety and no memory deficiencies (Winkler et al., 2018). However, the $Dlg2^{-/-}$ mutants showed defective motor learning. While elevated levels of DLG2 protein expression were found in the $Dlg4^{+/-}$, mRNA expression was not raised. It was proposed that DLG2 levels had increased in order to compensate for the absence of DLG4 (Winkler et al., 2018). However, in contrast to the above study, another recent study which used a battery of behavioural tests revealed that DLG2 loss in mice results in a rise in repetitive behaviour and reduced sociability (Yoo et al., 2020). When compared to WT controls, DLG2 deficient mice ($Dlg2^{-/-}$) in this study displayed decreased activity in a novel environment and hypo-sociability (Yoo et al., 2020).

It is important to recognise that in addition to engaging with NMDAR, additional interactors, which have been linked to SCZ, also interact with DLG2, such ErbB4 (Garcia et al., 2000), a gene of which rare variants have been implicated in SCZ risk (Buonanno, 2010; Rico and Marin, 2011). Additional studies are necessary in order to investigate the neuronal DLG2 interactor's function.

1.5.4. DLG2 in schizophrenia

De novo CNV research first identified a potential link between disruption of *DLG2* and an enhanced risk of developing SCZ (Kirov et al., 2012). This study evaluated rare genetic variants in parent-proband trios, identifying CNVs that are present in individuals with SCZ but not in either of their parents. Potentially pathogenic CNVs were recognised in *DLG2* (Kirov et al., 2012). The CNVs detected were specific to *DLG2* (11q14.1) (Bertini et al., 2022), rather than affecting numerous genes, that is, they did not span multiple genes (Kirov et al., 2012). GWAS research has not yet identified *DLG2* as a genetic risk marker for SCZ (Trubetskoy et al., 2022). Moreover, studies utilising exome sequencing techniques have detected rare *de novo* loss-of-function SNVs in *DLG2* in individuals with SCZ (Fromer et al., 2014; Purcell et al., 2014).

In patients with SCZ, post-mortem PFC samples have shown changes in the expression of *DLG2* mRNA and protein, although this has not been investigated in other diseases (Table 1.2). The role of DLG2 in the governance of synaptic plasticity has been suggested from data obtained from animal models, and *Dlg2* mutant mice exhibit cognitive dysfunction and chronic long-term potentiation abnormalities (Table 1.2). Using western blot and *in situ* hybridization, post-mortem studies of brains from schizophrenic individuals noted elevated levels of *DLG2* expression in the anterior cingulate cortex but there were no changes in the expression of *DLG2* within the dorsolateral prefrontal cortex. However, protein expression was reduced in the anterior cingulate cortex, suggesting that dysfunctional translation and/or rapid protein breakdown could be present (Table 1.2) (Kristiansen et al., 2006).

Expression and animal model studies on *DLG2* gene	
Expression studies	**Functional studies**
-Increased *DLG2* mRNA and decreased protein expression in anterior cingulate cortex of patients with schizophrenia (Kristiansen et al., 2006).	-*Dlg2* mutant mice displayed deficits in long-term potentiation (Carlisle et al., 2008). -*Dlg2* mutant mice exhibited cognitive abnormalities (Nithianantharajah et al., 2013).

Table 1.2 Expression and animal model studies on *DLG2* gene.

It has been postulated that the *DLG2* mutations may increase the risk of psychiatric disorders, especially SCZ (Kushima et al., 2022). A lack of *Dlg2* in the hippocampus reduces synaptic long-term potentiation (Carlisle et al., 2008), a process linked to synaptic plasticity which has been found to be disrupted in SCZ (Frantseva et al., 2008). However, the *Dlg4*$^{-/-}$ and *Dlg2*$^{-/-}$ deficiency impacted long-term potentiation differently in the hippocampus. The *Dlg4*$^{-/-}$ mice showed impaired long-term depression and facilitated long-term potentiation. However, the *Dlg2*$^{-/-}$ mice showed no effect on long-term depression but impaired long-term potentiation, indicating that *Dlg4* and *Dlg2* have distinct roles (Carlisle et al., 2008). Nevertheless, both DLG2 and DLG4 are necessary for the maturation of synapses, the formation of NMDA receptor supercomplexes and are essential for development during the postnatal phase (Frank et al., 2016; Oliva et al., 2012).

A cornerstone study, highly pertinent to schizophrenia, reported equivalent impairments in cognitive function in *DLG2* heterozygous humans and mice with *Dlg2* homozygous KO (Nithianantharajah et al., 2013). They evaluated these utilising corresponding tasks on touchscreen testing utilizing the Cambridge Neurophysiological Test Automated Battery. A succession of progressively intricate cognitive activities was assessed in three KO murine models (generated by the breeding of heterozygous mice) - *Dlg2*$^{-/-}$, *Dlg3*$^{-/-}$ and

Dlg4$^{-/-}$ - as well as in heterozygotic *Dlg1*$^{+/-}$ rodents. The latter was heterozygote, since the *Dlg1*$^{-/-}$ homozygote KO was unable to survive (embryonic lethal). Evidence has been produced in support of the embryonic lethal nature of *dlg* homozygous KOs in Drosophila (Woods et al., 1996). This implies that, unlike *Dlg2* and other paralogues, *Dlg1* continues to possess the invertebrate gene functions that are essential for development. During the assessment of cognitive function, the *Dlg1*$^{+/-}$ mice responded parallel to WT mice in all tasks. In contrast to the *Dlg1*$^{+/-}$ mice, straightforward associative learning measures were compromised in the *Dlg4*$^{-/-}$ mice, but normal in the *Dlg2*$^{-/-}$ and *Dlg3*$^{-/-}$ mice. This preliminary finding displays that dissimilar to *Dlg4*, both *Dlg3 a*nd *Dlg2* are not essential for associate learning and simple forms of conditioning. Also, the outcome reveals dissimilarities in the *Dlg* paralogue's function. More complex cognitive requirements, e.g. the paired association object-location task and activities which necessitated cognitive flexibility, reversal and extinction, *Dlg2*$^{-/-}$ mice performed less well than WT rodents. Members of this gene family seem to be important for different cognitive tasks. Of note, is that the specific cognitive impairments seen in *Dlg2*$^{-/-}$ mice were also documented in humans with DLG2 heterozygous mutations. Three of the four patients had been diagnosed as having schizophrenia with all four-having mutations in the *DLG2* coding region. They were similarly assessed for cognitive function utilizing the Cambridge Neurophysiological Test Automated Battery. In keeping with the features of the *Dlg2*$^{-/-}$ mice, the humans carrying DLG2 heterozygous mutations exhibited more errors than controls in visual discrimination acquisition and cognitive flexibility tasks, visuo-spatial memory and learning (paired associates learning). Moreover, in a test for sustained attention, they showed reduced precision in responding compared to controls (Nithianantharajah et al., 2013). The delayed reversal learning seen in *Dlg2*$^{-/-}$ mice and the humans carrying *DLG2* heterozygous mutations is a marker of poor cognitive flexibility and a diminished capacity to respond behaviourally to changes in the environment (Armbruster et al., 2012; Nithianantharajah et al., 2013). Moreover, the cognitive deficits observed in *Dlg2*$^{-/-}$ mice (such as deficits in reversal learning) are similar to those observed in SCZ patients (Leeson et al., 2009). However, humans that carry *DLG2* mutations are heterozygotes, as opposed to the full *Dlg2* KO (homozygous) employed in murine studies (Nithianantharajah et al., 2013). It would therefore be useful to investigate the display of these phenotypes in homozygote humans.

These data suggest that DLG2 gene mutations may increase the risk of developing schizophrenia and cognitive dysfunction as a consequence of encoding for factors that influence synaptic function within the glutamatergic system.

1.5.5. DLG2 in neurodevelopment

Research so far into the role of *DLG2* has focused on its involvement in synaptic transmission in mature, glutamatergic neurons. Data is scarce regarding the temporal expression of *DLG2* in mouse or human from early time points of development. *In situ* hybridisation at the level of mRNA has revealed expression of *Dlg2* at multiple development time points from day 18 embryonically to day 56 postnatally from cortex, striatum and cerebellum. Remarkably, throughout early development expression of *Dlg2* seems to be particularly high, increasing to the end of day 7 postnatally, contrary to adult brains (P56) where *Dlg2* expression is at lower levels (Yoo et al., 2020). Furthermore, RNA sequencing data of stem cell cortical differentiations (excitatory neurons) and human embryos (post-mortem human brains) showed an early RNA expression of *DLG2* gene (Figure 1.9 A and B). The first study investigated human cortical development through stem cell cortical differentiations (*in vitro*) using RNA sequencing at distinct time intervals (http://cortecon.neuralsci.org) (van de Leemput et al., 2014), which identified strong *DLG2* expression at all time periods and throughout early neurogenesis of the *in vitro* cortical differentiation process, which involved day 0 of hESCs (Figure 1.9 B). In the latest timepoint (contains mature neurons in culture), the *DLG2* level is only two to three times that found throughout initial timepoints (day 7) of cortical excitatory differentiation (Figure 1.9 B). The outcome of this study indicates that *DLG2* is express former to neurodevelopment because deep layer neurons will start to generate from day 26 of differentiation (van de Leemput et al., 2014). The second study, briefly, within the neocortex and across brain areas (Figure 1.9 A), the data for *DLG2* at all investigated time points display expression throughout the adult and developing brain. Because of the methodological limitations, at the early development point (period 1), expression at fewer than eight weeks post conception (WPC) is of little importance; however, during the succeeding 8-10 weeks after conception period (period 2), *DLG2* expression is substantial in all examined brain areas (Kang et al., 2011). When considering human brain development, 8-10 WPC corresponds with the neurogenesis stage (see blue shape in Figure 1.9 C) and is precedes the beginning of synapse development (see purple shape in Figure 1.9 C), which occurs about 20 weeks after conception (Figure 1.9 A and C) (Semple et al., 2013). In combination, these data indicate that *DLG2* may have a function during early cortical development other than synaptic transmission in the synapses of mature neurons. Moreover, recently, a study has shown that DLG2 protein is expressed early at day 30 of cortical excitatory neurons differentiated from hESC (Sanders et al., 2022). These findings indicate that *DLG2* is expressed prior to neurodevelopment.

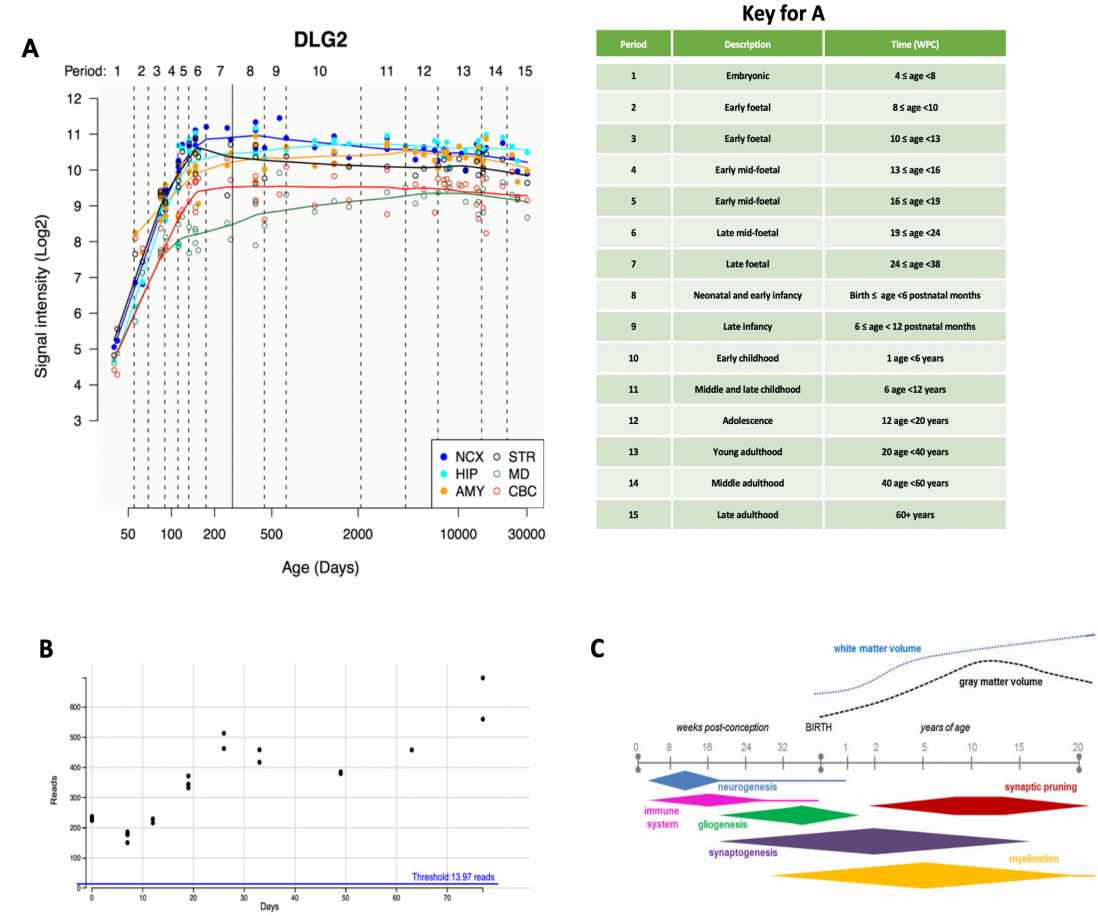


Figure 1.9 The RNA expression of DLG2 gene. A) displays dynamic gene expression of DLG2 gene (*In vitro*) during development and adulthood in the cerebellar cortex (CBC), mediodorsal nucleus of the thalamus (MD), striatum (STR), amygdala (AMY), hippocampus (HIP) and 11 areas of neocortex (NCX), weeks post conception (WPC). Modified from (https://hbatlas.org). **B)** Shows the mRNA DLG2 gene expression in stem cell differentiation (*In vitro*). From (http://cortecon.neuralsci.org)(van de Leemput et al., 2014). **C)** The time course of human brain neurodevelopmental process. From (Semple et al., 2013).

It is known that NMDA-glutamate receptors exhibit calcium permeability and the rise in intracellular calcium operates as a signal that induces numerous modifications in the postsynaptic neuron, comprising changes that result in long or short term alterations in synaptic strength (Hansen et al., 2018). Also considering DLG2 function as a scaffold protein, linking cell surface receptors (e.g. NMDAR) to intracellular signalling signal transduction pathways, DLG2 is required for the proper formation of NMDAR complexes within the synapse (Won et al., 2017; Zhu et al., 2016), DLG2 also control trafficking and grouping of ion channels (Al-Hallaq et al., 2001; Oliva et al., 2012). It is possible that DLG2 may affect the tyrosine phosphorylation-based regulation of NMDA receptors, thus modifying their signalling processes (Sato et al., 2008). Also, it is possible that DLG2 acting directly in the nucleus controlling the transcription together with other proteins. This can be postulated as DLG2 has GK domain and Calcium/calmodulin-dependent serine protein kinase (CASK- other synaptic protein) which has GK domain does act as a transcriptional regulator (Qi et al., 2005; Zheng et al., 2011; Zhu et al., 2016). Therefore, *DLG2* might regulate gene expression by serving as a connection between

cell-surface receptors and signal transduction pathways that regulate the activation of neurogenic transcriptional programs (related to cell-cell communication, neuron/synaptic development, synaptic transmission or duration, adhesion, normal neuronal circuitry regulation or impacts the quantity of released neurotransmitter).

1.6. Aim and objectives

Human *DLG2* mRNA can be identified during all the differentiation phases of hESCs to cortical excitatory neurones (van de Leemput et al., 2014), as well as from eight weeks after conception (Kang et al., 2011). Moreover, DLG2 protein has been identified at day 30 of human cortical excitatory neurons (Sanders et al., 2022). DLG2 is necessary during the synapse formation/development process or maturation growth for NMDA receptor complex formation (Frank et al., 2016). These complexes govern the transcriptional activation schemes that underpin longitudinal alterations in the synaptic function, and corresponding rare mutations are identified in SCZ patients (Fromer et al., 2014; Genovese et al., 2016; Kirov et al., 2012; Kushima et al., 2022; Pocklington et al., 2015; Purcell et al., 2014; Szatkiewicz et al., 2014). Some SCZ patients carry genes in which *de novo* deletions within *DLG2* have occurred (Kirov et al., 2012). Nonetheless, given the connection between the DLG2 gene and perturbed excitatory cortical neurogenesis/SCZ (Kirov et al., 2012; Kushima et al., 2022; Nithianantharajah et al., 2013; Sanders et al., 2022), it has been hypothesized that DLG2-dependent signalling may influence the regulation of early interneuron development. This role is important in SCZ because SCZ have a well-established neurodevelopmental component with risk factors impacting the cortical interneurons and their precursors (Akbarian and Huang, 2006; Alherz et al., 2017; Clifton et al., 2019; Goto et al., 2009; Guidotti et al., 2005; Hashimoto et al., 2008; Kaar et al., 2019; Lewis et al., 2005; Murray and Lewis, 1987; Polioudakis et al., 2019; Trubetskoy et al., 2022; Walker et al., 2019; Yoon et al., 2010). A recent study on *DLG2*$^{-/-}$ revealed a developmental function for *DLG2* in regulating neural stem/progenitor cells gene expression, migration, morphology, and elucidating the activation of transcriptional programs throughout early neocortical development (Sanders et al., 2022). It is thus possible that *DLG2* may be needed during cortical interneuron development for the signalling pathway's formation of initial neurogenesis expression programs that are related to the pathophysiology of SCZ.

This study will be the first to explore DLG2's function in interneurons and their progenitors, since the role of DLG2 in human cortical interneuron during development has, to date, not been reported. Furthermore, earlier mouse studies have not provided data describing the effect of the *Dlg2* KO on MGE and CGE cells (from NPCs and

newborn neurons) or on interneurons in cortex. The hypothesis of this study is that DLG2 may play a role in human cortical interneuron development.

Aim (**Investigating the role of DLG2 gene during the development of cortical interneuron**)

- **First objective:** Using a cortical interneuron differentiation protocol, explore the function of *DLG2* during cortical interneuron generation from hESCs through phenotyping and characterizing the WT and *DLG2* KO neuronal precursors and interneurons.
- **Second objective:** To characterize *Dlg2*$^{-/-}$, *Dlg2*$^{+/-}$ and *Dlg2*$^{+/+}$ embryonic mouse brains and investigate overlapping phenotypes upon DLG2 deficiency in human cells *in vitro* and mouse *in vivo* models.
- **Third objective:** To confirm the role of DLG2 in human cortical interneuron development and to gain an insight into possible mechanisms underlying the *DLG2*$^{-/-}$ phenotypes and their relevance to schizophrenia. To achieve this single-cell RNA seq (scRNAseq) of human cortical neural progenitors and interneurons of *DLG2*$^{-/-}$ and *DLG2*$^{+/+}$ hESCs will be used.

2. Materials and methods

2.1. Cell culture

All cell work was performed under sterile settings using a tissue culture hood (Maxisafe 2020, Thermo Scientific). Cell cultures were maintained under normal condition at 37°C and 5% CO_2 (Galaxy 170 R, New Brunswick).

2.1.1. Maintenance of human embryonic stem cells

Human embryonic stem cell (hESC) lines were used in this study. Two *DLG2*-deficient cell lines (JSD2 (KO1) and JSD21 (KO2)) were generated previously from H7 hESCs (WiCell) by Dr Eunju Shin, using gene-editing technology (CRISPR-Cas9) (Sanders et al., 2022). JSD4 (WT) was used as a control, which went through the gene editing procedure but remained unaltered, thereby serving as a WT sister line. First, cryopreserved hESC lines were thawed as follows. Three universal tubes were prepared with 9 ml of Essential 8™ Medium (Stem Cell Technologies) and a Matrigel® (Corning, BD Pharmagen, Cat No 346231)-coated six-well plate (Thermo Scientific) (1:100 diluted in DMEM/F-12 for 1 hour at 37°C, Gibco) with 2 ml of Essential 8™ (E8) Medium + RevitaCell (100x, Gibco™). The hESC cryovials (three cell lines) were removed from the liquid nitrogen storage tank (-196°C) onto dry ice. The cryovials were quickly thawed in a 37°C water bath by swirling until only a small ice pellet remained. Then, very gently, hESCs (1 ml) were pipetted from the vials into the prepared universal tube containing 9 ml media and centrifuged at 200g for 5 minutes. Next, the media was removed, and the pellet was resuspended in 1 ml of media + RevitaCell (1x) then transferred to the Matrigel-coated wells for a final volume of 2 ml. Then, the plate was shaken to distribute cells and incubated under standard conditions (37°C, 5% CO_2). The medium was replaced every day. After two to three days, cells had become 90% confluent and were ready to be passaged and expanded. For passaging, the E8 medium was aspirated, cells were washed 1x with DPBS (Gibco™, Fisher-Scientific), and 0.5 ml of Gentle Cell Dissociation Reagent Versene Solution (Gibco®, Thermo-Fisher) was added per well of a six-well plate and incubated for 2-3 minutes at 37°C in an incubator. Then, the solution was aspirated and 1 ml of E8 medium was added into each well. The cells were scraped gently with a 5 ml serological pipette (STARLAB) (vertical, horizontal, two diagonals and around the edge of the well). Afterwards, cells were collected from the wells and seeded onto Matrigel®-coated plates containing an appropriate volume of media based on the desired split ratio (either 1:3 or 1:6 –these ratios yield a seeding density of approximately 50,000-150,000 cells/cm^2; the final volume of medium was 2 ml per well of a six-well plate). Lastly, the plate was returned to the incubator and gently shaken from right to left and back to front to obtain even cell distribution. The hESCs were fed daily by aspirating off and replacing 2 ml media. Cells were washed with 1x DPBS if large numbers of dead

cells were observed and passaged when cells were around 80-90% confluent (usually 2-3 days).

2.1.2. Differentiation of hESCs into cortical interneurons

One WT and two KO lines were differentiated into cortical interneurons using the optimised protocol from the lab using existing protocols from Kim *et al.*, Maroof *et al.* and Chambers *et al.* (Chambers et al., 2009; Kim et al., 2014; Maroof et al., 2013). All three cell lines were plated onto these coated plates at a density of around 100,000 cells/well. The cells were maintained in E8 media until almost 90% confluence. Next, E8 was replaced with differentiation medium-N2B27-RA (day 0): 100 ml DMEM-F12, 50 ml Neurobasal, 1 ml N_2 supplement, 1 ml B27 without RA, 2 mM L-glutamine, penicillin-streptomycin, and 0.1 mM β-mercaptoethanol (all from Gibco). This was mixed with supplements SB431542 (SB; 10 μM; Stratech Scientific Ltd), XAV939 (1 μM; Stratech Scientific Ltd), a Wnt pathway inhibitor and LDN193189 (LDN; 100 nM; Tocris). Every other day, half of the media was replaced for the purpose of maintaining the balance of the media. On day 10, cells were pre-incubated with ROCK inhibitor (Y-27632 2HCl, 10 μM; Stratech Scientific Ltd) under standard conditions for one hour and passaged using Gibco® Versene Solution, but were maintained in large clusters. Then cells were transferred onto plates coated with fibronectin (diluted in DPBS 15 μg/ml, 2 hours at 37°C; Millipore) at a ratio of 2:3 and maintained in N2B27-RA supplemented with 100 ng/ml FGF8 (PeproTech), 200 ng/ml SHH (R&D) and 1 μM purmorphamine (Millipore). A second passage was performed at day 19 or 20 onto plates coated with poly-D-lysine and laminin with a ratio of 1:4 following dissociation with StemPro-Accutase®. Cells were kept in Accutase® for 10 minutes and shaken vigorously every 5 minutes to separate clusters into single cells, and then centrifuged with media. To coat the plates, Poly-D-lysine (10 μg/ml in ddH_2O) was added for 2 hours at 37°C, plates were washed three times with DPBS and then Laminin (10 μg/ml; Sigma) was added and incubated for 2 hours at 37°C. Until day 26, cells were maintained in N2B27-RA. Then the media was replaced with N2B27+RA (RA, Gibco). A final passage was performed onto poly-D-lysine-laminin coated plates on day 32, with a ratio of 1:2 following dissociation with StemPro-Accutase®. To support the survival of interneurons until the end of the differentiation on day 50, cells were maintained in N2B27+RA with BDNF (10 ng/ml; Peprotech). Cells were fixed for immunocytochemistry on days 20 and 50.

2.2. Immunocytochemistry

Fixation of cells using 3.7% paraformaldehyde (Sigma) in PBS was performed at room temperature (RT) for around 20 minutes. Then, cells were washed and blocked for 1 hour at RT using 5% donkey serum (Biosera) in phosphate-buffered saline (PBST,

Sigma) with 0.3% Triton-X-100. Next, primary antibodies (see Table 2.1) were mixed with the blocking solution, added to the cells and incubated overnight at 4°C. Then, cells were washed three times with PBS and then incubated in blocking solution with secondary antibodies (1:500, AlexaFluor anti-donkey anti-mouse/rabbit/rat/goat 647, 594, 488; Life Technologies) for 2 hours at RT in the dark. After that, DAPI (1/1000 dilution; Molecular Probes) in PBS was applied for 5 minutes at RT, and cells were washed twice before mounting. Then, using DAKO Fluorescence mounting medium (Agilent, Life Technologies), cells were mounted with glass coverslips (VWR).

The stained cells were visualised and imaged using a Leica DMI6000B inverted fluorescent microscope. Cell counting was performed on images taken from a minimum of three fields per well (randomly picked fields), using CellProfiler (2.2.0) (automated counting) and ImageJ (manual counting).

Antigen	Species	Supplier	Code	Dilution
CALBINDIN	Rabbit	Swant	CB30	1:1000
CALRETININ	Mouse	Swant	7697	1:500
COUPTF2	Mouse	Biotechne	PP-H7147-00	1:100
FOXG1	Rabitt	Abcam	ab18259	1:250
FOXP2	Goat	Abcam	ab1307	1:200
GABA	Rabbit	Sigma	A2052	1:500
GAD67	Mouse	Milipore	mab5406	1:500
GFP	Rabbit	Thermo Fisher	A11122	1:500
MAP2	Mouse	Millipore	AB5622	1:500
NANOG	Rabbit	Cell Signalling	D73G4	1:500
NEUN	Rabbit	Milipore	mab377	1:250
NKX2.1	Rabbit	Abcam	ab40880	1:1000
OCT4	Rabbit	Cell Signalling	C30A3	1:500
OLIG2	Goat	Sigma	p3088	1:200
PAX6	Mouse	DSHB	O42348	1:1000
PV	Mouse	Sigma	p3088	1:1000
SOX2	Mouse	SantaCruze	sc17320	1:500
SOX2	Rabbit	Cell Signalling	D6D9	1:500
SST	Rat	Milipore	mab354	1:50
TUJ1	Mouse	Sigma	T8860	1:500

Table 2.1. List of primary antibodies used for immunocytochemistry staining.

2.3. Migration assay

On day 32 of differentiation, cells from one WT and two KO lines were re-plated into a 24-well plate at low density (100,000 cell/well) using the same protocol as the day 20 passage. Then, on day 44, cells were infected with 1 µl of GFP lentivirus (hPGK-GFP). They were then incubated under standard conditions (37°C, 5% CO_2) for around one week to enable complete transduction of GFP. On day 52, they were washed with sterile DPBS (Gibco™, Fisher-scientific), fed with N2B27+RA and placed in the IncuCyte S3 (Sartorius) Live Cell Analysis System. They were left for 20 minutes to warm up, and then the software was set to capture 36 number of fields in each well of 24 wells at 20X magnification every hour for 24 hours. The videos were exported and analysed using FIJI (ImageJ) and PRISM 8 software. For ImageJ, a plugin called StackReg (Thévenaz, 2011) was used to stack the time-lapse images correctly, so that the field of view would be precisely the same each time. The speed and distance of interneuron migration were calculated. The formula used to measure the speed is $SQRT(X^2+Y^2)/24$. Using Prism, X and Y points of four single neurons over 24 hours (at 1-hour intervals) were plotted in order to show an example movement of a single migrating interneuron from *DLG2*$^{+/+}$ and *DLG2*$^{-/-}$ cultures.

2.4. Morphology Assay

GFP-lentivirus-transduced cells in low-density culture were generated as described above and fixed on day 54 using 3.7% paraformaldehyde (Sigma) at room temperature for around 20 minutes. Then, cells were washed with DPBS and stained with anti-GFP antibody (Table 2.1) using the protocol described in 2.2. Immunocytochemistry. Fields were picked randomly, visualised and imaged using a Leica DMI600B inverted microscope. In total, 42 neurons were analysed from two independent rounds of differentiation for each cell line. Using Neurolucida 360 (MBF Bioscience), cells expressing GFP with a clear neuronal phenotype were quantified. The Neurolucida 360 neuron tracing (branched structure analysis tool) and Neurolucida 360 analysis software packages were used to characterise soma size, the number of primary and secondary branches, and neurite length.

2.5. Adhesion assay

On day 25 of differentiation, triplicates per cell line were used for the adhesion assay. The ECM540 kit from EMD Millipore was used and the manufacturer's procedure was followed. Briefly, 200-µl DPBS was added to the ECM540 Cell Adhesion Array Kit, which is a 96-well plate pre-coated with several extracellular matrix (ECM) substrates (collagen I, collagen II, collagen IV, fibronectin, laminin, tenascin and vitronectin) and BSA as a negative control. The plate was left at RT for around 10 minutes. Then, cells were

prepared for passaging as follows: cells were washed once with DPBS, and then 500 µl of StemPro-Accutase® was added and incubated for 10 minutes. Then, the suspension was collected in a universal tube and centrifuged at 200 g for 5 minutes. Next, 5 ml of N2B27 media was used to resuspend cell pellets. Afterwards, DPBS was removed from the substrate-coated plate and 100,000 cells/well were seeded onto the seven different ECM substrates and incubated for 2 hours (37°C, 5% CO_2). Next, the media was removed, and the wells were washed twice with 200 µl of assay buffer (the assay buffer contains several salts such as magnesium chloride and calcium chloride and is meant to remove any cells that did not bind to the ECM proteins).

Then, 100 µl of the kit stain (the stain solution contains crystal violet, which is a DNA intercalating dye that allows to quantify DNA) was added and incubated for 5 minutes at RT. Afterwards, the stain was aspirated, and washed three times using 200 µl DPBS then left to dry. Lastly, 100 µl of extraction buffer (the extraction buffer contains alcohol and acetate-based compounds to solubilize the bound dye) was added per well and incubated for 5 to 10 minutes. The extraction buffer was added to lyse the cells. This will release the dye from the cells (more cells attached means darker colour) and this is how there is stronger colour hence higher reading with 560nm. So, colour intensity of cell adhesion was measured (absorbance readings) at 560 nm for each ECM substrate, using a colorimetric plate reader (CLARIOstar: BMG Labtech). This experiment was done twice using two independent rounds of differentiation.

2.6. CNV analysis on hESCs

For DNA extraction, QIAamp DNA Mini (QIAGEN) was used, and the manufacturer's instructions were followed. Briefly, two wells of hESCs per cell line (WT, KO1, KO2) were scratched with a 5ml serological pipette in 1 ml of DPBS and collected in microcentrifuge tubes. Then they were centrifuged at 900 g for 5 minutes. Next, supernatant was aspirated before the pellet was resuspended in a mix of 200-µl lysis buffer and 20 µl of QIAGEN Protease for around 15 seconds. Then samples were incubated for 10 minutes at 56°C before adding 200-µl ethanol. The samples were vortexed for 15 seconds and quick spin was applied. Then, the mixture was applied to the QIAamp Mini spin column and centrifuged for 1 minute at 6000 g. Next, 500-µl Buffer AW1 was added to each tube before performing another centrifuging step for 1 minute at 6000 g. Then, the QIAamp Mini spin column was placed in a fresh 2 ml collection tube. Next, 500 µl of Buffer AW2 was added to the samples and centrifuged for 3 minutes at 14,000 g. The QIAamp Mini spin column was placed in a new 1.5 ml microcentrifuge tube. Lastly, 200 µl of Buffer AE was added to the QIAamp Mini spin columns before incubating them at RT (16–25°C) for one minute, and they were then centrifuged for 1 minute at 6000 g. Then, DNA

concentration was measured using NanoDrop™ (Thermo Scientific™) ensuring the concentration was 50-60 ng/µl.

Then, for CNV analysis, Ngoc-Nga Vin (MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University) run the samples on the Illuminia Psych Array (v1.3) after DNA amplification and fragmentation according to the Illuminia assay protocol. The data were analysed using PennCNV with hg19 as a reference assembly.

2.7. Peptide affinity purification

2.7.1. Protein Extraction

Six wells per cell line (WT, KO1 and KO2) at multiple time points (day 20 and day 50), at each time point, four samples per line were taken from two separate differentiations. They were scratched with DPBS (Gibco), collected and centrifuged at 900 g for 5 minutes. Then the cell pellet was homogenised (containing 1 mg of protein) in 1 ml of cold DOC buffer (50 mM Trizma (Tris) pH 9;1% w/v 1 g sodium deoxycholate (Sigma – D6750-10G) and 1X Phosphatase and Protease inhibitor-Sigma MSSAFE-5VL) with instant pipetting. For one hour, the pellet was vortexed every 10 minutes and incubated on ice between vortexing sessions. Next, it was centrifuged for 2 hours at 4°C at 21,300 g. Then the supernatant was collected by decanting, and was stored at –80°C.

2.7.2. Peptide binding to resin (Affi-Gel® 10)

Next, 2 mg of NMDA receptor subunit NR2 C-term peptide “SIESDV” (Pepceuticals – custom peptide synthesis) was completely dissolved in 90% v/v 900-µl methanol (Sigma – 34885) + 100-µl 1M HEPES pH 7 (Sigma - H0887), vortexed and incubated at RT for 30 minutes with agitation. Then, 200-µl Affi-Gel 10 resin (Bio-Rad – 1536099) was pipetted slowly into a fresh Eppendorf tube. Then, 1 ml of methanol was added and spun at 3000 g for 2 minutes, following which the supernatant was removed. This step was repeated twice. Next, the peptide was added to the resin and incubated on a roller mixer at RT overnight. The following day, 200-µl 1 M Tris pH9 was added to block the unreacted nitrogen-hydrogen sulphate groups then incubated at RT on a roller mixer for 2 hours. Next, the samples were centrifuged at 3000 g for 2 minutes and supernatant removed. Then, 1 ml of DOC buffer was added, centrifuged at 3000 g for 2 minutes, and supernatant was removed. This was repeated twice.

2.7.3. Protein-resin binding

The Affi-Gel 10 resin (Bio-Rad – 1536099) was aliquoted into three tubes. The supernatant from the protein extraction was added to 50-µl of washed peptide-bound resin and incubated overnight at 4°C on a roller mixer.

2.7.4. Washing and elution

The protein-resin binding tubes were spun at 3000 g for 2 minutes at 4°C. The supernatant was removed. Next, 1.5 ml of cold DOC buffer was added, then the tubes were incubated on ice for 2 minutes, spun at 3000 g for 2 minutes at 4°C, then the supernatant was removed. This step was repeated four times. Lastly, protein was eluted by adding 50 µl of 5% SDS (Sigma – L3771-25G), heating the tube at 70°C for 15 minutes, and spun at 3000 g for 2 minutes. Supernatant (~50 µl) was collected and stored at –80°C.

Finally, the protein samples were quantified using mass spectrometry; this was undertaken by Mark Collins (Department of Biomedical Science, University of Sheffield) following the protocol described in (Sanders et al., 2022).

2.8. Analysis of Dlg2 transgenic embryonic mouse brain

2.8.1. The Dlg2 transgenic mouse model

The mouse model presented here was generated by the European Conditional Mouse Mutagenesis Program (EUCOMM) as part of a project that aimed to create models with knockout (KO) for all protein-coding genes. The generation of the *Dlg2*tm1a(EUCOMM)Wtsi mice was done by the Wellcome Trust Sanger Institute (WTSI) (transmission of first germline 2014). This was achieved by inserting a cassette in chromosome 7 upstream of critical exon 14 (upstream of SH3 and some part of SH3) on a C57BL/6N-Atm1Brd background (Figures 2.1 and 2.2). A frame shift mutation was the outcome of this knockout. The tm1c way was used to generate the KO (Figure 2.3). Mice of both genotypes (homozygous and heterozygous) are viable.

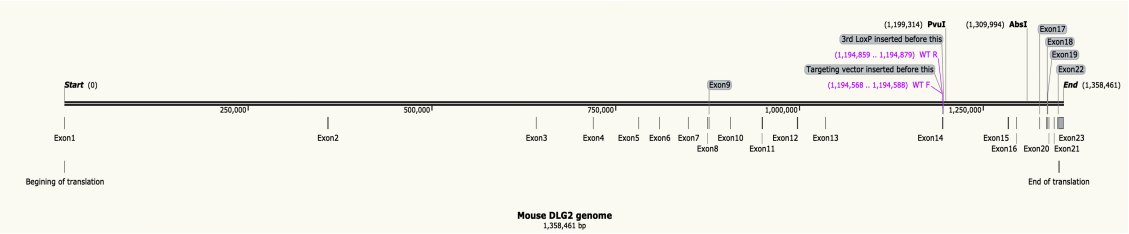


Figure 2.1. Mouse *Dlg2* location of the genetic lesion. Diagram showing the location of the genetic lesion of mouse *Dlg2* genome used in this chapter. Map assembled with downloaded sequences from Ensemble (https://www.snapgene.com). WT R is a wild-type reverse primer and WT F is a wild-type forward primer.

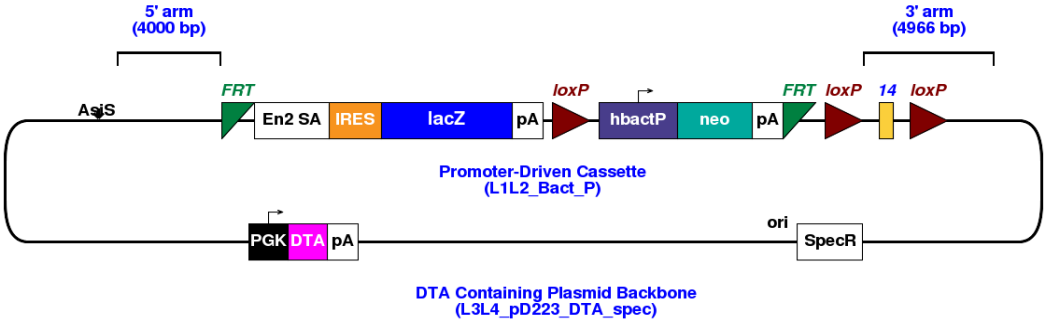


Figure 2.2. An illustration of the vector used to target and generate the *Dlg2*$^{tm1a(EUCOMM)Wtsi}$ knockout mouse (Skarnes et al., 2011). An insertion at the (92285553) position of the L1L2_Bact_P cassette of Chromosome 7 upstream of the critical exon(s) (Build GRCm38). The cassette is comprised of an FRT site followed by a lacZ sequence and a loxP site. This primary loxP site is followed by a gene that is resistant to neomycin under the control of the human beta-actin promoter, SV40 polyA, a second FRT site and a second loxP site. A third loxP site is inserted at position 92286873 downstream of the targeted exon(s). The critical exon(s) is/are therefore flanked by loxP sites. The map of the vector was taken from http://www.mousephenotype.org/data/genes/MGI:1344351.

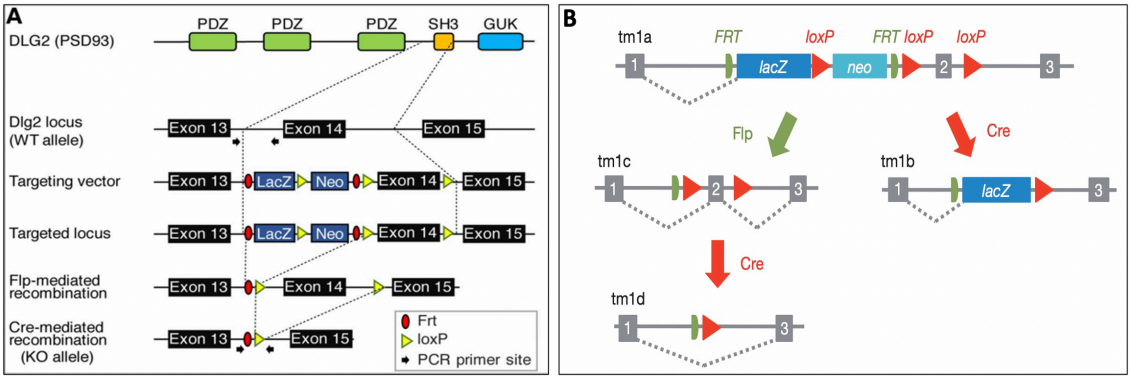


Figure 2.3. Diagram of the conditional allele first-knockout. A) The general strategy for *Dlg2* knockout (Yoo et al., 2020). **B)** The first-knockout allele (tm1a) has a trapping cassette (IRES:lacZ) and a neo cassette (a floxed promoter-driven) which was inserted into the intron of a gene, disrupting the function of the gene. The Flp changes the 'knockout-first' allele to a conditional allele (tm1c), returning the activity of the gene. The promoter-driven selection cassette will be deleted by Cre and tm1a allele floxed exon to produce tm1b (a lacZ-tagged allele) or erase the floxed exon of the tm1c allele to create a frameshift mutation (tm1d), activating nonsense mediated decay of the deleted transcript. From (Skarnes et al., 2011).

2.8.2. Housing and breeding

The mice were bred and culled under approved Home Office project licence 30/3135 and all the procedures were performed according to the University Animal Care Committee's regulations and European Directive 2010/63/EU on the protection of animals used for scientific purposes. The mice were housed in JBIOS in the School of Psychology at Cardiff University.

The breeding was done by Jenny Carter, Dr Rachel Pass and Dr Eunju Shin. One male and two or three female mice were left together between 4-5 p.m. then separated the following morning. The morning of separation was counted as E0.5. On E15.5, the pregnant dam was culled by cervical dislocation. The abdomen was sprayed with 70%

EtOH and opened, and the uterus containing embryos was collected into ice-cold Leibovitz's L-15 medium (Sigma).

2.8.3. Collection and dissection of the mouse embryos

Embryo collection and dissection was performed by Dr Eunju Shin. The uterus was placed in a Petri dish, embryos were retrieved from the sac, and brains were dissected out (n = 12) under a compact stereomicroscope (Olympus SZX7, USA). The dissection was started by separating the embryos from the placenta. Then the size of each embryo was measured by placing the ruler over the embryo and measuring the crown rump length (Table 2.2). Then, an incision was made just above the eyes using scissors, and the skin and skull were peeled off using forceps. Then the brains were freed, and the hindbrain or spinal cord was cut to release the brain. Finally, after isolation, the brains were transferred into 4% PFA overnight at 4°C, and then stored in 25% Sucrose at 4°C.

Embryo ID	05(4) +/+	J4(2) +/+	N3(2) +/+	N3(4) +/+	05(8) +/-	05(9) +/-	N3(1) +/-	N3(3) +/-	DLG13 -/-	K7(1) -/-	K7(3) -/-	K7(2) -/-
CRL (mm)	13.0	11.0	13.5	13.0	13.0	13.0	13.0	12.5	Data not available	12.5	13.0	11.5

Table 2.2. Mouse brains' ID and crown rump length (CRL) measurements.

2.8.4. Embryonic mouse brain genotyping

Genotyping was done by Jenny Carter and Dr Rachel Pass. The embryos' tails were used to perform genotyping. Once the embryo had been collected, the tail was removed and placed into a 1.5 ml tube and then chopped the tail with scissors to make the tissue smaller for easier lysis. Then, using a genomic QIAGEN DNA tissue extraction kit (QIAGEN, Manchester), the genomic DNA was extracted. The manufacturer's instructions were followed to extract genomic DNA. PCR was carried out to identify the genotype.

DNA was extracted from the ear for those used for breeding or the tail in case of embryos. The primers and the PCR cycling protocol for genotyping were provided by the Wellcome Trust Sanger Institute (Table 2.3). A communal *Tm* for the mutant plus WT PCR primers permitted multiplex and simplex reactions. Utilising the kit of MyTaqTM DNA polymerase (Bioline, London, UK) the total master mixes per reaction was 20 µl. Simplex reactions included: 9.2 µl of ddH_2O, 4 µl of 5x reaction buffer, 0.8 µl of reverse and forward primers each, and 0.2 µl of MyTaqTM plus 5 µl of genomic DNA. Multiplex reactions included: 7.6 µl of ddH_2O, 4 µl of 5x buffer, 1.6 µl of forward primer, 0.8 µl of reverse primer 1 and 2, 0.2 µl of MyTaqTM plus 5 µl of genomic DNA. A Biorad Thermal Cycler (T100 BioRadTM, Herts, UK) was used to run the mutant and WT samples at 94°C for 5 minutes, followed by 34 cycles at 94°C for 30 seconds, 58°C for 30 seconds plus 72°C

for 45 seconds, and a last extension for 5 minutes at 72°C; samples were then kept indefinitely at 4°C.

Primer	Reaction	Sequence (5' > 3')	Reaction Temp	Expected band size (bp)
Dlg2_42053_F	Wild type	CCAGAATGTACTTCAGCACCA	58	312
Dlg2_42053_R		TGTGTGTATGTGTGGCTGTTT		
Dlg2_42053_F	Mutant	CCAGAATGTACTTCAGCACCA		222
CAS_R1_Term		TCGTGGTATCGTTATGCGCC		

Table 2.3. PCR primers for *Dlg2*tm1a genotyping. Wellcome Trust Sanger Institute's sequences of PCR primer and temperature reaction for *Dlg2*tm1a(EUCOMM)Wtsi genotyping. F forward, R reverse and CAS_R1_Term targeting cassette. From Rachel Pass's thesis (Pass et al., 2022).

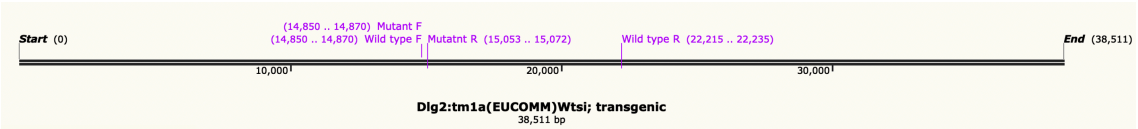


Figure 2.4. Primers binding sites. All primers' binding sites on the vector (*Dlg2*tm1a(EUCOMM)Wtsi). F forward primer and R reverse primer. Assembled on https://www.snapgene.com.

An agarose gel (2%) was prepared with Tris-acetate-EDTA (TAE) (1%) buffer (w/v) and DNA Stain SYBR Safe Gel (1:1000, ThermoFisherScientific, UK). Gel electrophoresis was run for 45 minutes at 95 V (10-µl PCR product loaded/well). The gels were visualised using Omega LumTM G imaging system (Apglegen, San Francisco, USA). The existence of separate WT (312 bp) and mutant (222 bp) bands indicated homozygous mice in simplex reactions (Figure 2.5b). However, in multiplex reactions, WT mice were defined by one band (312 bp), whereas heterozygous mice were defined by two bands (312 bp and 222 bp). Lastly, homozygous mice were defined by a mutant (222 bp) but not a WT (312bp) band. The primers' sequences (Table 2.3) and the PCR protocol for genotyping were provided by the Wellcome Trust Sanger Institute (Pass et al., 2022). The testing of mRNA and protein expression was done by Dr Rachel Pass. The reduction in *Dlg2*$^{+/-}$ protein and mRNA expression in the prefrontal cortex of eight-week-old heterozygote mice was confirmed (Pass et al., 2022). Also, *Dlg2*$^{+/-}$ mice showed no change in the mRNA expression of DLG1, DLG3 and DLG4 across the cerebellum, hippocampus and prefrontal cortex at 8 weeks (Pass et al., 2022).

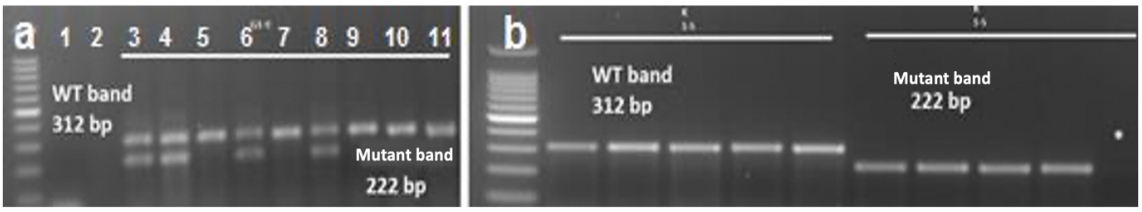


Figure 2.5. Multiplex (a) and simplex (b) genotyping reactions to screen for WT and mutant bands in both WT and Dlg2$^{+/-}$ mice. (a) Lanes 1 and 2 are reaction controls. Lanes 3, 4, 6 and 8 show two bands, indicating heterozygotes. The remaining lanes define WT. From Rachel Pass's thesis (Pass et al., 2022).

Brain ID & Genotype		
05(4) +/+	05(8) +/-	DLG13 -/-
J4(2) +/+	05(9) +/-	K7(1) -/-
N3(2) +/+	N3(1) +/-	K7(3) -/-
N3(4) +/+	N3(3) +/-	K7(2) -/-

Table 2.4. Mouse brain ID and genotype results.

2.8.5. Embedding of the embryonic mouse brains

Firstly, the embryonic mouse brains were removed from the sucrose and rinsed with PBS. Square plastic base moulds (24 x 24 x 5 mm, Fisher Scientific, UK) were prepared by adding a small amount of embedding medium OCT (Agar Scientific Ltd, UK) to cover the base of the mould. Then, the embryonic mouse brain was placed in a linear way using forceps, where the forebrain was at the top and the hindbrain at the bottom of the mould. Next, the mould was placed on dry ice to allow the mouse brain to adhere in that specific orientation. Then, the whole brain was covered with the OCT and left on dry ice until the OCT solidified. Lastly, the embedded moulds were covered with foil and stored at –80°C until sectioning.

2.8.6. Cryostat sectioning of embryonic mouse brain

The Leica Cryostat (CM1860UV) from Leica Biosystems, Germany was used for sectioning. Coronal sections of embedded mouse brains were made at –21°C; sections were 12-µm thick. Ten poly-L-lysine coated slides were labelled from 1A, B to 10A, B. The sections started from olfactory bulb tissue. Sections were collected from the 1st slide to the 10th slide, and the 11th section was attached to the 1st slide, resulting in a 108-µm (12-µm x 9) rostrocaudal distance between the two sections in the slide. The collection pattern was kept the same for all brains, resulting in 16 slices per slide. Afterwards, the slides were maintained at –80°C until the day of staining.

2.8.7. Immunohistochemistry

On day one, the cryostat tissue sections were thawed at room temperature and left to dry. A water bath was pre-heated to 90°C and a Coplin jar filled with citrate buffer (10x antigen retrieval from SIGMA diluted with PBS) place in the bath. The tissue sections were placed in the Coplin jar for 10 minutes. Then, the slides moved to a PBS-filled Coplin jar and then they were dried with a paper towel. Blocking buffer (5% donkey serum (Sigma-Aldrich) and 0.3%PBS+Triton-X-100) was applied (1 ml/slide) for one hour at RT.

Tissue sections were incubated overnight at 4°C with diluted primary antibody (1 ml/slide) (Table 2.5) in blocking solution.

The primary antibody was washed off by rinsing three times in PBS for 5 minutes. Then, the diluted secondary antibody was added (1 ml/slide) and slides were incubated in the dark at RT for 1 hour. Then, slides were rinsed three times in PBS before applying 1 ml of diluted DAPI counterstain (1:1000 in PBS) (Sigma-Aldrich) leaving in the dark at RT for 5 minutes. Then, the DAPI was rinsed off with PBS. Lastly, the slides were mounted using a fluorescent mounting solution (Sigma-Aldrich) and glass coverslips by gently pressing the glass coverslips on top of the slides. Slides were stored at 4°C for one day to allow complete mounting.

Primary antibody	Host Species	Supplier	Code	Dilution	Secondary Ab and company	Dilution
COUPTF2	Rabbit	Millipore	ABE2586	1/1000	Anti-Rabbit488 (Invitrogen-ThermoFisher)	1/500
NKX2.1	Rabbit	Abcam	ab40880	1/1000	Anti-Rabbit488 (Invitrogen-ThermoFisher)	1/500
OLIG2	Goat	R&D	af2418	1/200	Anti-Goat 594 (Invitrogen-ThermoFisher)	1/500

Table 2.5. Antibodies used for staining, including host species and dilution.

2.8.8. Imaging and counting

The stained sections were scanned using a slide scanner machine (Zeiss Axioscan Z1, ZEISS, Germany). The scanning profile was fixed to take images using a 20X objective for the whole brain. For each brain/slide, there was a total of 16 slices. The images were processed using ZEN Imaging Software (ZEISS, Germany). Next, a rostral, a caudal and a medial image that was representative of those sections for all brains were selected. Then, a square region of interest in the brain was selected (MGE, CGE and where NKX2.1 dispersal can be seen) with specified height and width in a specific area across all brain sections.

The regions were picked according to what has been reported in the literature (Cauli et al., 2014; Kessaris et al., 2014) and in the embryonic mouse brain atlas (http://www.epmba.org) and (https://developingmouse.brain-map.org). Selected regions of interest were as follows: for NKX2.1 and OLIG2 (frontorostral (width) 540 pixels x

(height) 750 pixels; intermediate rostral (width) 932 pixels x (height) 1085 pixels; medial (width) 790 pixels x (height) 1325 pixels; caudal (width) 895 pixels x (height) 804 pixels). For COUPTF2 (one medial (width) 601 pixel x (height) 628 pixels, two caudal (width) 985 pixels x (height) 816 pixels) and for NKX2.1 cells where dispersal can be seen in intermediate rostral (width) 448 pixels x (height) 424 pixels). Then, counting was performed manually, using the cell counter plugin on Fiji Image J (NIH, Bethesda, Maryland, USA). Positive cells were counted by merging DAPI and markers of interest.

2.9. Statistical analysis

PRISM 8 GraphPad software (LLC, USA) was used to create graphs and statistically analyse the wet experimental data. All data are shown as mean ± SEM unless specified in figure legend. Overall statistical significance was tested using one-way or two-way ANOVA with Dunnett's test for multiple comparisons; p-values of <0.05 were considered as statistically significant. In figure legends, n represents the number of replicates used in phenotypic experiments such as immunocytochemistry and adhesion. In the context of immunocytochemistry quantification, each replicate is a mean value for a cell culture well taken from microscope fields selected randomly.

Data analysis of single cell RNA sequencing (scRNAseq) was performed as set out below. Where more than one test was performed, p-values were corrected for multiple testing using the Bonferroni correction. The difference was considered statistically significant if the p-value, corrected in case of multiple tests, is less than 0.05.

2.10. Single cell RNA sequencing: data generation

Single cell RNA sequencing was done on nine samples from cell cultures at three time points during cortical interneuron differentiation from hESCs (days 20, 30 and 50) as follows: Sample 1 (Day 20 - $WT^{DLG2\ +/+}$), sample 2 (Day 20 – $KO1^{DLG2\ -/-}$), sample 3 (Day 20 – $KO2^{DLG2\ -/-}$), sample 4 (Day 30 - $WT^{DLG2\ +/+}$), sample 5 (Day 30 – $KO1^{DLG2\ -/-}$), sample 6 (Day 30 – $KO2^{DLG2\ -/-}$), sample 7 (Day 50 - $WT^{DLG2\ +/+}$), sample 8 (Day 50 – $KO1^{DLG2\ -/-}$) and sample 9 (Day 50 – $KO2^{DLG2\ -/-}$).

2.10.1. Single cell preparation for 10x Chromium

To prepare the cell suspension of the nine samples, first the medium (N2B27) was aspirated from one well of a 12-well plate before the well was washed once with DPBS (Gibco). Then, 500 µl of StemPro-Accutase® was added and incubated for around 10-15 minutes. When the cells started to detach, 1 ml of medium (N2B27) was added to the well and cells were dissociated by pipetting; they were then transferred to a 1.5 ml sterile centrifuge tube and centrifuged at 200 x g for 10 minutes. Next, the medium was

aspirated; 2 ml of 0.04% BSA (diluted in N2B27) was added; and cells were resuspended (three to five times) and passed through the FACS tube (Corning). To perform cell counting, 10 µl of cell suspension was taken and mixed well with 40 µl of trypan blue (Sigma). Then, 5 µl of mixture was taken and pipetted gently onto a haemocytometer (Acromec). Manual counting was performed on four squares. The volume needed for 120,000 cells was calculated and the cell suspension was made up to 100-µl with 0.04% BSA in N2B27. An automated cell counter (Countess II) was used to verify the manual cell count and viability. Samples were transported and handled on ice to perform single cell capture, RNA extraction and cDNA and library preparation, which was performed by Angela Marchbank (School of Biosciences, Cardiff University).

2.10.2. GEM Generation, barcoding and cDNA amplification

The Chromium Single Cell 3' Reagent Kits v3 (10x Genomics) was used and the manufacturer's instructions were followed. Briefly, 42.6 µl volume of nuclease-free water plus 4 µl of cell suspension to obtain a final cell count of 3000 cells (1200 cells/µl) was added to the master mix. Next, the master mix containing cells was gently pipetted five times with a pipette tip set to 90-µl. Finally, the master mix containing cells was loaded onto the chip. Also, gel beads and partitioning oil were loaded onto the same chip. The mRNA was captured, and primed RNA was reverse transcribed into complementary DNA (cDNA) before incubation. Silane magnetic beads were used to purify cDNA from the post-GEM reaction mixture. The full-length cDNA was amplified through PCR. Qubit 3.0 was used to analyse the concentration of the cDNA and completed libraries. Tape-station 2200 was used to look at the quality of the cDNA and libraries, and the amplification steps were carried out on a standard thermal cycler. The Chromium Controller (10x Genomics) was used to handle single-cell suspensions. The Chromium Single Cell 3′ Reagent kit version 3 was used. On average, 3000 cells from each 10x reaction were put directly into one of the 10x Genomics chip's inlets. Each step was carried out in accordance with the manufacturer's instructions.

2.10.3. Sequencing

The sequencing was carried out by Joanne Morgan (MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University). HiSeq 4000 (Illumina) was used to sequence barcoded libraries containing 75-bp paired-end reads.

2.11. Single cell RNA sequencing: data analysis

2.11.1. Reads alignment and gene expression quantification

To implement barcode processing, sample demultiplexing and single-cell 3′ gene counting, the Cell Ranger Single-Cell Software Suite (version 3.1.0) was used

(https://software.10xgenomics.com/single-cell-gene-expression). In brief, the *cellranger mkfastq* command was used to demultiplex the Illumina sequencer's base call files and produce FASTQs (for the Read1 and Read2 paired-end reads) enclosing the barcode and read data. By using the FastQC tool (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), these reads were assessed for quality and then processed by the *cellranger count* pipeline using standard parameters that implemented the alignment on the reference transcriptome (GRCh38-3.0.0), filtering and counting of the unique molecular identifier (UMI).

2.11.2. Selecting cells and normalization

The cells were selected by considering the number of genes per barcode, the fraction of counts from mitochondrial genes per barcode and the number of counts per barcode (count depth). Initially, cells that had a mitochondrial read rate < 15% were selected. Next, the low-quality cells were filtered out, keeping only the cells with at least 5% of expressed genes and a number of counts > 5% with respect to the entire cell population. Additionally, potential doublets were filtered out, removing cells with more than 95% of expressed genes and a number of counts > 95% with respect to the entire population. The R package Seurat (v. 3.1.1) (Satija et al., 2015) was used to log-normalize the raw counts, applying the *NormalizeData* function with *scale.factor* parameter set to 10,000. In particular, gene counts for each cell were divided by the total counts for that cell and multiplied by the *scale.factor*. This normalized value was then log transformed using the natural logarithm of value +1.

2.11.3. Dimensionality reduction and clustering

The implemented functions in the Seurat package were used for dimensionality reduction and clustering. The dimensionality reduction was performed using the principal component analysis (PCA), and the statistically significant principal component was determined using the *JackStraw* method with 100 iterations (*JackStraw* and *ScoreJackStraw* functions) and the Elbow approach. Louvain clustering was performed using the *FindClusters* function with the resolution parameter set to 1. The t-distributed stochastic neighbour embedding (t-SNE) technique (Van der Maaten and Hinton, 2008), implemented in the *RunTSNE* function, was used to present data in the form of two-dimensional coordinates (Supplementary File 2 and 3).

2.11.4. Differential expression analysis

Differential expression (DE) analyses among the cell clusters were conducted using the Seurat function *FindAllMarkers*. DE analysis was performed on 1) WT cells only, comparing the cells in each specific cluster versus all other cells (Supplementary File 2

and Supplementary Table 2); 2) on all cells, comparing KO cells versus WT cells in each specific cluster (Supplementary File 3 and Supplementary Table 3); 3) on the identified cell type groups (eight higher groups), comparing the cells in each group versus all other cells (KO cells versus WT) (see Section 5.4). Briefly, a comparison of one cluster (or group) of cells with the rest of the cells was performed using a Wilcoxon Rank Sum test on expressed protein coding genes. After Bonferroni correction, genes with a p-value < 0.05 were considered statistically significant. In differential gene expression analysis, both knockout samples, (KO1 and KO2) were merged. More details about the analysis are presented in Chapter 5.

2.12. Gene set enrichment and gene ontology (GO) analysis

The scRNAseq dataset from Nowakowski *et al.* was used to confirm the precise and accurate identification of cell types. Only inhibitory cell types were selected (MGE RG: MGE-RG1, MGE-RG2; MGE IPC: MGE-IPC1, MGE-IPC2, MGE-IPC3; MGE DIV: MGE-div; MGE Newborn: nIN1, nIN2, nIN3, nIN4, nIN5; MGE mature: IN-CTX- MGE1, IN-CTX-MGE2; CGE mature: IN-CTX- CGE1, IN-CTX-CGE2) (Nowakowski et al., 2017). Using the Seurat function *FindMarkers,* the differential expression analyses between each cell type versus the other cell types were conducted. The number of DE genes obtained for each comparison were: IN-MGE vs others: 465; IN-CGE vs others: 493; MGE-RG vs others: 1347; MGE-Newborn vs others: 752; MGE-IPC vs others: 777; MGE-DIV vs others: 702. The above analysis was performed by Dr Daniel D'Andrea (MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University).

The differential expressed (DE) genes for each Nowakowski cell type were tested for enrichment in any of the DE gene set obtained from the comparison of each WT cluster versus the others. A script provided by Dr Andrew Pocklington (MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University) was used. All genes from hESCs derived cortical interneurons data and in Nowakowski's dataset were used as background set. Fisher's exact test was used to assess the degree of overlap between the two gene sets and a set was considered significantly enriched if the p-value was less than 0.05 after Bonferroni correction.

The same script for the enrichment gene set analysis was performed on the DE genes obtained from the KO and WT comparisons for each cluster and group. The background set contained all coding genes that were expressed in both KO and WT samples. An iterative refinement procedure was employed to determine a semi-independent subset of over-represented annotations from the results of GO term tests. In brief, the gene set with the highest enrichment odds ratio (OR) were chosen; deleted all genes in this set

from all other over-represented annotations; re-tested these reduced gene-sets for over-representation; and removed gene-sets with P ≥ 0.05 (after Bonferroni-correction). This method was repeated until there were no remaining sets with a corrected P < 0.05 (Sanders et al., 2022).

2.13. Genetic association (SCZ) analysis

All common variant gene-set enrichment analyses were carried out using the competitive gene-set enrichment test in MAGMA version 1.08 (de Leeuw et al., 2015), conditioning on all coding genes that were expressed utilizing the *condition-residualize* function. Gene locations were referenced from human genome build 37 (https://ctg.cncr.nl/software/magma). An enrichment test was performed on UP/DOWN regulated genes in the cluster (or group): this was done in order to test whether UP and DOWN regulated genes in a specific cluster (or group) captured more of the SCZ association than expected, after accounting for SCZ association in all expressed genes. The schizophrenia gene-level association statistics were taken from The Schizophrenia Working Group of the Psychiatric Genomics Consortium *et al.* (Pardinas et al., 2018; Trubetskoy et al., 2022).

3. Effect of DLG2 deficiency on human interneuron development using hESCs *in vitro*

3.1. Introduction

DLG2 works as a scaffolding protein within mature neuronal PSD. It plays a part in the functions of the localisation, assembly, and neurotransmitter receptor complexes that are linked with plasticity governance and synaptic transmission (Won et al., 2017). It is only at a somewhat advanced phase of human neurogenesis that synaptogenesis occurs (Silbereis et al., 2016), but significant *in vitro* and *in vivo* expression of *DLG2* mRNA is observed in the first phases of neurogenesis prior to synapse formation, as discussed in Section 1.5.5. Examination of phenotypes generated by the disruption of expression could validate that DLG2 is expressed during early neurogenesis, as well as promote understanding of its role. Consequently, the function of *DLG2* in early human interneuron development can be best explored by using an appropriate model system with gene-level manipulation. Such manipulation can be undertaken in different ways, including expression minimisation through knockdown (Boettcher and McManus, 2015) or expression elimination through knockout (Hsu et al., 2014). The latter was chosen for *DLG2* because it enables identification of the gene's role.

DLG2's function in early human cortical interneurons was analysed via the hESCs model system and *DLG2* editing was performed via CIRSPR/Cas9, as it is cost-effective and displays superior targeting performance (Hsu et al., 2014; Ran et al., 2013). The homozygous knockout lines of *DLG2* were generated because the aim was not to model disease directly, but to understand the function of *DLG2* during cortical interneuron development. Moreover, the differentiation of hESCs into cortical interneurons was undertaken through a developmental patterning method that did not speed up or circumvent the NPC phase. This was justified by the fact that the work was concerned with the function of DLG2 in early neurogenesis instead of in fully developed functional neurons.

No large CNVs were different in KO cell lines compared to WT. However, there remains the possibility that there are changes/other variants or off-target effects which has not been picked up using CNVs analysis. The KO mutations and their locations in the two KO lines are different (Sanders et al., 2022) so flanking changes will not be the same. Also, off-target effects were investigated in Sanders *et al.* with no differences found (Sanders et al., 2022). Therefore, to fully rule out any differences (for maximum certainty), whole-genome sequencing should be conducted to rule out any potential off-target effects in CRISPR-Cas9-derived stem cell lines (WT, KO1 and KO2).

There is a possibility that phenotypes are caused by DLG2 KO, but expression of these phenotypes may depend on genetic background. Introducing KO mutations into multiple hESC lines would allow to explore this possibility. However, it should be noted that in rodent KO studies, it is also standard to modify a gene in a single inbred strain to study the effect of the gene KO. Producing an identical (or at least as identical as possible) genetic background is key, because using different backgrounds makes it difficult to assign genotype-phenotype associations (Sigmund, 2000). And this is achieved by using an inbred line. Typical use of one inbred strain in KO studies is due to time and resources needed to generate one KO line and it applies to the same to the new human pluripotent stem cell line establishment. This is not to say though that the KO effect will be the same for all the genetic backgrounds, it is best to study in different backgrounds to truly define the exact role of the gene in each circumstance, however, it needs to be started from somewhere. When talking about using the patient iPSC lines, it requires more lines as there will be much greater genetic variation between the lines (Germain and Testa, 2017; Volpato and Webber, 2020). How many patient and healthy iPSC lines will be required to truly identify the effect of risk alleles like CNVs and SNPs is not exactly known as it would depend on various things in each experiment (Germain and Testa, 2017; Volpato and Webber, 2020). However, the number of samples used in GWAS and exome sequencing studies guarantee that large numbers of iPSC lines will be required to identify the biological effect of risk alleles. This suggests that most studies using patient iPSC lines are in fact most likely to be underpowered as only handful of iPSC lines are used in most cases. Overall, since 1) the edited hESC line possessed no extra CNVs compared to the wildtype sister line; 2) no alteration has been found in off-target sites; and 3) each KO had different resulting DNA sequence after CRISPR, it seems most likely that the phenotypes observed are due to DLG2 KO, although the additional work mentioned above would help rule out alternatives and consolidate the DLG2 phenotypes that we have identified.

The work discussed in this chapter sought to identify any additional functions of DLG2 besides its known involvement in PSD; this was achieved by analysing the *DLG2*$^{-/-}$ phenotypes in human NPCs and interneurons. To validate two pre-established *DLG2*$^{-/-}$ hESC lines, mass spectrometry was used to assess the complete knockout of DLG2 and its pattern of expression in WT cells in the context of cortical interneuron differentiation. For additional validation, CNV analysis and expression of pluripotency markers were conducted. Neuronal expressing markers and morphology were characterised via immunocytochemistry after the *DLG2*$^{-/-}$ lines and WT controls had differentiated into cortical interneurons. Moreover, a migration analysis of the interneurons was performed to give more insight into the effect of *DLG2*$^{-/-}$ on interneuron migration. Therefore, the

research presented in this chapter covers one aim: phenotyping and characterising the *DLG2* WT and KOs' NPC and interneurons by using an interneuron differentiation protocol.

3.2. Experimental outline

The present work employed two homozygous DLG2 knockout hESC lines created earlier by implementing the CRISPR/Cas9-D10A nickase system (Chiang et al., 2016) for H7 hESC genome editing; the cleavage of the DLG2 exon 22 that encodes a portion of the initial PDZ domain of the protein, was guided by two gRNAs (Sanders et al., 2022). Several methods were employed to validate the KO1 and KO2 of *DLG2*$^{-/-}$ hESC lines that had been created earlier by Dr Eunju Shin (NMHRI, Cardiff University). Protein-level DLG2 knockout and WT DLG2 expression in the context of early cortical interneuron differentiation was examined by using peptide affinity pulldown and mass spectrometry, the latter being undertaken by Dr Mark Collins (Department of Biomedical Science, University of Sheffield). The CNV analysis was done for additional validation of the *DLG2*$^{-/-}$ hESC lines. Moreover, pluripotency was verified by examining the expression of SOX2, OCT4 and NANOG via immunocytochemistry.

After confirmation, a developmental patterning procedure was adopted for the differentiation of the two *DLG2*$^{-/-}$ lines and a genetically WT sister line from hESCs, which went through the CRISPR-Cas9 genome editing pipeline, but remained genetically unaltered. Immunocytochemistry was used to examine the expression of several known neural progenitor and interneuron markers found both *in vivo* and *in vitro* neurogenesis research in order to define the *DLG2*$^{-/-}$ phenotype at two time points in the context of cortical differentiation, namely days 20 and 50. FOXG1 (Fitzgerald et al., 2020; Maroof et al., 2013), SOX2 (Ferri et al., 2013) and NKX2.1 (Butt et al., 2008; Goulburn et al., 2011; Sussel et al., 1999) were amongst the markers employed for NPC profiling and validation of a ventral telencephalic fate. Meanwhile, MAP2 (Soltani et al., 2005) and NEUN (Duan et al., 2016) served as post-mitotic neural markers. TUJ1 worked as a marker of premature/immature postmitotic neurons (Menezes and Luskin, 1994), whilst OLIG2 and COUPTF2 (Alzu'bi et al., 2017a) were associated with cortical interneuron progenitors of MGE and CGE respectively and SST/PV/GABA/GAD67/CALBINDIN/CALRETININ (Fitzgerald et al., 2020; Maroof et al., 2013; Wonders and Anderson, 2006) were associated with mature cortical interneurons. Additionally, neurons GFP^{+} area (boundaries of neuron morphology) was traced and outlined by conducting a morphometric analysis on *DLG2*$^{-/-}$ and WT neurons on day 54 after transduction with a GFP lentivirus. Also, on day 52, GFP-expressing *DLG2*$^{-/-}$ and WT neurons were recorded for 24 hours and traced for the purpose of investigating their migration phenotype.

Chapter 2 contains a detailed procedure of the experiments performed for this chapter. If not specified otherwise in figure legends, at least two separate differentiations were used to obtain the results.

3.3. DLG2 protein expression in WT and DLG2 KO lines

The levels of DLG2 protein expression were analysed via mass spectrometry to establish the authenticity of KO1 and KO2. Consequently, the differentiation of KO1, KO2 and WT hESCs into cortical interneurons was undertaken. Protein samples were collected at two differentiation time points and analysed using mass spectrometry. The latest time point selected was day 50, as this coincides with known function of DLG2 in PSD. This decision was based on the assumption of higher content of mature neurons with synapses, and, implicitly, relatively high DLG2 expression in WT cells at this point. Furthermore, to verify DLG2 expression in NPCs, the time point of day 20 was chosen (Figure 3.1). As protein extracts would be heterogeneous, with a minor proportion being DLG2, possibly lower than the identification threshold for mass spectrometry, proteins containing PDZ domains were enriched. The NR2 subunit of the NMDA receptor (NR2 C-terminus) peptide (SIESDV) is an established PDZ domain interactor, which was employed to perform enrichment via peptide-affinity pulldown (Husi and Grant, 2001). Quantitative mass spectrometry was conducted from days 20 and 50 on the generated protein samples enriched with the PDZ domain interacting peptide. The results were examined for the number of DLG2 unique peptides, as well as the label-free quantification (LFQ) intensity of those peptides; LFQ is a cross-sample normalised intensity measure. The expression of DLG2 occurred both in day 50 samples, as anticipated, and in day 20 at a lower level. Moreover, both KO1 and KO2 cell lines were confirmed as full DLG2 knockouts, which were completely lacking in unique peptides of DLG2 at days 20 and 50 (Figure 3.1). For WT samples, unique DLG2 peptides were absent in two samples on day 20; this suggests that the washing steps during the pulldown sample preparation were suboptimal. Also, the number of unique peptides in the day-50 samples does not differ to those of the day-20 samples; this too might have occurred due to suboptimal sample preparation or immature PSD at this time point (day 50).

In supplementary table 7, there are list of proteins that are significantly high or low in KO samples. They are not a long list of proteins (highlighted in red and green) (Supplementary Table 7). However, it is still hard to interpret the meaning of this as it could be due to at least two cases. First, the significantly decreased proteins can be those that are a part of the DLG2 complexes, hence when DLG2 is absent they cannot bind to the NR2B c-terminal peptide, leading to less amount found in the sample. The second possibility is that their expression itself is reduced due to the lack of DLG2. When

interpreting the increased proteins, it could also be at least the two cases. Firstly, if they are usually binding to the NR2B c-terminal, then by absence of DLG2, there is less competition to the binding, hence more of them could end up in eluted samples. Second case is that their expression itself is increased due to the lack of DLG2. This makes the data for other proteins difficult to interpret unless extra protein experiments are performed. And the primary aim of this experiment was to prove lack of DLG2 protein expression in KO neural precursors and interneurons.

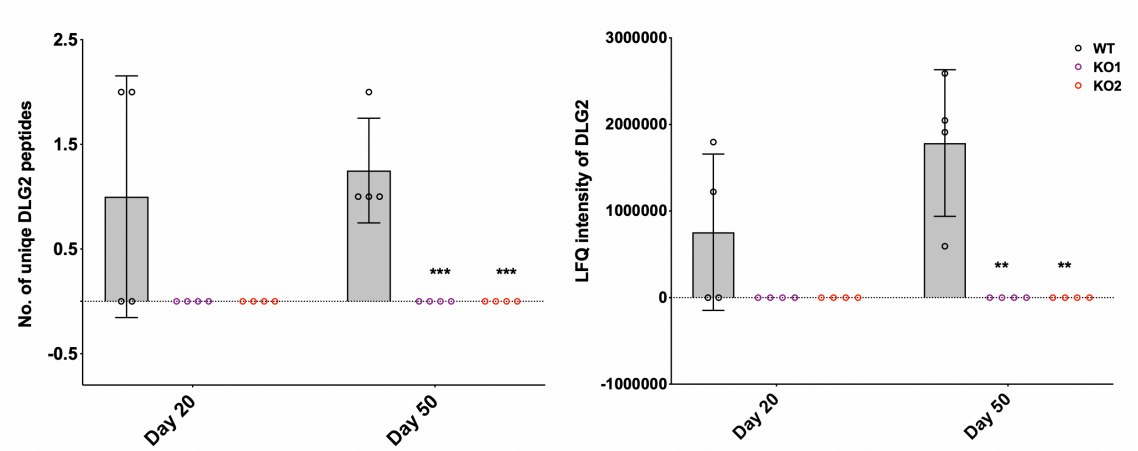


Figure 3.1. Quantification of DLG2 expression by mass-spectrometry in *DLG2*$^{-/-}$ and WT cell lines. DLG2 expression quantification from two *DLG2*$^{-/-}$ knockouts and control (WT) cells lines at two time points (day 20 and 50) of cortical differentiation in PDZ-binding peptide (SIESDV) affinity pulldowns samples utilising LC-MS/MS analysis. The DLG2 unique peptide number and label free quantification (LFQ) intensity established the presence of DLG2 in two samples at day 20 and all day 50 samples of WT, whilst no DLG2 was expressed by either of the *DLG2*$^{-/-}$ (KO1 & KO2) lines at the same time points. One-way ANOVA analysis: unique peptide at day 20 ($F_{2,9}$= 3.000; P=0.1004; n=4); LFQ intensity at day 20 ($F_{2,9}$= 2.797; P=0.1136; n=4); unique peptide at day 50 ($F_{2,9}$= 25.00; P=0.0002; n=4); LFQ intensity at day 50 ($F_{2,9}$= 17.79; P=0.0007; n=4). All data are shown as mean ± SD. Stars represent corrected-Bonferroni multiple comparison test: P<0.05 (*); P<0.01(**); P<0.001(***) vs. WT.

3.4. Further validation of *DLG2*$^{-/-}$ lines

CNV analysis was conducted to verify whether the KO1 and KO2 genomes contained any duplicated regions or if there were additional deletions compared to the WT line, thus further confirming the validity of the lines. This CNV analysis revealed that, in comparison to the reference genome, every hESC line, including the WT line, had two deleted and duplicated regions that were not pathogenic in nature. In spite of this, the KO lines and WT did not differ; hence, the lines were used in the current research (Figure 3.2).

All CNVs were checked using the UCSC genome browser (https://genome-euro.ucsc.edu/cgi-bin/hgGateway?redirect=manual&source=genome.ucsc.edu) (The February 2009 human reference sequence (GRCh37)). The genes in each region were listed. Each gene was checked on NCBI (for GO terms, links to papers) to check if the gene is involved in synaptogenesis or involved in neurodevelopment (Supplementary File 4). In addition, using the WT one cluster versus the rest analysis, the lists of significant DEGs for each cluster were combined to create a single 'developmentally

regulated' set of genes (in total 4975 significant DEGs). Genes in the 'developmentally regulated' set hit by each CNV were then identified. As well as identifying the individual genes, the percentage of developmentally regulated genes for each CNV was calculated. As a comparator, a set of 'all expressed' genes (genes expressed in one or more WT cluster, 16735 genes in total) was created and the percentage of these in the developmentally regulated set was checked. The 'developmentally regulated' set contains 4975 genes. The CNVs file contains 460 genes (Supplementary File 4), and the CNV genes associated with neurodevelopment/synaptogenesis were 28 genes (checked on NCBI for GO terms, links to papers-highlighted in yellow in Supplementary File 4). The intersection between CNV neurodevelopment/synaptogenesis genes (28 genes) and developmental regulated DEGs (4975 genes) is 15 genes (ATP9A, STMN3, KCNQ2, PHACTR3, EEF1A2, TMEM230, TGIF2, SLC32A1, PTPRA, SLC12A5, SNAP25, PTPRT, SS18L1, FLRT3, FOXA2). So, 28 out of 460 genes hit by CNVs (i.e. 6% cnv-hit genes) have previously been shown to play a role in synaptogenesis or neurodevelopment. The intersection between all CNV genes (460) and DEGs (4975 genes) is 136 genes. So, 136 out of 460 genes hit by CNVs (i.e. 30% cnv-hit genes) were developmentally regulated in the scRNAseq data, compare to 4975 out of all 16735 expressed genes (= 30% of all genes). So CNVs were not particularly enriched for developmentally regulated genes in my dataset (they look about average). It is noteworthy that the same CNVs were defined in all three cell lines (WT, KO1 and KO2) and they are not known to be pathogenic in nature.

The authenticity of KO1, KO2 and WT as hESC lines was validated by carrying out immunocytochemistry, evaluating the expression of major pluripotency genes (SOX2, OCT4 and NANOG). These genes are key transcriptional regulators that are highly expressed in pluripotent cells, and the loss of these markers indicates the loss of pluripotency (Boyer et al., 2005). The expression of these markers was identified in almost all cells in both KO and WT cultures, confirming the retention of the pluripotency in these lines. These results collectively demonstrated that both *DLG2*$^{-/-}$ knockouts (KO1 and KO2) were DLG2 homozygous knockout hESC lines that had not acquired any other unintended genetic alterations compared to the control sister WT line (Figure 3.3).

Cell line	CNV location	Duplication/Deletions	Position	Approximate Length
WT, KO1, KO2	2q37.3	Deletion	Ch2:242,900,000-243,100,000	300,000Kb
WT, KO1, KO2	3q26.1	Duplication	Chr3:166,000,000-166,400,000	400,000Kb
WT, KO1, KO2	20p12-p11.1	Deletion	Ch20:63244-25889302	25,826,059Mb
WT, KO1, KO2	20q11.21-q13.33	Duplication	Chr20:30,000,000-62,500,000	23,500,000Mb

Figure 3.2. WT and DLG2 KO CNV analysis. Using the Illumina PsychArray v1.3, DNA genomic samples of all cell lines (WT and two KOs) were analysed (using PennCNV). During PennCNV QC, CNVs comprising fewer than 10 SNPs and smaller than 100Kb were filtered out. Between the two genotypes (WT & KO), no further CNVs were identified.

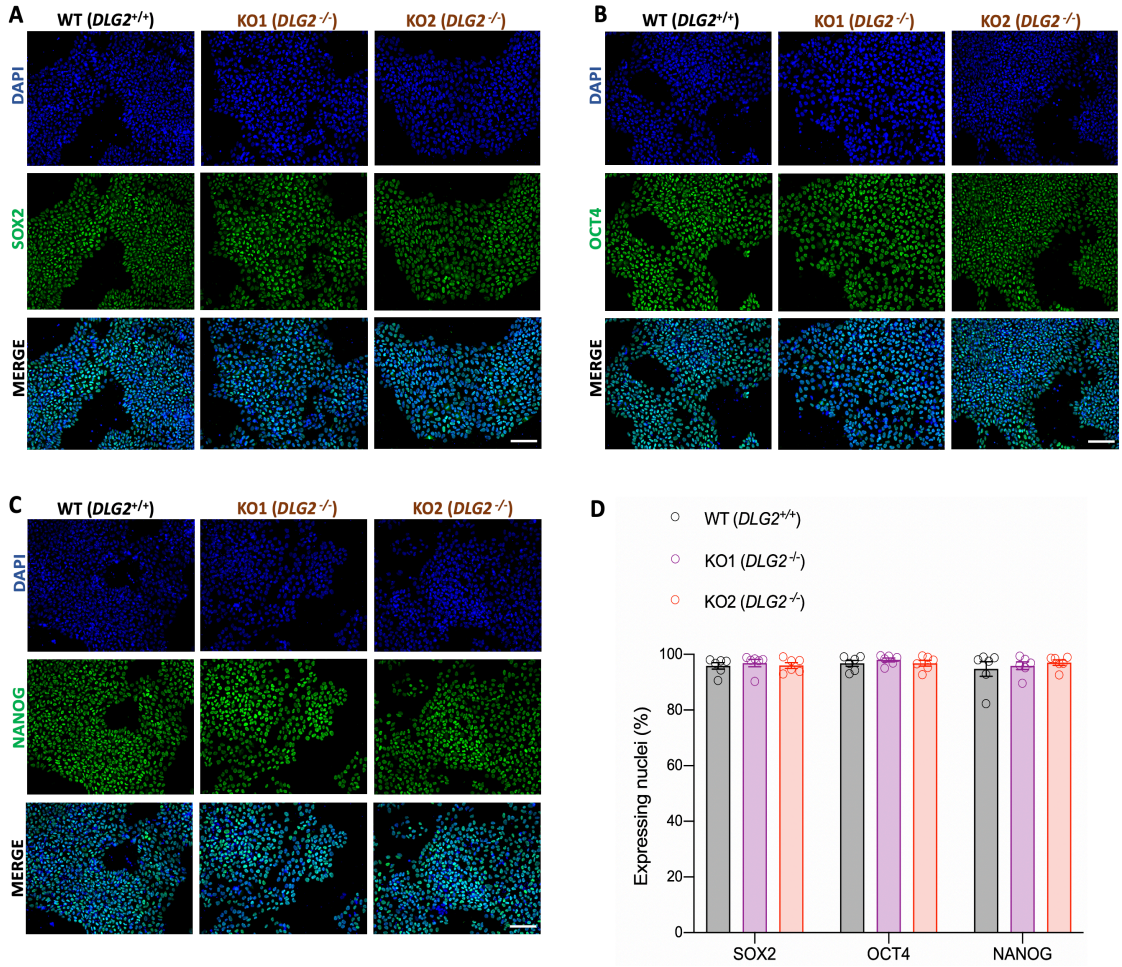


Figure 3.3. Expression of pluripotency markers in both *DLG2*$^{-/-}$ KOs and WT hESCs. Representative immunocytochemistry images of nuclear counterstain DAPI with SOX2 **(A)**, OCT4 **(B)** and NANOG **(C)** expression in WT and both (*DLG2*$^{-/-}$) KO hESC lines. **(D)** Quantification of nuclei expressing SOX2, OCT4 or NANOG in two (*DLG2*$^{-/-}$) KO and WT hESC lines. It is clear that all cell lines (including KOs) express core pluripotency genes. For each marker, analysis was performed using one-way ANOVA: SOX2 $F_{2,15}$= 0.2238; P=0.8021; n=6, OCT4 $F_{2,15}$= 0.5516; P=0.5873; n=6; NANOG $F_{2,15}$= 0.3948; P=0.6806; n=6. The genotype did not show any significant effect on the expression of these pluripotency genes. The bar graph presents the mean ± SEM. All scale bars are 100µm.

3.5. Characterisation of day 20 cortical interneuron cultures derived from DLG2 KO and WT hESCs

The fate of both *DLG2*$^{-/-}$ KOs and WT interneuron progenitors was explored. To characterise their phenotypes, WT and *DLG2*$^{-/-}$ KO cells were differentiated using a cortical interneuron differentiation protocol and immunocytochemistry was performed on day 20 of differentiation.

The immunocytochemistry findings demonstrate that as expected, the cells originating from all lines were mostly ventral forebrain-fated (MGE fate) neural precursor cells, which in WT, expressed typical markers, such as FOXG1 (77%), SOX2 (87%) and NKX2.1(76%). Conversely, scarce neuronal marker expression was observed, such as the marker of immature postmitotic neurons, TUJ1 (1.8% in WT). This reaffirmed that

cells at this time point are mainly neural precursors, and not mature neurons or stem cells.

The day 20 immunocytochemistry evaluation of FOXG1, SOX2 and TUJ1 displayed no evident trends or significant variations amongst cell lines, implying expression by comparable cell proportions in the WT and *DLG2*$^{-/-}$ lines. However, significantly less OLIG2 was observed in the deficient *DLG2*$^{-/-}$ cells compared to WT (2.9% KO1, 3.2% KO2 vs 4.7% WT). Furthermore, compared to WT, cells stained for NKX2.1 (expressed in MGE-derived progenitors) showed a significant decrease in NKX2.1 (53.1% KO1, 60.4% KO2 vs 76% WT) and a slight increase in COUPTF2 (expressed in CGE-derived progenitors – not significant) was observed in the DLG2-deficient cell lines (Figures 3.4 A & B). Finally, at day 20, cells stained for PAX6 (excitatory progenitors) and FOXP2 (marker for LGE progenitors) showed no expression in the three cell lines (data not shown). This confirms that the desired identity of the cultures and DLG2 deficiency did not have an impact upon the gross cell fate change within the forebrain. If resources and time had allowed, the performance of other techniques, such western blot, would have been useful to see if the level of protein expression of these markers in each cell would decrease where the percentage of cells expressing the protein was comparable. Despite immunocytochemistry-based population level analysis being enlightening and informative, owing to restricted resources, it is unable to reflect every detail of *DLG2*$^{-/-}$ progenitors. A single cell-level high throughput approach, such as single-cell RNA sequencing, would complement the current approach.

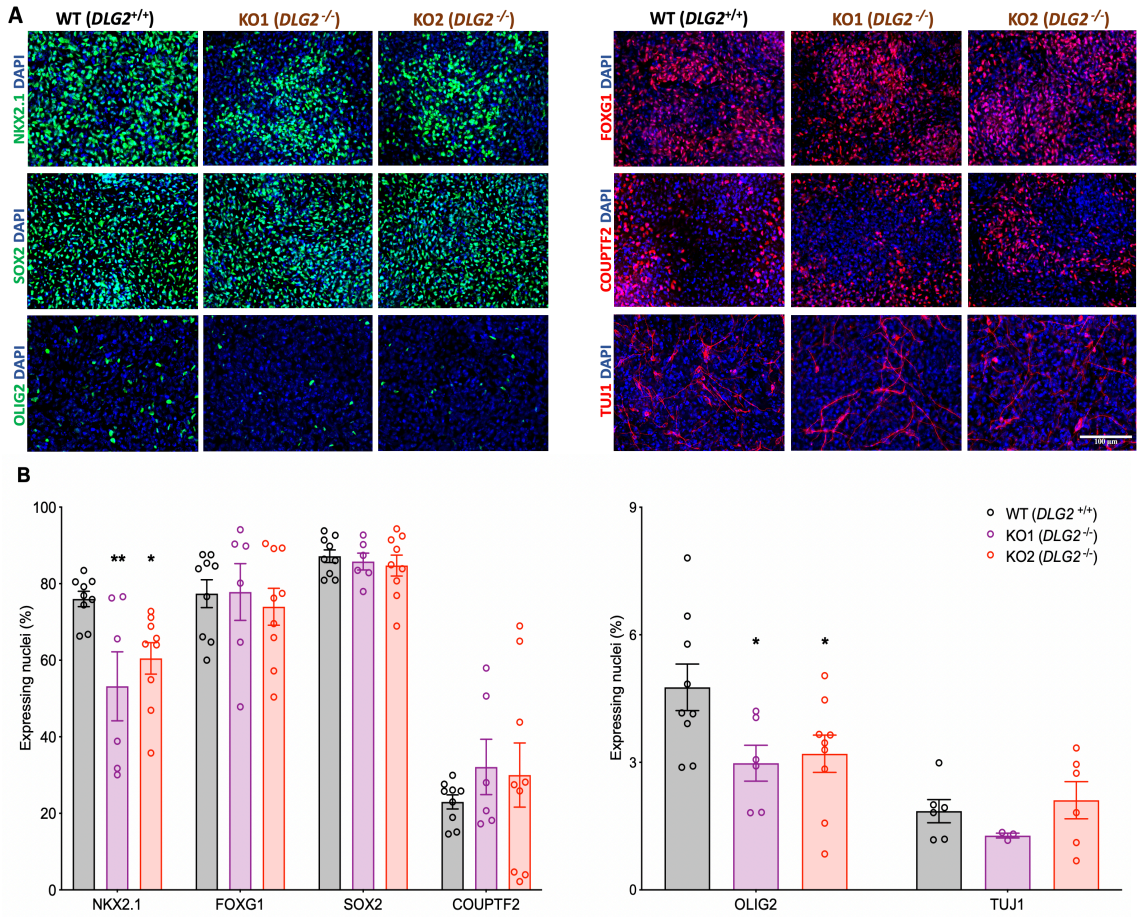


Figure 3.4. Day 20 immunocytochemistry phenotyping and characterisation of *DLG2*$^{-/-}$ and WT cortical progenitors. A) Representative immunocytochemistry images of NKX2.1, FOXG1, SOX2, COUPTF2, OLIG2 and TUJ1 with DAPI nuclear counterstain in WT and two *DLG2*$^{-/-}$ KO cell lines at day 20 of cortical interneuron differentiation. **B)** Quantification of nuclei expressing NKX2.1, FOXG1, SOX2, COUPTF2, OLIG2 and TUJ1 in WT and two *DLG2*$^{-/-}$ KO cells. As determined by one-way ANOVA, genotype did not show any significant effect on expression levels of FOXG1($F_{2,21}$=0.1802; P=0.8363; n≥6), SOX2 ($F_{2,21}$= 0.3185; P=0.7307; n≥6), COUPTF2 ($F_{2,21}$=0.5565; P=0.5814; n≥6) and TUJ1 ($F_{2,12}$=1.042; P=0.3826; n≥3, from one differentiation only for KO1). However, the genotype did show a significant effect on the expression levels of NKX2.1 ($F_{2,21}$=5.607; P=0.0112; n≥6) and OLIG2 ($F_{2,21}$=4.005; P=0.0336; n≥6). Post hoc analysis (Dunnett's multiple comparison test) showed significant differences of the individual marker NKX2.1 with p value of WT vs. KO1 P=0.0090 and WT vs. KO2 P=0.0474 and OLIG2 with p value of WT vs. KO1 P=0.0445 and WT vs. KO2 P=0.0497. The stars represent significance in the corrected-Bonferroni multiple comparison test: P<0.05 (*); P<0.01(**) vs. WT.

3.6. Characterisation of day 50 cortical interneuron cultures derived from DLG2 KO and WT hESCs

To investigate the effect of DLG2 deficiency in more developed neurons, immunocytochemistry was performed at day 50 of cortical differentiation. It was apparent that cells from both KOs and WT lines could differentiate from neural progenitors into postmitotic neurons with interneuron-specific gene expression, such as PV, SST, GABA, GAD67, CALRETININ and CALBINDIN (Figure 3.6).

Immunocytochemistry pertaining to NEUN, together with the quantification of NEUN-positive cells, demonstrated no significant change in the percentage of postmitotic neurons arising from the *DLG2*$^{-/-}$ KO cells versus control cells. This was in keeping with

the similar fraction of neural precursor cells observed in all three lines (WT, KO1 and KO2) at day 20 and equivalent degrees of TUJ1, NEUN and SOX2 expression. SOX2-expressing cells were similar across day-50 cell lines (Figure 3.5 B); however, their percentage was lower compared to day 20. This is not surprising, because of the increased number of NEUN (postmitotic neurons) (22% in WT) at day 50. The double staining of NeuN and neuronal subtype markers would give an accurate number of mature cells expressing the marker, however a rough estimation was conducted to calculate the neuronal subtypes (SST, PV, CALRETININ, CALBINDIN, GABA and GAD67) percentage out of NeuN percent (average of NeuN%). There were similar number of KO SST^+ cells compared to WT (12% KO1, 11% KO2 vs 10.3% WT). For PV^+ cells, there was very low expression when considering WT cells; however, a similar number of PV^+ cells were observed (0.53% KO1, 0.51% KO2 vs 0.48% WT). Also, similar percentages of CALRETININ, CALBINDIN and GAD67 positive cells were observed. However, there were decreasing trend in the cells exhibiting GABA (17.2% KO1, 24.7% KO2 vs 28.7% WT) at cortical differentiation day 50. Thus, lack of *DLG2* does not seem to influence the speed at which neurons are synthesised from NPCs, but it may hinder the expression of cortical interneurons' identity. It was not clear whether fewer cells expressing NKX2.1 at day 20 led to fewer cells expressing GABA, or whether this was a separate event – i.e. whether delayed expression in this gene occurred independently (GABA). Due to the restricted resources, it was not possible to investigate all key MGE/GE-specific TF expression. Further research into the expression of genes such as DLX1, 2, 5, 6, OTX1 and SIX3 might reveal a view of the general feature in TFs expressed in MGE when DLG2 is deficient.

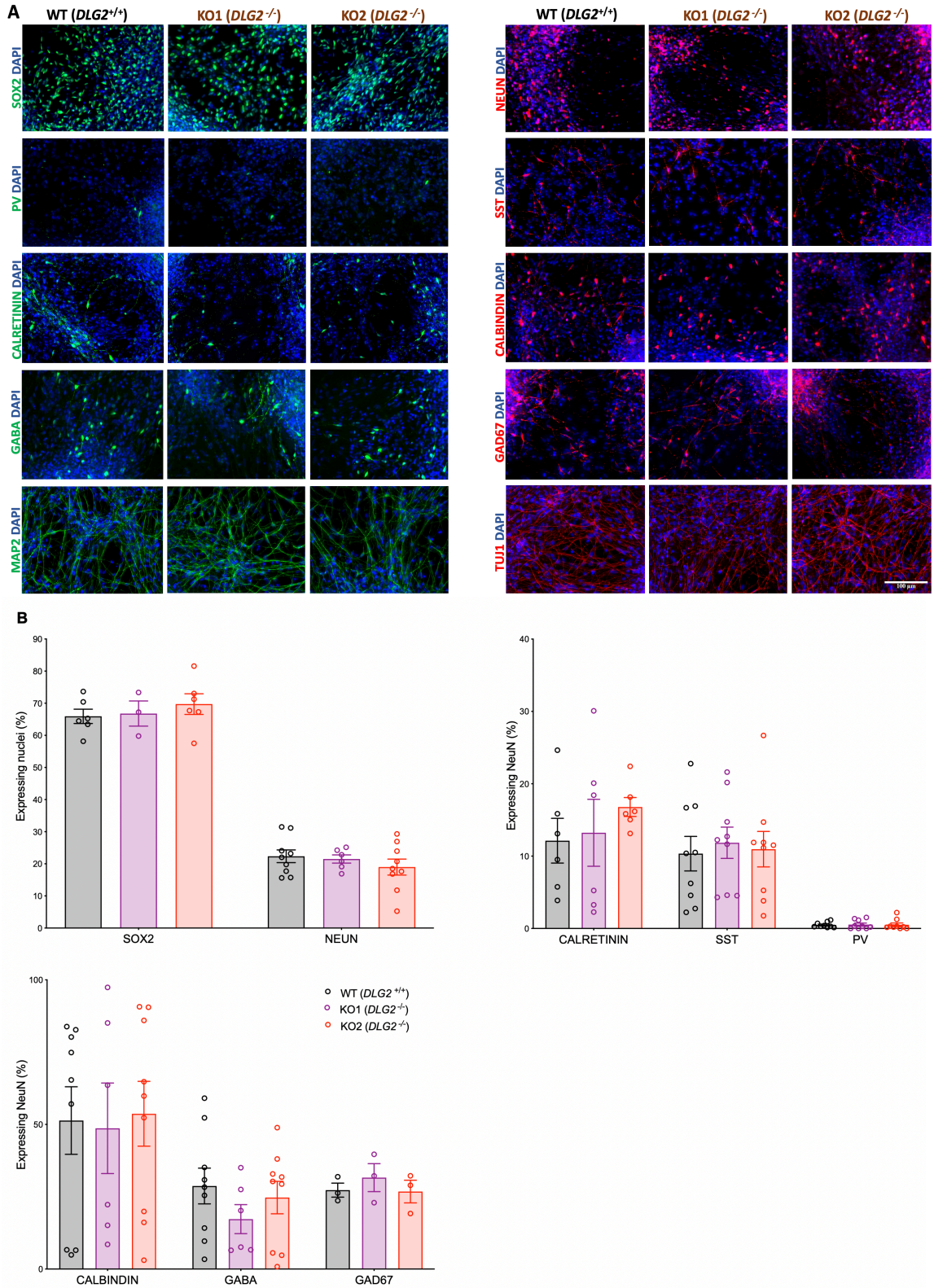


Figure 3.5. Day 50 immunocytochemistry phenotyping and characterisation of *DLG2*$^{-/-}$ and WT cortical interneurons. A) Representative immunocytochemistry images of PV, SST, SOX2, NEUN, CALRETININ, CALBINDIN, GAD67, GABA, MAP2, and TUJ1 with DAPI nuclear counterstain in WT and two *DLG2*$^{-/-}$ KO cell lines at day 50 of cortical interneuron differentiation. **B)** Quantification of nuclei expressing SOX2, NEUN and quantification of NeuN expressing PV, SST, CALRETININ, CALBINDIN, GAD67 and GABA in WT and two *DLG2*$^{-/-}$ KO cells. As determined by one-way ANOVA, genotype did not show any significant effect on expression levels of SOX2 ($F_{2,12}$=0.506; P=0.6155; n≥3, from one differentiation only for KO1), NEUN ($F_{2,21}$= 0.712; P=0.5022; n≥6), CALRETININ ($F_{2,15}$=0.546; P=0.5901; n=6), SST ($F_{2,24}$=0.104; P=0.9012; n=9), PV ($F_{2,24}$=0.0211; P=0.9791; n=9), CALBINDIN ($F_{2,21}$=0.0478; P=0.9534; n≥6), GABA

($F_{2,21}$=0.863; P=0.4362; n≥6) and GAD67 ($F_{2,6}$=0.469; P=0.6465; n=3, from one differentiation only).

3.7. Effect of DLG2 deficiency on cortical interneuron migration

The same proportion of neurons was generated from KO and WT hESCs: hence, the migratory function of the neurons generated was compared. A migration assay was performed on the *DLG2*$^{+/+}$ and *DLG2*$^{-/-}$ neurons at day 52. Cells were infected with GFP lentivirus (hPGK-GFP), and a scanning plan was set to scan every hour for 24 hours. Measurement of migration was performed on 30 neurons/cell line (15 neurons/round). The GFP-expressing neurons were randomly picked and their migration was traced to estimate the average speed and the final displacement at 24 hours. One-way ANOVA and Dunnett's multiple comparisons tests revealed statistically significant differences in both criteria (Figures 3.6 A & B). Neurons derived from the DLG2 deficient hESC lines showed disruption in their movement, suggesting a role for DLG2 in migration.

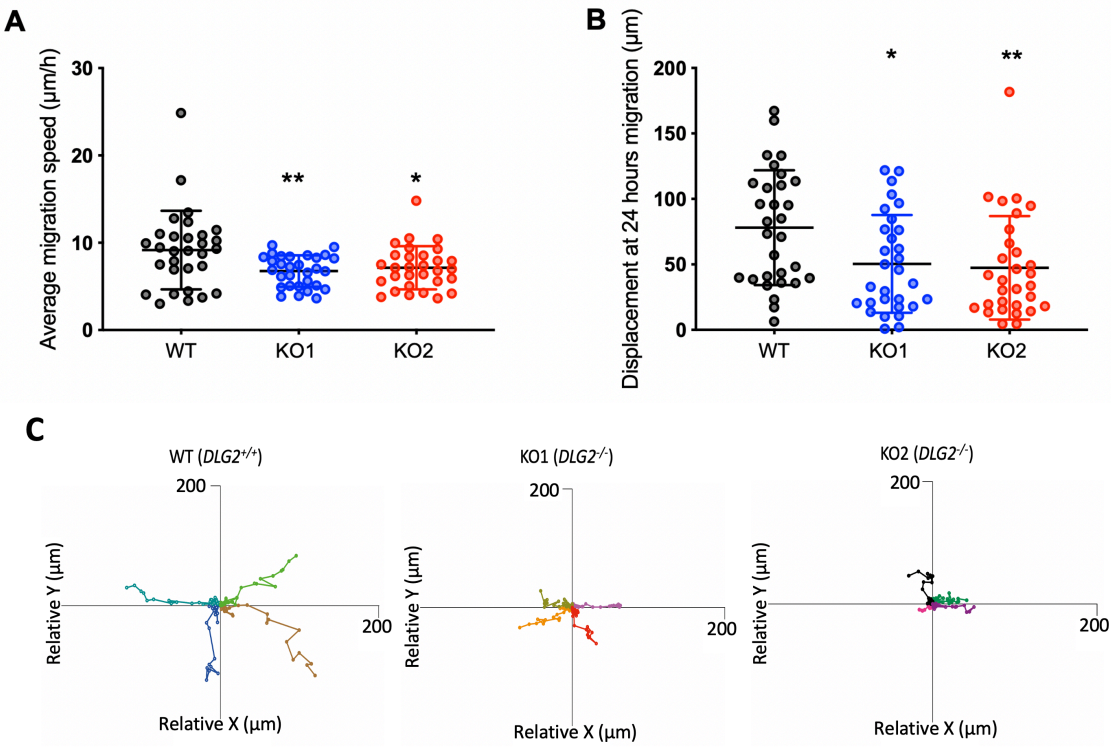


Figure 3.6. Migration of interneurons differentiated from WT and *DLG2*$^{-/-}$ hESCs. A) Shows how quickly the interneurons moved (the average speed of individual neurons per hour). One-way ANOVA and Dunnett's multiple comparisons test showed a statistically significant difference ($F_{2,87}$ = 5.04, P=0.0085) compared to WT. **B)** Shows the final displacement of every single neuron at 24 hours. One-way ANOVA and Dunnett's multiple comparisons test showed a statistically significant difference ($F_{2,87}$ = 5.27, P=0.0069). The bar represents the mean ± SD. **C)** Example movements (traces) of migrating interneurons from a given origin in *DLG2*$^{+/+}$ and *DLG2*$^{-/-}$ cultures (the total movement and the final location point of neurons). *DLG2*$^{-/-}$ neurons show disruption in the movement and final displacement of *DLG2*$^{-/-}$ neurons is less compared to the WT neurons. The stars represent significance in the corrected-Bonferroni multiple comparison test: P<0.05 (*); P<0.01(**) vs. WT.

3.8. DLG2 deficiency affects cortical interneuron morphology

Having established a migratory defect in *DLG2*$^{-/-}$ cortical interneurons, a morphometric assessment of *DLG2*$^{-/-}$ and WT interneurons was conducted. More advanced (i.e. day 54) *DLG2*$^{-/-}$ and control neurons were tracked following transduction with a lentivirus that expressed GFP; this enabled the principal neuronal configurations to be discerned. A marked decrease in the number of secondary neurites; that is, those that protrude from primary branches as opposed to the cell body was observed in *DLG2*$^{-/-}$ neurons. Since the *DLG2*$^{-/-}$ neurons exhibited a similar soma area, total neurite length and number of primary neurites compared to WT, this gives rise to an obvious phenotype of elongated neurites with less branching in KO than WT (Figure 3.7).

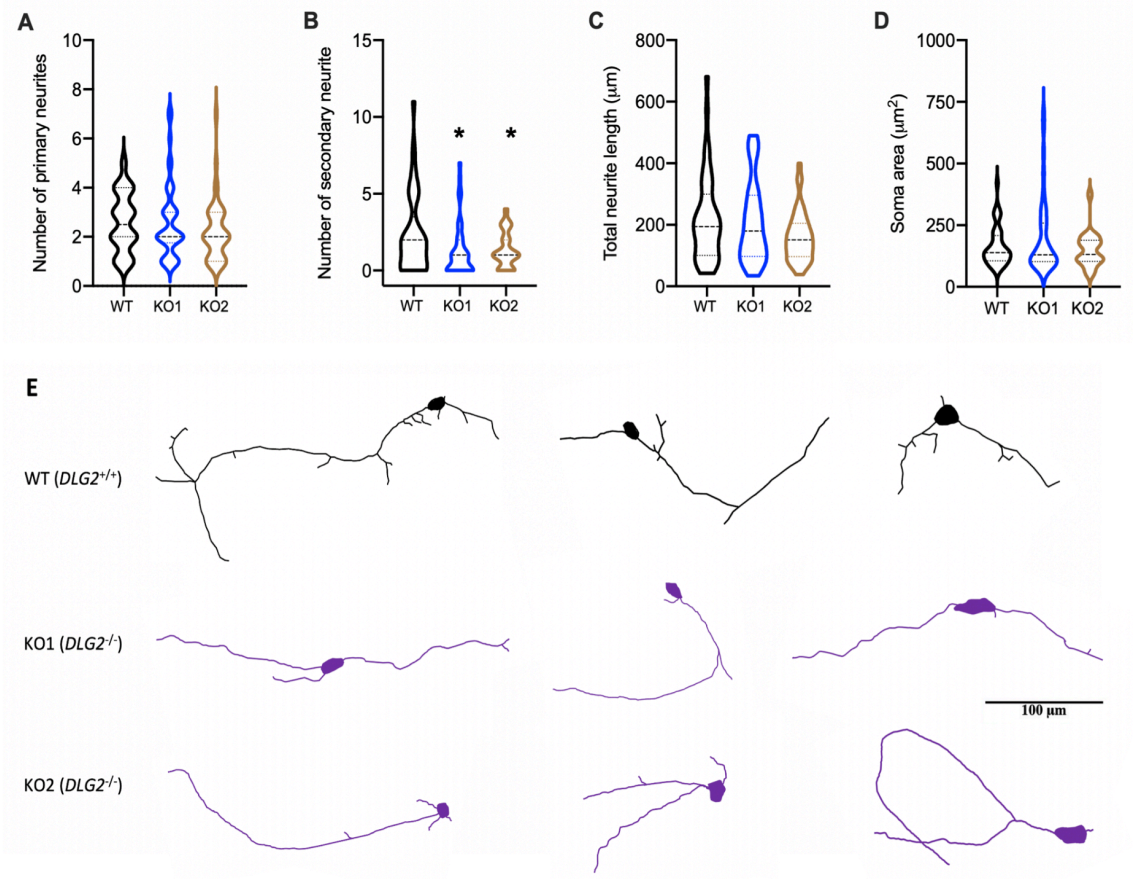


Figure 3.7. Characterisation of neuronal morphology in *DLG2*$^{+/+}$ and *DLG2*$^{-/-}$ cultures. (A-D) *DLG2*$^{-/-}$ and WT neurons' morphological quantification at day 54 of differentiation. **A)** Genotype did not have a significant effect on primary neurite numbers (projecting from soma) ($F_{2,123}$=1.061; P=0.3491; n=42). **B)** Genotype and Dunnett's multiple comparisons test did have a significant effect on secondary neurite numbers (projecting from primary branches) ($F_{2,123}$=4.186; P=0.0174; n=42). **C)** Genotype did not have a significant effect on total neurite length ($F_{2,123}$=2.867; P=0.0606; n=42). **D)** Genotype did not have a significant effect on the area of soma ($F_{2,123}$=2.220; P=0.1129; n=42). The bar graph presents the mean ± SEM. **E)** Representative tracing of neurons. The stars represent significance in the corrected-Bonferroni multiple comparison test: P<0.05 (*) vs. WT.

3.9. Discussion

3.9.1. Absent expression of DLG2 in knockouts and presence in WT cells

Using CRISPR/Cas9-D10a nickase genome modification, a pair of *DLG2*$^{-/-}$ lines, KO1 and KO2, were established. After a peptide affinity pulldown of PDZ domain-containing proteins, mass spectrometry was performed to verify that WT express DLG2 protein and both KO cell lines were complete knockouts, lacking the ability to produce DLG2 protein. This was carried out on both KO lines and WT cells at cortical differentiation on days 20 and 50. Mass spectrometry analysis confirmed that DLG2 protein was absent in the knockouts but present in WT on days 20 and 50. These data suggest that performing mass spectrometry on WT specimens from earlier phases of cortical differentiation, such as day 20, was advantageous, as it elucidated the longitudinal nature of DLG2 expression *in vitro* more precisely. There is evidence to indicate that, prior to neurite outgrowth, synaptic proteins such as GAP-43, SNAP-25 and NRGN were initially expressed and secreted by progenitors of neurons between day 10 and day 25 of iPSC-derived cortical neurons' differentiation (Nazir et al., 2018). Therefore, this may lead to the possibility that DLG2 has a role to play during the early stage of development, before the presence of PSD, because at this time point it is unlikely that effective or significant synapse development has happened (Semple et al., 2013; Silbereis et al., 2016). This finding could confirm the findings from in Cortecon's study, reported in Section 1.5.5, in which *DLG2* mRNA expression was exhibited at all time periods and throughout early neurogenesis of *in vitro* cortical differentiation, involving day 0 of hESCs (van de Leemput et al., 2014), indicating that *DLG2* is present prior to neurodevelopment or synaptogenesis.

The emphasis of this study was to find a role of canonical DLG2 that contains three PDZ, one SH3 and one GK domain during cortical interneuron differentiation, enrichment utilising a PDZ-peptide interactor, such as 'SIESDV', was a valid approach as no suitable DLG2 antibody was identified. If enrichment of the entire set of DLG2 isoforms was necessary, including those with absent PDZ domains, alternative strategies could be deployed, such as extracting protein from synaptosomes, rather than the entire cells. This could be a valid approach for mature interneurons (perhaps the duration for differentiation should be lengthened, for example, until day 60).

3.9.2. Successful validation of *DLG2*$^{-/-}$ hESCs for CNV and expression of pluripotency marker

In order to recognise any replicated or deleted genomic areas, CNV testing was conducted. Data from that analysis reveal that *DLG2*$^{-/-}$ lines and a WT sister line displayed two duplicated and two deleted regions that are not known to be pathogenic

in nature when judged against the reference genome (genes in each CNVs region detected were checked if they were involved in synaptogenesis or neurodevelopment (see Section 3.4)). This could be the consequence of passaging to sustain the H7 (parental hESC line), which has been inferred in CNV insertion (Liu et al., 2014). Independently, some laboratories have reported chromosome alterations in H7 after a high number of passages, such as 60 (Catalina et al., 2008; Draper et al., 2004; Maitra et al., 2005). However, of the identified CNVs, a minimum of one – the duplication of 20q11.21 – seems to be characteristic of the cell line, possibly arising from its inception (Laurent et al., 2011). Despite not comparing every single variant but comparing large structural variants, all three lines including parent cell line, sister line and KOs have the same CNVs. Also, the fact that sister line was used for control gives better justification in experimental design than using the parent line. Given that the two *DLG2*$^{-/-}$ lines (KO1 & KO2) and the WT sister line were equivalent (express pluripotency markers OCT4, NANOG and SOX2), with no extra CNVs, this additionally implies that the genome modification was successful.

3.9.3. Use of *DLG2*$^{-/-}$ hESCs directed differentiation to model human cortical interneuron development

Utilising a developmental patterning protocol optimised from Kim *et al*., Maroof *et al*. and Chambers *et al*., (Chambers et al., 2009; Kim et al., 2014; Maroof et al., 2013) several guided differentiations from hESCs to cortical interneurons were carried out on KO1, KO2 and WT cell lines. It was fully expected that the *DLG2*$^{-/-}$ hESCs (KO1 & KO2) would achieve full neuronal differentiation, since they express the typical pluripotency genes OCT4, NANOG and SOX2 at hESC stage (Boyer et al., 2005). Also, since the technique *in vitro* essentially repeats the developmental pathway *in vivo* (see Section 1.1 and 1.3) (Fitzgerald et al., 2020), it was anticipated that cells from all cell lines would demonstrate gradual progression, developing from pluripotent hESCs to neural precursors, early immature and ultimately into postmitotic interneurons.

The two *DLG2*$^{-/-}$ lines (KO1 & KO2) successfully differentiated from being hESCs displaying above pluripotency markers, to being neural precursors at day 20 that expressed SOX2, FOXG1 and NKX2.1. The last two markers are of particular importance, as they validate that the neural precursors have a ventral telencephalic fate (Butt et al., 2008; Fitzgerald et al., 2020; Goulburn et al., 2011; Maroof et al., 2013; Sussel et al., 1999), and thus are antecedents to cortical interneurons, verifying the correct lineage of intended differentiation. The expression of these markers is similar to those in a previously reported study (Close et al., 2017): which showed 86% of FOXG1^{+}

and 74.5% NKX2.1^{+} cells at day 24 of cortical differentiation; the current data shows 77% of FOXG1^{+} and 76% of NKX2.1^{+} cells on day 20.

Immature/developing neurons in culture observed on the 50th day, expressed the postmitotic neuronal gene, NEUN, together with GABA, GAD67, CALBINDIN and CALRETININ, which are markers of cortical interneurons. This is in keeping with the usual neural cortical development, where intraneuronal markers are present at high levels at later time points (Kim et al., 2014; Nicholas et al., 2013). Thus, the differentiation of *DLG2*$^{-/-}$ hESC lines into ventral telencephalic NPCs and cortical interneurons that express markers typical of these cells was achieved. This verified that these lines can be utilised effectively as a model for the interneuron neurogenesis of the human cortex.

It is important to note that the numbers of cIN subtypes found at day 50 of differentiation was at low level because the maturation of cINS necessitates stimulation from the surrounding environment (Patz et al., 2004). This could be achieved by co-culturing cells with mouse glial cells or excitatory cells or glial cells in combination with mouse neurons. This method has been shown to accelerate the cells maturation (Maroof et al., 2013; Nicholas et al., 2013; Patz et al., 2004). Another strategy would be elongating the protocol time (e.g. day 70 or 100) to give more cINs numbers.

Also, particularly, the very low levels of PV numbers which is one of major difficulties because PV neurons are born a bit later (Chu and Anderson, 2015; Xu et al., 2010b) with early expression of PV reported in mouse postnatally from day 4 onward (Davila et al., 2005) and they need longer time to become fully functional (Letinic and Kostovic, 1998). So, cINs will need to be cultured for around 10 months to reach postnatal stage for human which is not possible as neurons may not survive by that time. Also, it is known that PV neurons are activity-dependent (Ferguson and Gao, 2018) and within the culture there is no activity or connection from pyramidal neurons or other types of neurons to trigger their activity.

3.9.4. Expression of cellular identity has been affected significantly by the absence of DLG2

Day 20 culture is supposed to be mostly neural precursor cells, and this was confirmed by SOX2 and TUJ1 staining on day 20. At the NPC stage (day 20), there was a significant decrease in the percentage of interneuron progenitors expressing NKX2.1 (which marks the MGE-derived progenitors) and OLIG2 (expressed in a subpopulation of MGE-derived progenitors). Also, a slight increasing trend in COUPTF2-expressing cells (expressed in CGE-derived progenitors) was observed. However, FOXG1 (a telencephalic marker), SOX2 (neuronal stem cells) and TUJ1 (immature postmitotic neurons) showed no

differences. It is known that FOXG1 is needed in the ventral telencephalon for region-specific regulation of telencephalic NPC specification and proliferation (Fitzgerald et al., 2020; Maroof et al., 2013). According to the published literature, nearly all cells express FOXG1 (a forebrain marker) and NKX2.1 (a MGE marker) on day 25 of hESC differentiation (Liu et al., 2013). Thus, in the current study similar amounts of FOXG1 (i.e. telencephalic or forebrain cells) were detectable in the WT and KO cultures on day 20, but it seems that their fate changed in the KO cell lines or that fate specification progression is slower in KOs. Moreover, the percentage of SOX2 was similar in all three cell lines, which means that the decrease in NKX2.1 was independent of SOX2 expression. SOX2 is needed for the early expression of NKX2.1, as demonstrated by E11.5 mouse *Sox2* mutants, in which NKX2.1 expression is absent in the telencephalon (Ferri et al., 2013). The NPCs lacking NKX2.1 expression did not seem to become a totally different lineage. This was clear when cells were stained for PAX6 (excitatory progenitors) and FOXP2 (marker for LGE progenitors). Therefore, the majority of cells have GE (except for LGE) identity, but there might be a shift of identity from MGE to CGE if this was not simply a delayed identity expression.

Lack of DLG2 does not seem to influence the speed of neurogenesis; that is, the differentiation of NPCs into postmitotic interneurons. On day 50 of cortical interneurons differentiation, no notable variation was observed in the postmitotic neuronal marker expression of NEUN in *DLG2*$^{-/-}$ cultures compared to WT. Although immunocytochemistry was performed for MAP2 and TUJ1 on day 50, it was hard to quantify positive cells, as neurons were clumped and their neurites had become mingled. This was the case for some SST^{+}, GABA^{+} and GAD67^{+} cells, although clearly identified positive cells were quantified. If time and resources had permitted, quantification of protein level using mass spectrometry or western blot would have given a clear indication of the DLG2$^{-/-}$ impact on the GABAergic population.

At the interneuron stage (day 50), all three cell lines had similar levels of SOX2 (neuronal stem cells) and NeuN (postmitotic neurons). Around 66% of cells were neuronal progenitors and 22% were postmitotic interneurons. Compared to WT cells, KO cells exhibited a declining trend in GABA. Perhaps it is not surprising to observe the decrease of GABA^{+} cortical interneurons in KO, as NKX2.1 (MGE progenitor) expression was decreased in the deficient cell lines at the NPC stage, and these are the precursor population to cortical interneurons (Goulburn et al., 2011; Sussel et al., 1999). Additionally, the reduction in GABA^{+} cells might be independent of the reduced NKX2.1^{+} population, again, this might have occurred as a delay in maturation, but to verify this, further investigation would be required at later time points.

Despite the fact that *DLG2*$^{-/-}$ cultures may yield neurons at an equivalent rate as WT, KO neurons exhibited a more immature structure with markedly diminished secondary branching of the neurites on day 54. This may be due to changes in the identity of *DLG2*$^{-/-}$ neurons, with a possibly less branched neuronal type, or a delay in morphological development. To obtain additional insight into neuronal development, morphological characteristics were analysed, which included analyses of primary neurite numbers, soma size and total neurite length showed no differences. This means that the morphological development of soma and primary neurites had not been affected by the lack of DLG2. However, a significant reduction in the number of secondary neurites was observed in KO cell lines. Similar effect has been observed in DLG2 KO excitatory neurons differentiated from hESC (at day 30 and 70) (Sanders et al., 2022). The immature shape of *DLG2*$^{-/-}$ neurons might potentially be interpreted by way of a delay in displaying their latent neural identity. However, a phenomenon like this needs to be checked over a longer time period (e.g. day 60). It is interesting to note that there seem to be at least two distinct developmental pathways: DLG2-independent and dependent. The former includes higher-level cell type progression, such as from NPCs to neurons (neurogenesis itself), whereas the latter includes delayed subtype identity expression of NPCs and interneurons.

3.9.5. Significant migration differences detected between WT and *DLG2*$^{-/-}$ neurons

ErbB4 has an intracellular C-terminal area, to which *DLG2* binds via PDZ domains. It is known that ErbB4 plays a key role in tangential migration of interneurons by responding to EGF and neuregulin family members (Flames et al., 2004; Garcia et al., 2000). Thus, on day 52, the migration of neurons was traced for 24 hours. Migration analysis of cortical interneurons showed a decrease in the movement of the deficient neurons compared to the controls. Most of the cells in the deficient cell line showed disruption in their movement (neurons in the KO cell lines migrated at lower speed and over less distance from the origin after 24 hours), suggesting a role for DLG2 in migration. Correspondingly, Flames *et al.* reported deficits in the migration of interneurons in mutant mice that lacked *ErbB4*. So, it is possible that loss of interaction between ErbB4 and DLG2 might have led to the migration defect seen in *DLG2*-deficient cell lines. It is not clear what exactly was driving the migration of neurons in the culture, but as ErbB4 is activated by Neuregulin1, and ErbB4–NRG1 signalling is crucial for the assembly of GABAergic circuits (Flames et al., 2004), the addition of NRG1 expressing cells might exaggerate the migration defect of *DLG2*$^{-/-}$ neurons if this deficit was elicited by the lack of interaction between ERBB4 and DLG2.

3.9.6. Postulated mechanism of *DLG2*$^{-/-}$ phenotype

Immunocytochemistry, migration and morphometric analysis results imply a function of DLG2 in the establishment of neuronal sub-identity. A smaller percentage of *DLG2*$^{-/-}$ neurons exhibit cortical GABA identity expression, their migration is reduced and their morphology is less developed. The findings do not elucidate an obvious mechanism that underpins this phenotype; however, two possible cases can be postulated – delayed neuronal development and/or altered neuronal subtype. In the case of the latter, it is plausible that this subtype may have been derived from the NPC stage. On day 20 of cortical interneuron differentiation, neural precursors are the principal cell type in *DLG2*$^{-/-}$ cell cultures, and while the expression of the CGE marker, COUPTF2, shows a rising trend, there is a significant decrease in the expression of the MGE markers, NKX2.1 and OLIG2. The precise way in which certain subtypes of neurons are produced from neural progenitors is still not fully understood, with some data indicating that NPCs with specific TF expression yield specific neuronal subtypes (Mi et al., 2018). Other studies suggest that the neuronal subtype identity may result from an intrinsic property of the NPCs' when NPCs exit the cell cycle (Brown et al., 2011; Kepecs and Fishell, 2014; Petros et al., 2015; Sultan et al., 2013; Wonders and Anderson, 2006). Therefore, two likely scenarios for the mechanism of the *DLG2*$^{-/-}$ phenotype in determining the identity of neuronal subtypes can be postulated. First, it might change the proportion of particular NPC groups within the culture, or secondly and probably more likely, it might modify the inherent status of NPCs before the onset of neurogenesis.

3.9.7. Summary

The two *DLG2*$^{-/-}$ hESC KO lines, developed earlier using the genome modification technique (CRISPR/Cas9-D10a nickase), were validated and demonstrated to produce no additional CNVs. Immunocytochemistry showed these cells exhibited pluripotency marker expression equivalent to that of WT cells, and mass spectrometry at the protein level, they were complete DLG2 knockouts. The latter analysis additionally provided evidence of DLG2 expression in WT at the NPC phase, suggesting an additional role in development prior to its known role in PSD. Upon DLG2 deficiency, hESCs were able to produce NPCs and interneurons throughout the course of the differentiation process at the same rate as WT. However, the expression of the NPC and interneurons' identities, migration and morphology seemed to be affected by DLG2 deficiency; further studies are required to establish the underlying mechanisms for this phenomenon.

4. Effect of *Dlg2* deficiency on mouse interneuron development *in vivo*

4.1. Introduction

The developmental role of *DLG2* gene has been investigated through phenotyping the DLG2 deficient and WT neuronal progenitors and interneurons derived from hESCs using a cortical interneuron differentiation protocol (see Chapter 3). NKX2.1 and OLIG2 expression showed significant decreases, while an increasing trend in COUPTF2 expression was seen in neural precursors at day 20. In addition, there was a decline trend in GABA (not significant) in maturing cortical interneurons at day 50. A *DLG2*$^{-/-}$ mouse model could be used alternatively or complementarily to validate the *in vitro* identified phenotypes *in vivo* during early embryogenesis and investigate their recovery at later stages or after birth. Murine models are used extensively for research (Rosenthal and Brown, 2007) into the delineation of the processes underlying human cortical neurogenesis (Arber and Li, 2013). In mice, the generation of neocortical interneurons starts at E9.5, reaching its peak between E12 and E15 before finishing at E18.5. (Batista-Brito and Fishell, 2009; Miyoshi et al., 2007; Miyoshi et al., 2010). During embryonic development, interneurons are produced mainly within the MGE and CGE (Hansen et al., 2013) (see Chapter 1).

However, to determine whether the results reported in Chapter 3 were *in vitro* or human cell specific phenotype, a characterisation of *Dlg2* deficiency (homozygous and heterozygous) embryonic mouse brain was conducted. To date, there is no study that focuses on the effect of *Dlg2* deficiency during cortical interneuron development *in vivo* has been reported. Thus, this chapter characterises both *Dlg2* heterozygous and homozygous mouse models at embryonic day 15.5 by using immunohistochemistry involving several markers. The heterozygous *Dlg2*$^{-/+}$ could be closer to the human condition, whereas the homozygous *Dlg2*$^{-/-}$ might offer a better understanding of *Dlg2* function during mouse cortical interneurons development. Therefore, the research presented in this chapter covers two aims. Firstly, characterisation of *Dlg2* KO embryonic mouse brain by using three genotypes (WT, *Dlg2*$^{+/-}$ and *Dlg2*$^{-/-}$), and secondly overlapping DLG2 deficient phenotypes in mouse *in vivo* and human *in vitro* models are investigated.

4.2. Experimental outline

Embryonic mouse brains were collected at E15.5 and genotyped using PCR (see detailed method in Chapter 2). The identities of each brain were not revealed to me until the end of the analysis to avoid experimenter bias.

The embryonic mouse brains were embedded and sectioned in a coronal position at 12 microns thickness. The same interval (10 slides/brain, 16 slices/slide) was used for all the brains in the collection. Next, immunohistochemistry was performed on brain tissues using antibodies for NKX2.1, OLIG2 and COUPTF2. Later, the stained sections were scanned using Zeiss slide scanner. A representative rostral image that was similar across all brains was selected; this process was repeated for medial and caudal sections. Then, a square region of interest with specified height and width in a specific area in the brain (MGE, CGE and where NKX2.1 dispersal can be seen) was selected across all brain sections. In some cases where parts of the region of interest were missing and/or appeared damaged during tissue processing or not clearly stained, the results were not included in the analysis. Counting was performed manually using Image J with the cell counter plugin; positive cells were counted by merging DAPI and the marker of interest. The percentage of positive cells over total cells were analysed. For statistical analysis, a two-way ANOVA with Dunnett's test for multiple comparisons was applied. All data are shown as mean ± SEM and P-values of <0.05 were considered to be statistically significant after multiple comparison corrections. Chapter 2 contains a detailed procedure of the experiments performed for this chapter.

4.3. Effect of *Dlg2* deficiency on cortical interneuron progenitors in embryonic mouse brains

The interneuron precursor markers (NKX2.1 and OLIG2) in four coronal sections – frontal rostral, intermediate rostral, medial and caudal were analysed (Figure 4.1). Medial and two caudal sections were investigated for COUPTF2 and one coronal section (intermediate rostral) was examined for NKX2.1 where dispersal can be seen. A total of 1100-6200 number of cells were analysed. The results shown here are from at least two independent brains/genotypes unless otherwise specified in the figure legends.

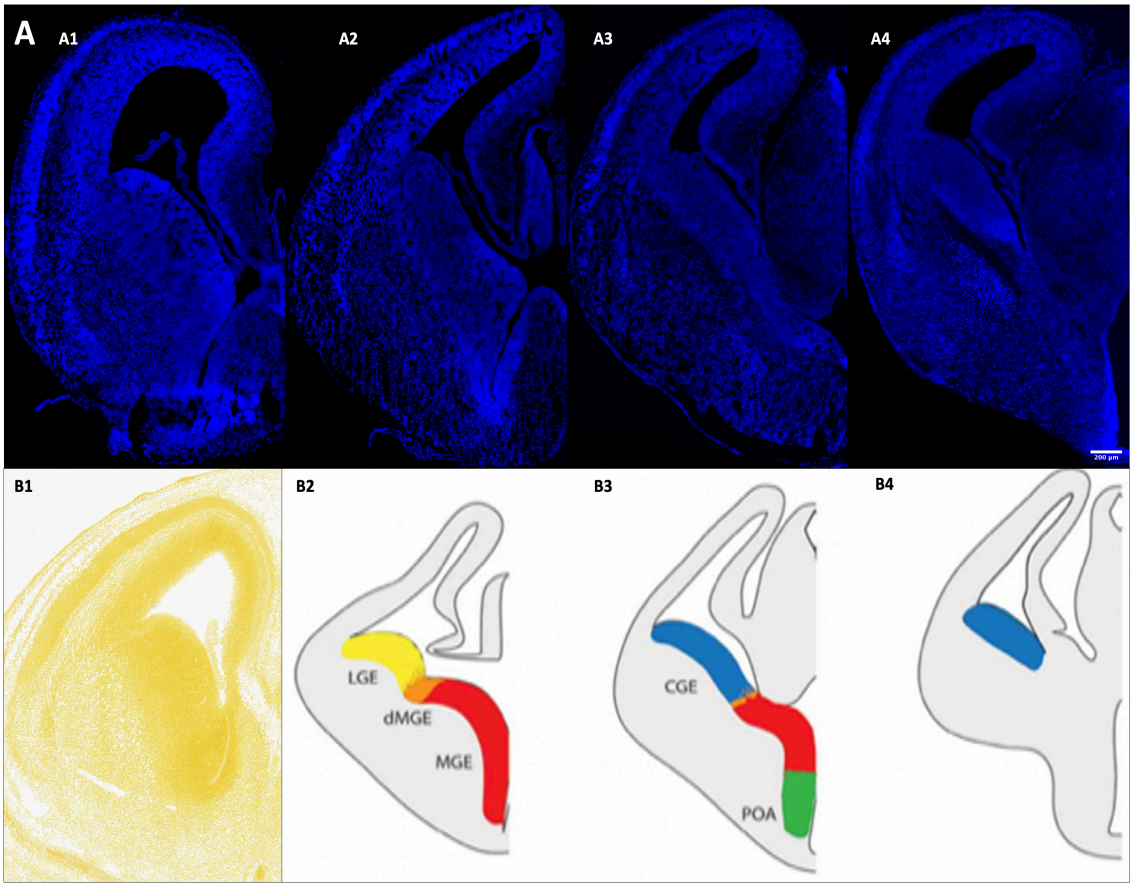


Figure 4.1. Embryonic mouse brain sections from different positions. A) Coronal sections (12 microns thickness) from E15.5 embryonic mouse brain stained with DAPI. A1 is frontal rostral section parallel to the published frontal rostral section in (Wichterle et al., 2001) and in B1, which is from E15.5 mouse brain standard atlas (https://developingmouse.brain-map.org), (A2, A3 and A4) are the comparable sections of B2, B3 and B4 (A2 Intermediate Rostral, A3 Medial and A4 Caudal). B2, 3 &4). Diagram showing the different zones/sections of embryonic mouse brain (B2 Intermediate rostral, B3 Medial, B4 Caudal) (Cauli et al., 2014). Sectioning was performed using Leica Cryostat (CM1860UV). Images were taken using a slide scanner machine (Zeiss Axioscan Z1, ZEISS, Germany) at 20X magnification. Scale bar 200 µm.

NKX2.1 is responsible for the production of the majority of neocortical interneurons in the mouse neocortex. NKX2.1 expression is evident in the mouse MGE area (VZ and SVZ) (Brown et al., 2011; Hernandez-Miranda et al., 2010; Sussel et al., 1999). The MGE zone was selected for NKX2.1^{+} cell counting in the frontal rostral, intermediate rostral and medial sections of the brains (see Figures 4.2-4.4). Furthermore, NKX2.1 is also expressed in the CGE area, which can be considered a caudal extension of the MGE and LGE (Brandao and Romcy-Pereira, 2015). Therefore, the CGE region in the caudal brain sections was also selected for NKX2.1-positive cells counting (see Figure 4.5).

When comparing sections of homozygous (*Dlg2*$^{-/-}$) and heterozygous (*Dlg2*$^{+/-}$) to WT, the gross morphology of the majority of the sections such frontal rostral, intermediate rostral and caudal looked similar; however, the medial section in the KO brain looked different (cortex shape). This could be because the sections picked for KO were slightly more caudal, or they were stretched during the process of when collecting the section. Also,

the slight differences might be attributable to the gene dosage effect (full amount, half amount, and no expression of *Dlg2*). Although the gross structure does look slightly different for medial sections of KOs, the analysis was still carried out.

In Figure 4.2, the pattern of expression in the homozygous KO (*Dlg2*$^{-/-}$) looks different from the WT and heterozygous (*Dlg2*$^{+/-}$) brains. The NKX2.1^{+} cells are distributed in a horizontal pattern, rather than along the ventricular zone (this was seen in one *Dlg2*$^{-/-}$ KO mouse brain section). Moreover, in Figure 4.4, there is a big difference in the distribution of NKX2.1^{+} cells in *Dlg2*$^{-/-}$ KO in this level (this was seen in two *Dlg2*$^{-/-}$ KO mouse brains). In both cases, cells close to the ventricle are not expressing the gene, which may have happened due to a loss of gene expression in cells proximal to the ventricle, or to the abnormal positioning of NKX2.1^{+} cells in the KO. However, it is not clear if this could be attributed to a common KO phenomenon.

The caudal part depicted in Figure 4.6 shows much less NKX2.1 (among all genotype) expression than the rostral and medial sections. This makes sense, as it is known that major expression of NKX2.1 is in the MGE area than the CGE area. The green patches in the caudal part of Figure 4.6 are blood vessels; hence, they were not counted as cells. *Dlg2* deficiency in the frontal rostral, intermediate rostral, medial and caudal brain sections does not prevent the expression of NKX2.1 overall (see Figures 4.2–4.5). No differences in mouse brain were observed between heterozygous (*Dlg2*$^{+/-}$) and WT across all sections (see Figures 4.2–4.6). However, a decreasing trend of NKX2.1 positive cells was observed in the *Dlg2*$^{-/-}$ frontal rostral MGE (mean 47.2% for WT and 32.3% for homozygous) and medial MGE (mean 55.3% for WT and 43.6% for homozygous) sections compared to WT (*Dlg2*$^{+/+}$) (Figures 4.2 and 4.4). For the caudal section, there was no difference between the means of WT (7.4%) and homozygous (5%). The level of brain section had a significant effect on the expression of NKX2.1^{+} cells, while no significant interaction was seen.

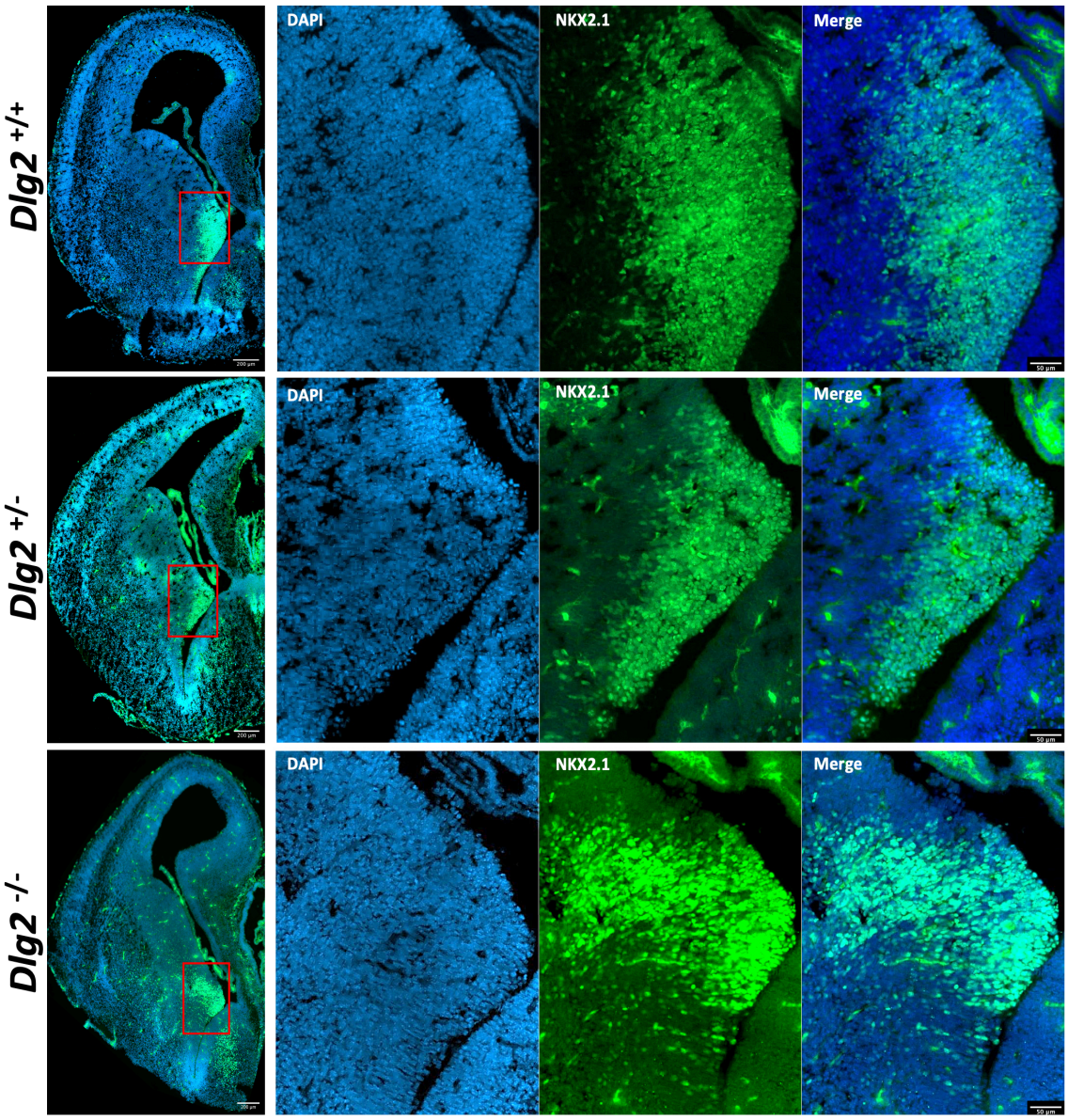


Figure 4.2. Effect of *Dlg2* deficiency on NKX2.1^{+} interneuron progenitors at the frontal rostral part of embryonic mouse brain. Immunohistochemistry images of E15.5 embryonic mouse brains of WT (*Dlg2*$^{+/+}$), heterozygous (*Dlg2*$^{+/-}$) and homozygous (*Dlg2*$^{-/-}$). Comparable embryonic mouse brain coronal sections were stained with DAPI (blue) and NKX2.1 (green). NKX2.1 marks MGE-derived progenitors. Less NKX2.1^{+} cells were observed in the *Dlg2* homozygous (*Dlg2*$^{-/-}$) compared to WT (*Dlg2*$^{+/+}$) mouse brains. Images were taken using a slide scanner (Zeiss Axioscan Z1, ZEISS, Germany) at 20X magnification. The red boxes show the ROI where the counting was performed (the box size is identical across all brains, (width) 540 x (height) 750 pixels). The scale bar for the brain images is 200 µm and for the ROI, it is 50 µm.

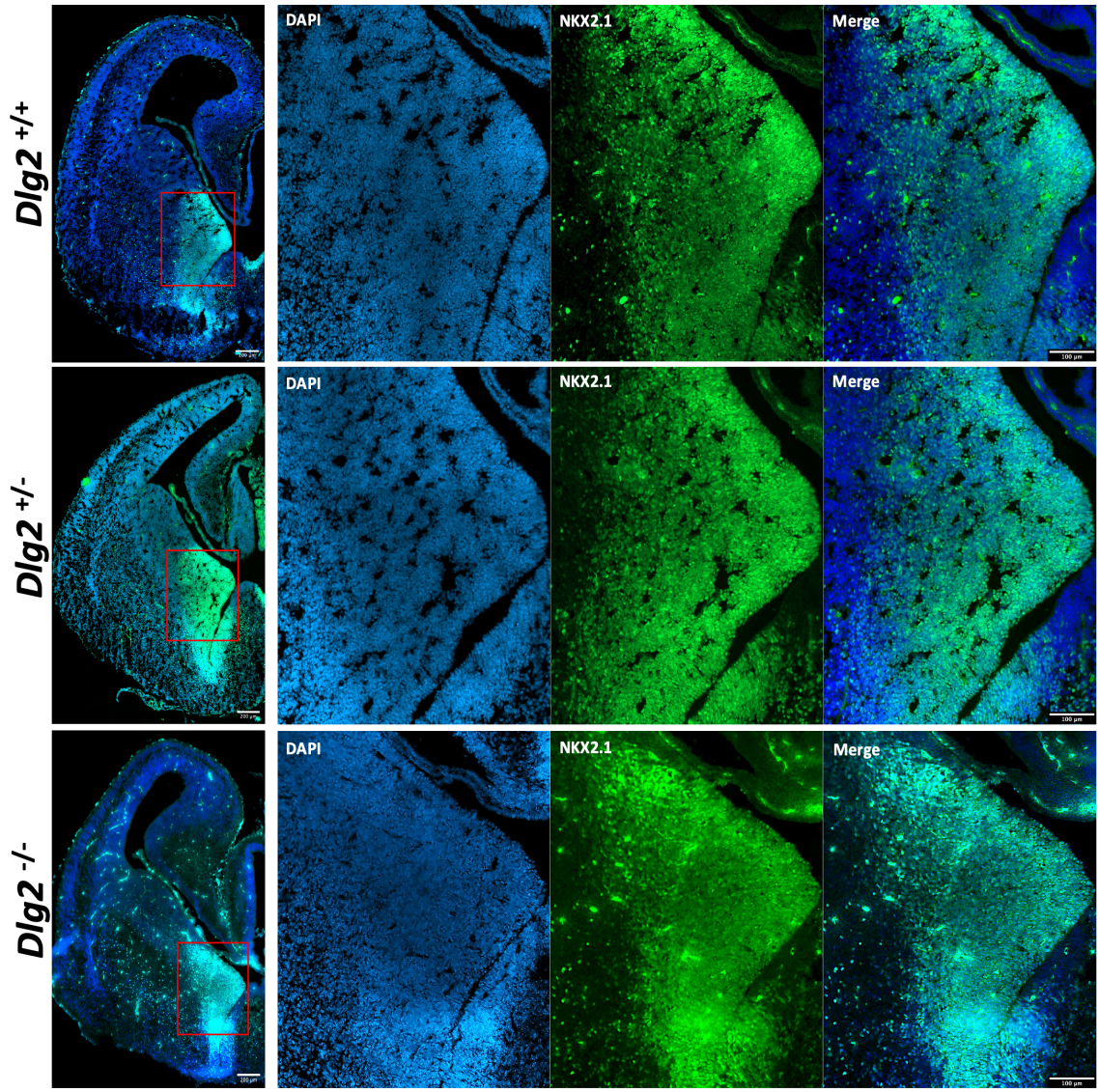


Figure 4.3. Effect of *Dlg2* deficiency on NKX2.1^{+} interneuron progenitors at the intermediate rostral part of embryonic mouse brain. Immunohistochemistry images of E15.5 embryonic mouse brains of WT (*Dlg2*$^{+/+}$), heterozygous (*Dlg2*$^{+/-}$) and homozygous (*Dlg2*$^{-/-}$). Comparable embryonic mouse brain coronal sections were stained with DAPI (blue) and NKX2.1 (green). NKX2.1 marks MGE-derived progenitors. No differences were observed for NKX2.1^{+} cells in the *Dlg2* homozygous (*Dlg2*$^{-/-}$) or heterozygous (*Dlg2*$^{+/-}$) compared to WT (*Dlg2*$^{+/+}$) mouse brains. Images were taken using a slide scanner machine (Zeiss Axioscan Z1, ZEISS, Germany) at 20X magnification. The red boxes show the ROI where the counting was performed (the box size is identical across all brains, (width) 932 x (height) 1085 pixels). The scale bar for the brain images is 200 µm and for the ROI images, it is 100 µm.

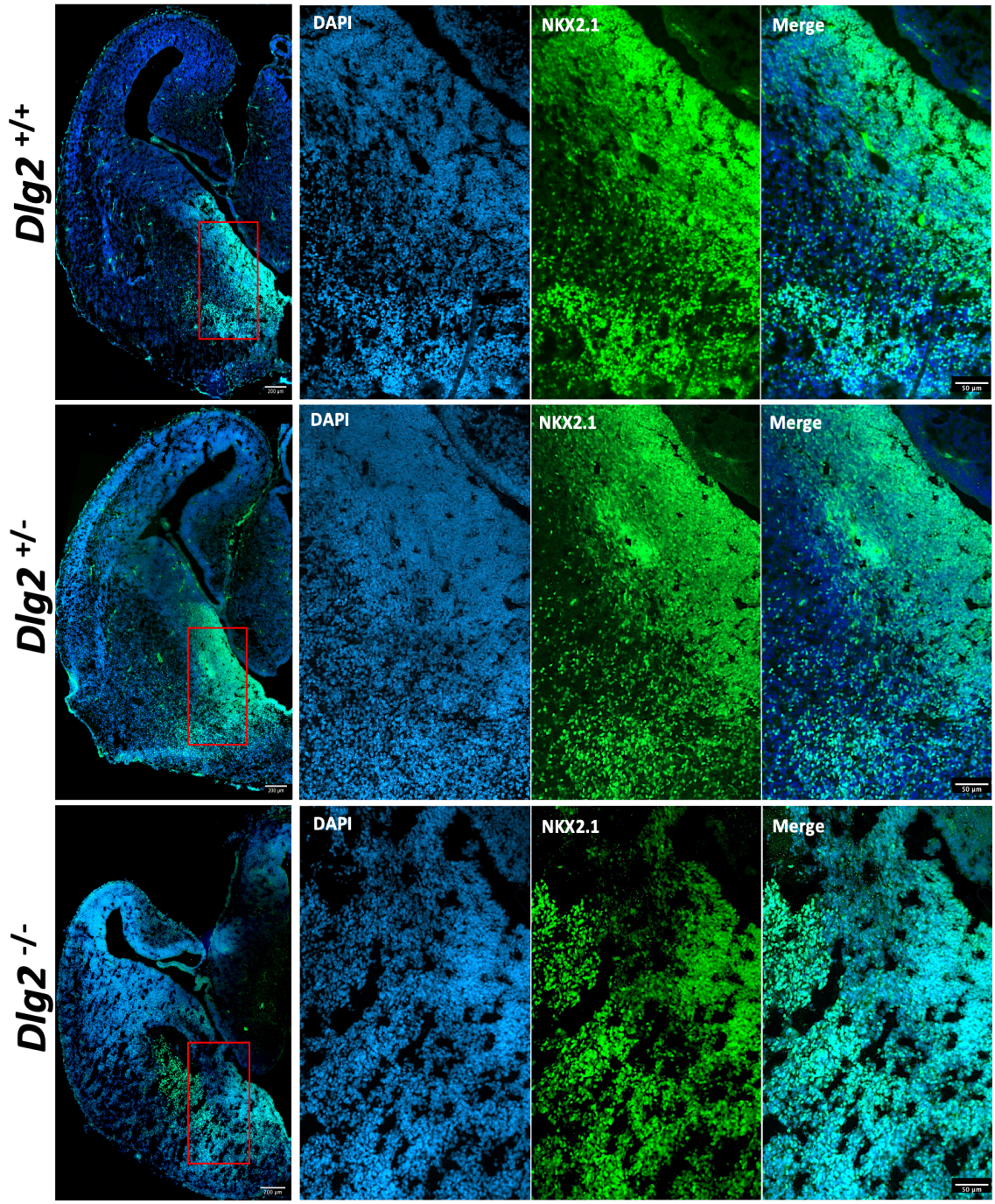


Figure 4.4. Effect of *Dlg2* deficiency on NKX2.1^{+} interneuron progenitors at the medial part of embryonic mouse brain. Immunohistochemistry images of E15.5 embryonic mouse brains of WT (*Dlg2*$^{+/+}$), heterozygous (*Dlg2*$^{+/-}$) and homozygous (*Dlg2*$^{-/-}$). Comparable embryonic mouse brain coronal sections were stained with DAPI (blue) and NKX2.1 (green). NKX2.1 marks MGE-derived progenitors. Fewer NKX2.1^{+} cells were observed in the *DLG2* homozygous (*Dlg2*$^{-/-}$) brains compared to WT (*Dlg2*$^{+/+}$) mouse brains only. Images were taken using a slide scanner machine (Zeiss Axioscan Z1, ZEISS, Germany) at 20X magnification. The red boxes show the ROI where the counting was performed (the box size is identical across all brains, (width) 790 x (height) 1325 pixels). The scale bar for the brain images is 200 µm and for the ROI images, it is 50 µm.

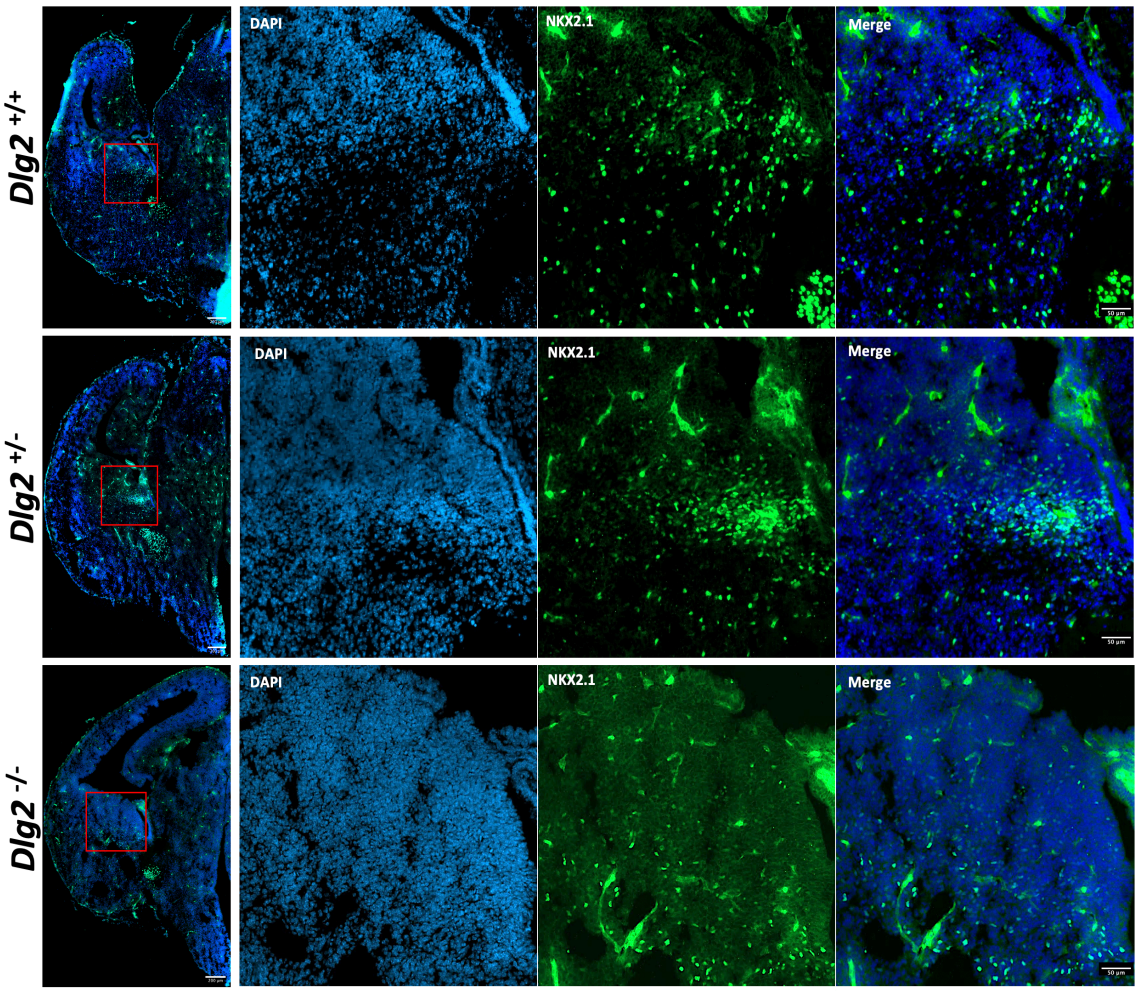


Figure 4.5. Effect of *Dlg2* deficiency on NKX2.1^{+} interneuron progenitors at the caudal part of embryonic mouse brain. Immunohistochemistry images of E15.5 embryonic mouse brains of WT (*Dlg2*$^{+/+}$), heterozygous (*Dlg2*$^{+/-}$) and homozygous (*Dlg2*$^{-/-}$). Comparable embryonic mouse brain coronal sections were stained with DAPI (blue) and NKX2.1 (green). NKX2.1 marks MGE-derived progenitors. No differences were observed for NKX2.1^{+} cells in the *Dlg2* homozygous (*Dlg2*$^{-/-}$) or heterozygous (*Dlg2*$^{+/-}$) compared to WT (*Dlg2*$^{+/+}$) mouse brains. Images were taken using a slide scanner machine (Zeiss Axioscan Z1, ZEISS, Germany) at 20X magnification. The red boxes show the ROI where the counting was performed (the box size is identical across all brains, (width) 895 x (height) 804 pixels). The scale bar for the brain images is 200 µm and for the ROI images, it is 50 µm.

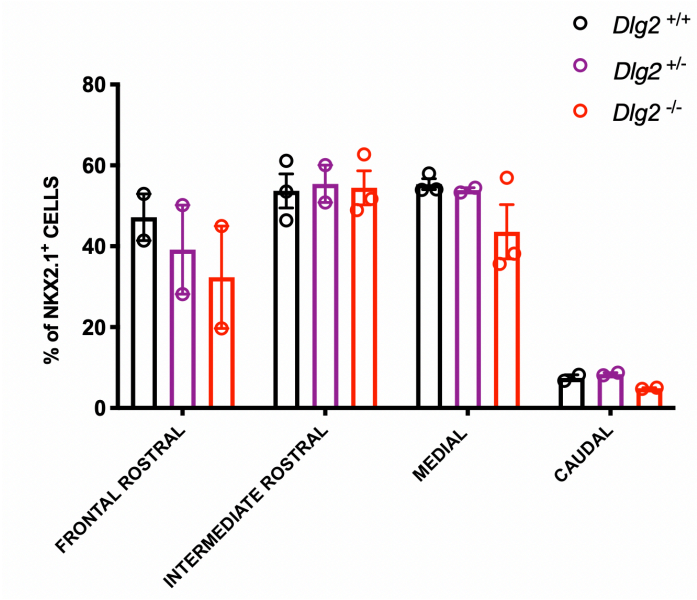


Figure 4.6. Characterisation of NKX2.1^{+} interneuron progenitors at the frontal rostral, intermediate rostral, medial and caudal parts of embryonic mouse brain from *Dlg2*$^{+/+}$, *Dlg2*$^{+/-}$ and *Dlg2*$^{-/-}$ genotypes. Immunohistochemistry on E15.5 embryonic mouse brains of WT (*Dlg2*$^{+/+}$), heterozygous (*Dlg2*$^{+/-}$) and homozygous (*Dlg2*$^{-/-}$) was used to visualise the marker expression. The percentages of positive cells over total cells were analysed using ImageJ (FIJI) and Prism8. According to the two-way ANOVA test, genotype ($F_{2, 16}$ =1.789; P=0.1990; n=3) did not have significant effect, while sections ($F_{3, 16}$=41.58; P<0.0001; n≥2) had significant effects on the % of NKX2.1 positive cells. There were no significant interactions between these two factors (interaction $F_{6, 16}$ =0.5102; P=0.7920). Post hoc analysis showed no significant differences between WT vs *Dlg2*$^{+/-}$ and WT vs *Dlg2*$^{-/-}$ at each brain level. Three WT mouse brains were examined for each intermediate rostral and medial level, except the frontal rostral and caudal were two brains. For *Dlg2*$^{+/-}$, two mouse brains were examined for each intermediate rostral, medial, frontal rostral and caudal level. For *Dlg2*$^{-/-}$, there were two mouse brains for frontal rostral and caudal levels, intermediate rostral and medial levels came from three mouse brains. The bar graph presents the mean ± SEM. Less than p<0.05 was considered statically significant.

OLIG2 is highly expressed in the *Nkx2.1*-positive domain of MGE and at a lower level in the CGE region (Hernandez-Miranda et al., 2010; Miyoshi et al., 2007). Thus, the MGE zone in the frontal rostral, intermediate rostral and medial brain sections (Figures 4.7-4.9) and the CGE region of the caudal brain (Figure 4.10) were selected for OLIG2^{+} cell counting.

Expression of OLIG2 in progenitor cells in MGE and CGE was investigated. Figures 4.8 and 4.9, show the area of the staining patterns in ventricular zone to be smaller than seen with NKX2.1 (Figures 4.3 and 4.4). Moreover, when considering staining patterns in caudal sections (Figure 4.10), it is obvious that OLIG2 is expressed in a larger area than NKX2.1 (Figure 4.5). This means that there are OLG2^{+} and NKX2.1^{-} cells in the caudal brain. These cells could be cholinergic neuron precursors (Furusho et al., 2006), as they were reported to originate from MGE (Olsson et al., 1998); alternatively, perhaps not every MGE precursor expresses NKX2.1. Or they could be oligodendrocyte progenitor cells, as they were expressed in the ventral telencephalon MGE (Nery et al., 2001; Tekki-Kessaris et al., 2001).

The quantification of OLIG2^{+} cells was carried out in an ROI in each section, as outlined in Figures 4.7–4.10. Afterwards, OLIG2^{+} cells were counted manually by merging DAPI with OLIG2 staining. Analysis of the percentage of positive cells over total cells were performed. *Dlg2* deficiency did not affect expression of OLIG2 in both heterozygous (*Dlg2*$^{+/-}$) and homozygous (*Dlg2*$^{-/-}$) mouse brains. No differences were observed in frontal rostral, intermediate rostral, medial and caudal OLIG2^{+} cells in all three genotypes (Figures 4.7–4.11). Genotype had no significant effect upon the expression of OLIG2^{+} cells; however, brain section level significantly affected the expression of OLIG2, although no significant interaction between the genotype and the level of brain section was seen. Similar to NKX2.1^{+}, OLIG2^{+} expression is more prevalent in rostral sections than caudal sections (Figure 4.11).

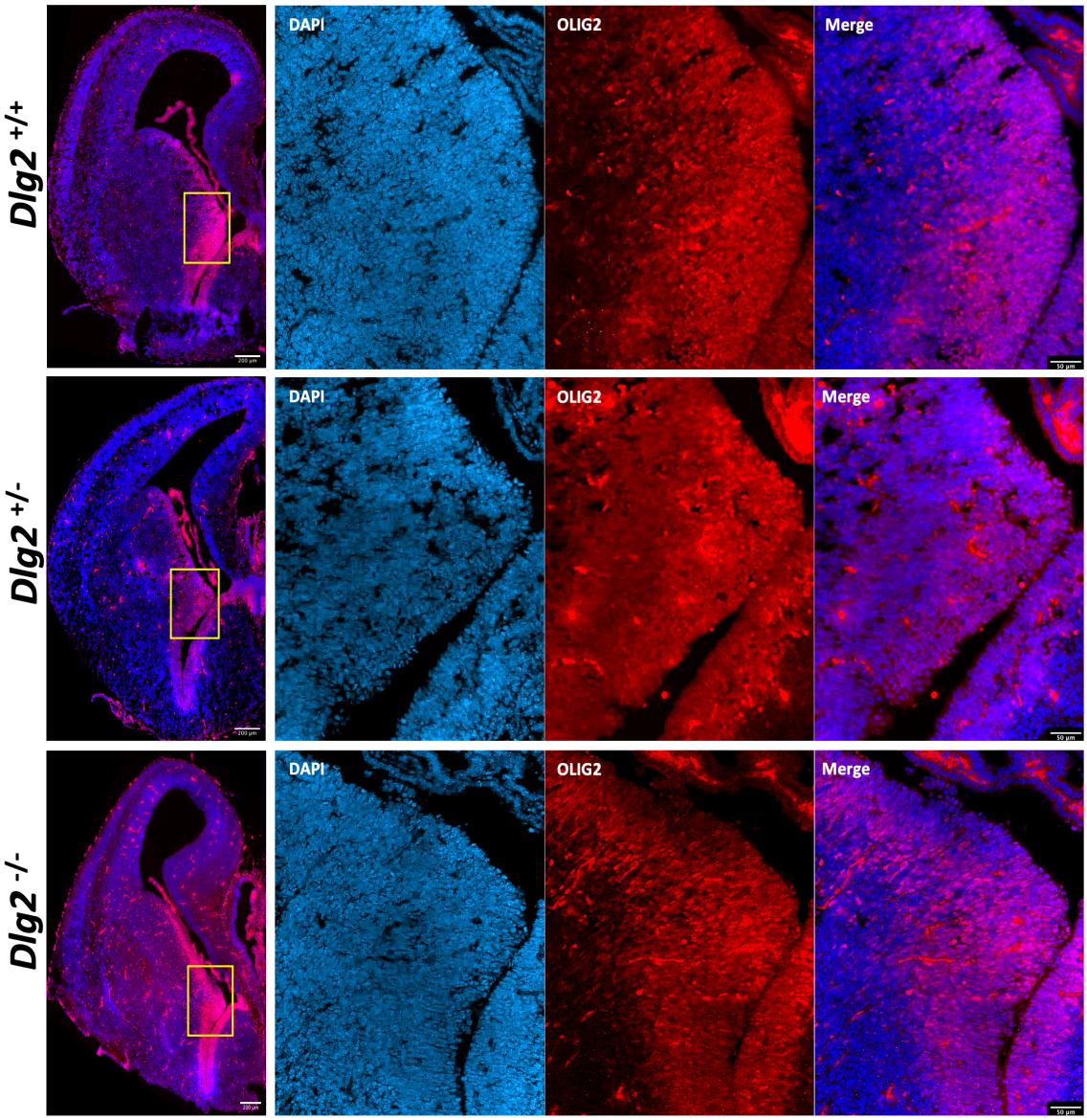


Figure 4.7. Effect of *Dlg2* deficiency on OLIG2^{+} interneuron progenitors at the frontal rostral part of embryonic mouse brain. Immunohistochemistry images of E15.5 embryonic mouse brains of WT (*Dlg2*$^{+/+}$), heterozygous (*Dlg2*$^{+/-}$) and homozygous (*Dlg2*$^{-/-}$). Comparable embryonic mouse brain coronal sections were stained with DAPI (blue) and OLIG2 (red). OLIG2 is expressed in a subpopulation of MGE-derived progenitors. No differences were observed for OLIG2^{+} cells in the *Dlg2* homozygous (*Dlg2*$^{-/-}$) or heterozygous (*Dlg2*$^{+/-}$) compared to WT (*Dlg2*$^{+/+}$) mouse brains. Images were taken using a slide scanner machine (Zeiss Axioscan Z1, ZEISS, Germany) at 20X magnification. The yellow boxes show the ROI where the counting was performed (the box size is identical across all brains, (width) 540 x (height) 750 pixels). The scale bar for the brain images is 200 µm and for the ROI images, it is 50 µm.

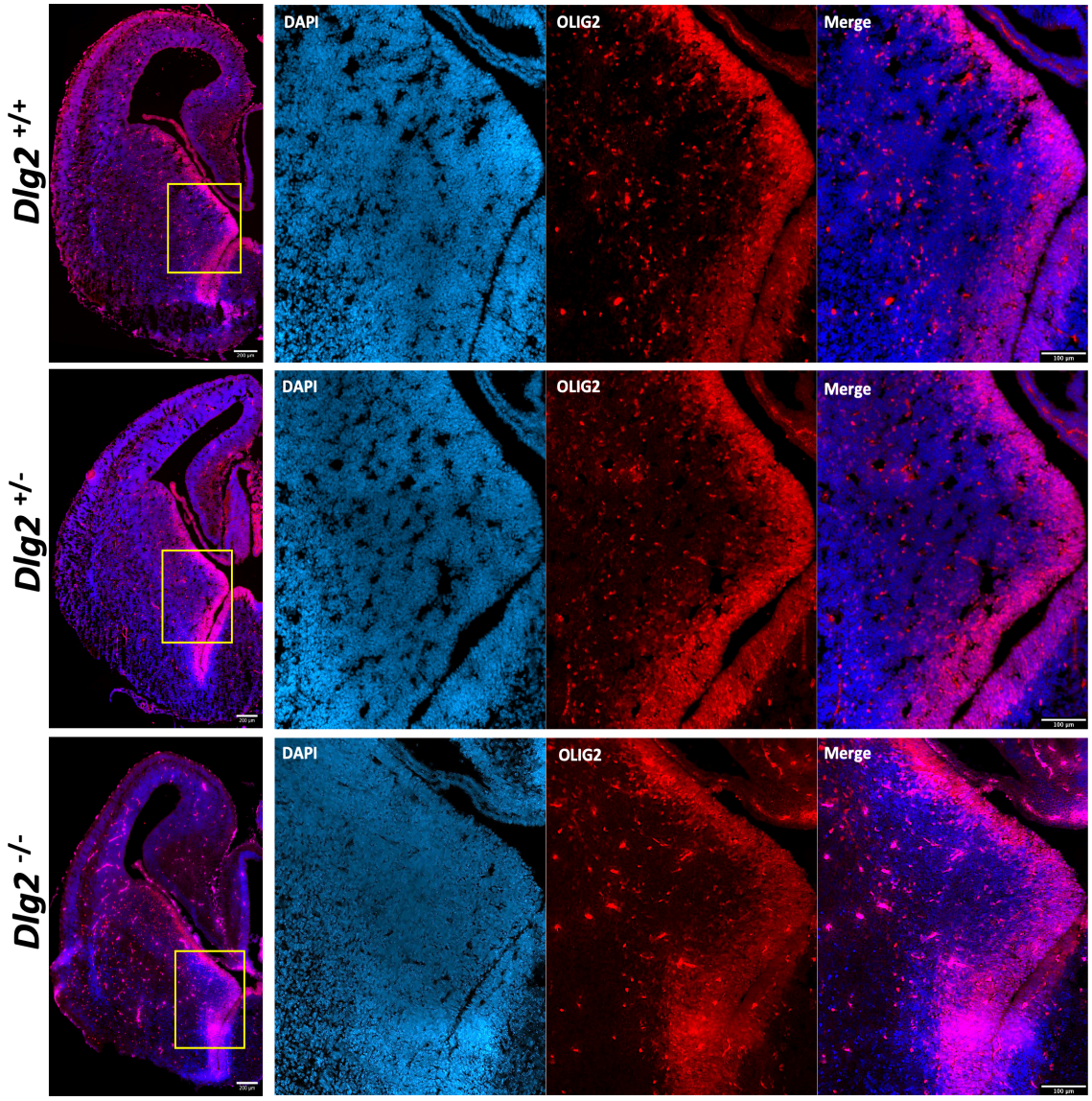


Figure 4.8. Effect of *Dlg2* deficiency on OLIG2^{+} interneuron progenitors at the intermediate rostral part of embryonic mouse brain. Immunohistochemistry images of E15.5 embryonic mouse brains WT (*Dlg2*$^{+/+}$), heterozygous (*Dlg2*$^{+/-}$) and homozygous (*Dlg2*$^{-/-}$). Comparable embryonic mouse brain coronal sections were stained with DAPI (blue) and OLIG2 (red). OLIG2 is expressed in a subpopulation of MGE-derived progenitors. No differences were observed for OLIG2^{+} cells in the *Dlg2* homozygous (*Dlg2*$^{-/-}$) or heterozygous (*Dlg2*$^{+/-}$) compared to WT (*Dlg2*$^{+/+}$) mouse brains. Images were taken using a slide scanner machine (Zeiss Axioscan Z1, ZEISS, Germany) at 20X magnification. The yellow boxes show the ROI where the counting was performed (the box size is identical across all brains, (width) 932 x (height) 1085 pixels). The scale bar for the brain images is 200 µm and for the ROI images, it is 100 µm.

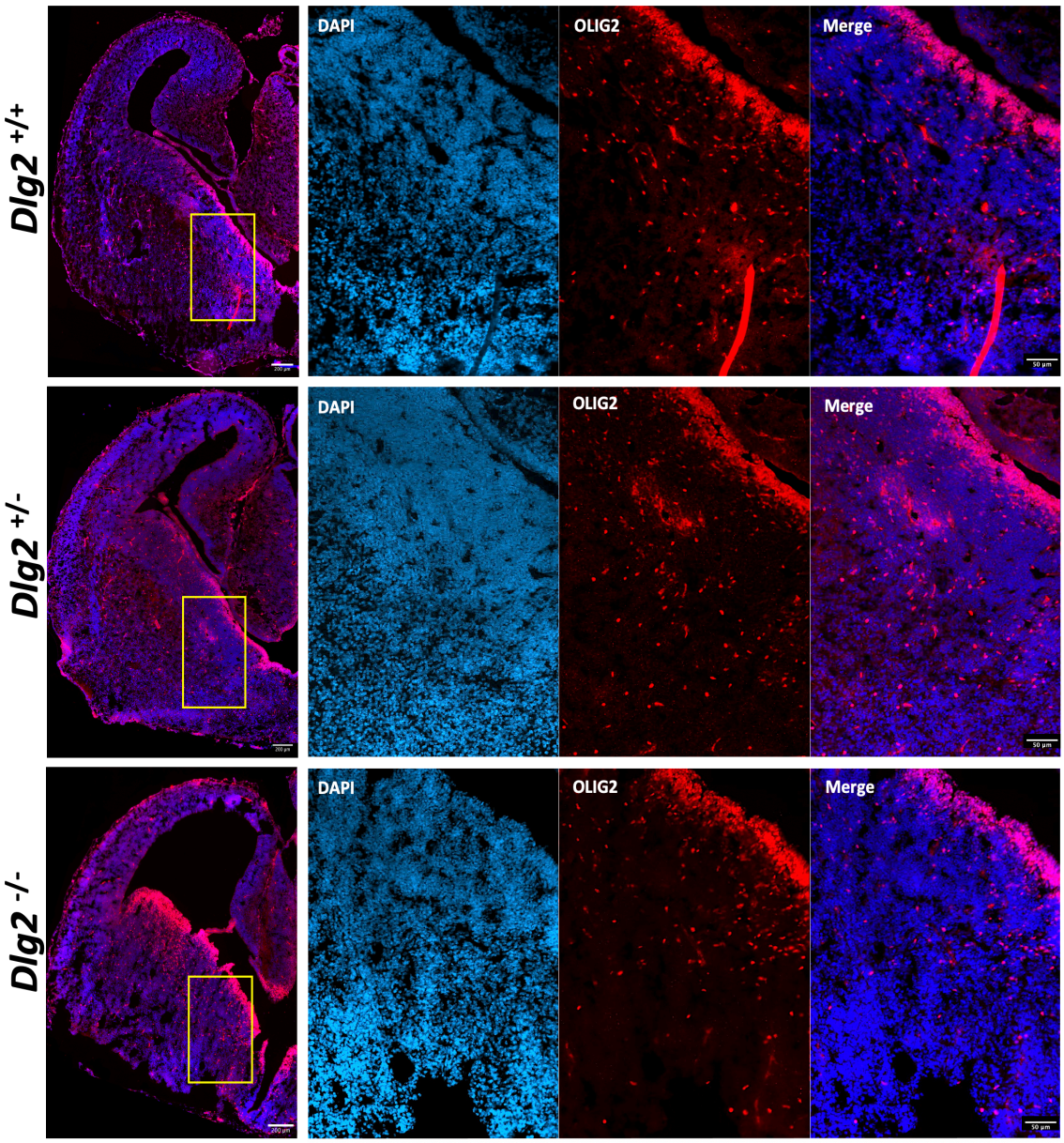


Figure 4.9. Effect of *Dlg2* deficiency on OLIG2^{+} interneuron progenitors at the medial part of embryonic mouse brain. Immunohistochemistry images of E15.5 embryonic mouse brains WT (*Dlg2*$^{+/+}$), heterozygous (*Dlg2*$^{+/-}$) and homozygous (*Dlg2*$^{-/-}$). Comparable embryonic mouse brain coronal sections were stained with DAPI (blue) and OLIG2 (red). OLIG2 is expressed in a subpopulation of MGE-derived progenitors. No differences were observed for OLIG2^{+} cells in the *Dlg2* homozygous (*Dlg2*$^{-/-}$) or heterozygous (*Dlg2*$^{+/-}$) compared to WT (*Dlg2*$^{+/+}$) mouse brains. Images were taken using a slide scanner machine (Zeiss Axioscan Z1, ZEISS, Germany) at 20X magnification. The yellow boxes show the ROI where the counting was performed (the box size is identical across all brains, (width) 790 x (height) 1325 pixels). The scale bar for the brain images is 200 µm and for the ROI images, it is 50 µm.

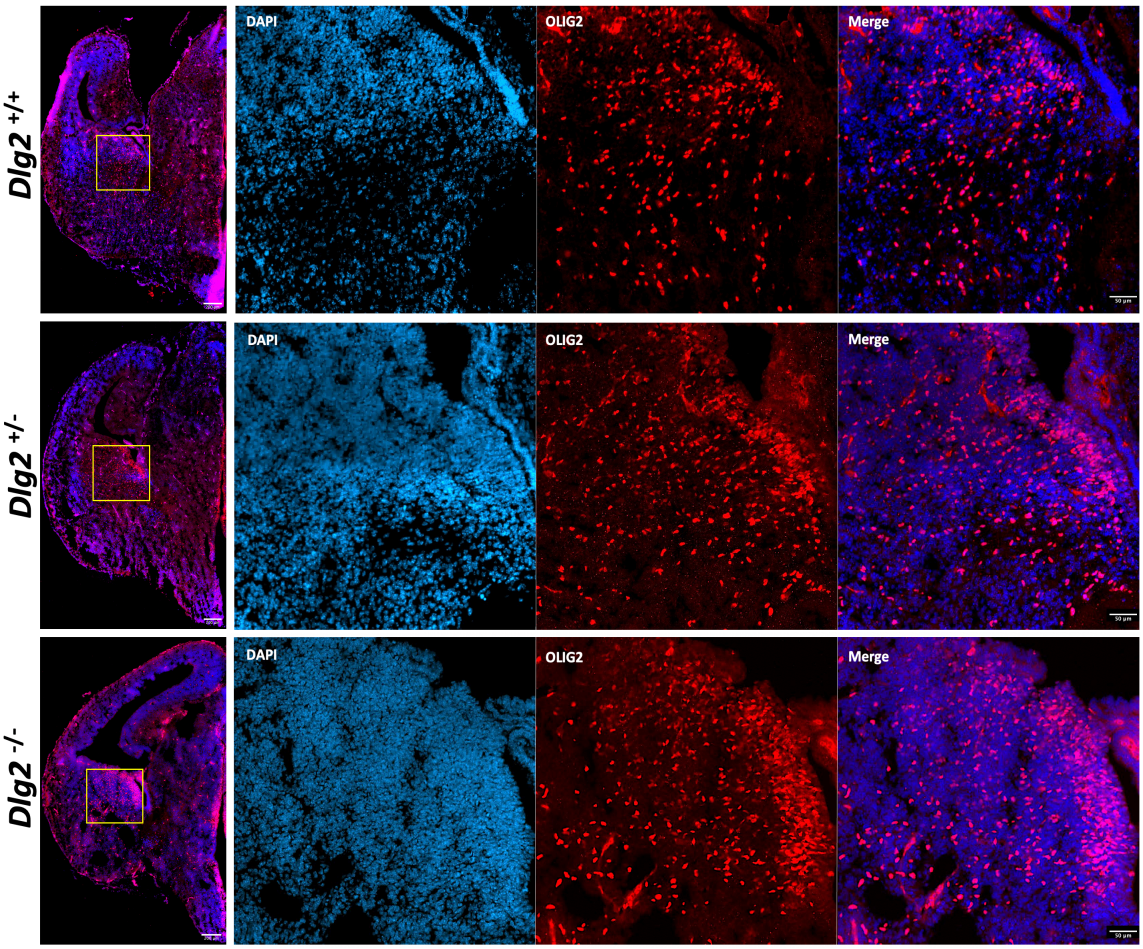


Figure 4.10. Effect of *Dlg2* deficiency on OLIG2$^+$ interneuron progenitors at the caudal part of embryonic mouse brain. Immunohistochemistry images of E15.5 embryonic mouse brains WT (*Dlg2*$^{+/+}$), heterozygous (*Dlg2*$^{+/-}$) and homozygous (*Dlg2*$^{-/-}$). Comparable embryonic mouse brain coronal sections were stained with DAPI (blue) and OLIG2 (red). OLIG2 is expressed in a subpopulation of MGE-derived progenitors. No differences were observed for OLIG2$^+$ cells in the *Dlg2* homozygous (*Dlg2*$^{-/-}$) or heterozygous (*Dlg2*$^{+/-}$) compared to WT (*Dlg2*$^{+/+}$) mouse brains. Images were taken using a slide scanner machine (Zeiss Axioscan Z1, ZEISS, Germany) at 20X magnification. The yellow boxes show the ROI where the counting was performed (the box size is identical across all brains, (width) 895 x (height) 804 pixels). The scale bar for the brain images is 200 µm and for the ROI images, it is 50 µm.

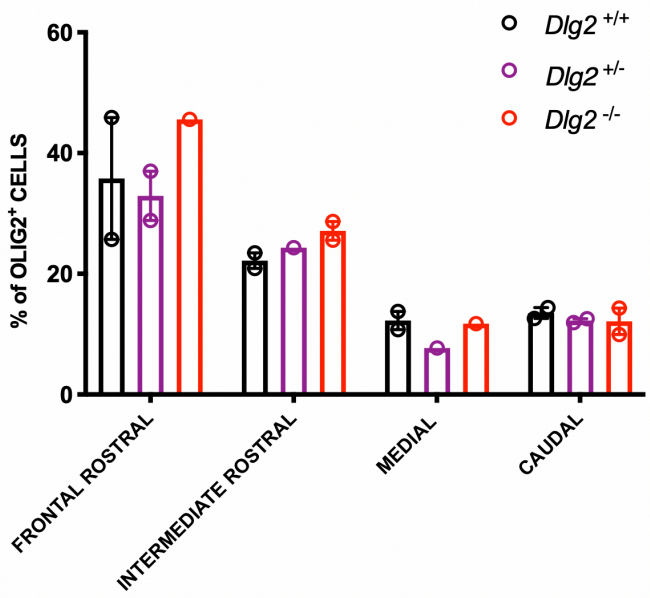


Figure 4.11. Characterisation of OLIG2^{+} interneuron progenitors at the frontal rostral, intermediate rostral, medial and caudal parts of embryonic mouse brain from *Dlg2*$^{+/+}$, *Dlg2*$^{+/-}$ and *Dlg2*$^{-/-}$ genotypes. Immunohistochemistry on E15.5 embryonic mouse brains of WT (*Dlg2*$^{+/+}$), heterozygous (*Dlg2*$^{+/-}$) and homozygous (*Dlg2*$^{-/-}$) was used to visualise the marker expression. The percentages of positive cells over total cells were analysed using ImageJ (FIJI) and Prism8. $p<0.05$ was considered statistically significant.

CGE and MGE regions were chosen for counting COUPTF2-positive cells. Reports indicate that COUPTF2 regulates the generation of CGE-derived interneurons and expressed in both the CGE and MGE areas (Cai et al., 2013; Hu et al., 2017), with higher expression in CGE (Kanatani et al., 2008). A medial and two caudal sections were stained for COUPTF2 (Figures 4.13-4.15).

Most of the sections look similar; however, KO medial section looked slightly different. This could be due to the section being stretched during collection. Also, the slight variances observed might be attributable to the gene dosage effect (full amount, no expression of *Dlg2*). In general, there were some dissimilarities in the gross morphology of KO brains; however, they were similar enough to be compared.

In general, the expression of COPUTF2^{+} cells in medial sections (Figure 4.13) was less than in caudal sections (Figures 4.13 & 4.14). This was expected, as it is known that COUPTF2 is expressed more strongly in the CGE than MGE. From the images of medial and caudal-1 (Figures 4.12–4.13), there is no difference in the expression pattern of COUPTF2^{+} cells among all genotypes. Nevertheless, caudal-2 (Figure 4.14) showed a lower percentage of COUPTF2^{+} in heterozygous and homozygous brain sections compared to WT (*Dlg2*$^{+/+}$).

Quantification was performed in the ROI of the MGE area in the medial brain section (yellow box in Figure 4.12) and the CGE area in the two caudal sections (yellow box in Figures 4.13 and 4.14). The level of brain section had no significant effect upon the expression of COUPTF2 with no significant interaction. At medial and caudal-1 brain sections, no differences were observed among all three genotypes (see Figure 4.12–4.14 & 4.15). However, a declining trend was observed in the caudal-2 CGE sections of COUPTF2^{+} cells in heterozygous *Dlg2*$^{+/-}$ (17.4%) and homozygous *Dlg2*$^{-/-}$ (16.6%) mouse brains compared to WT *Dlg2*$^{+/+}$ (33.3%) (Figures 4.14 and 4.15).

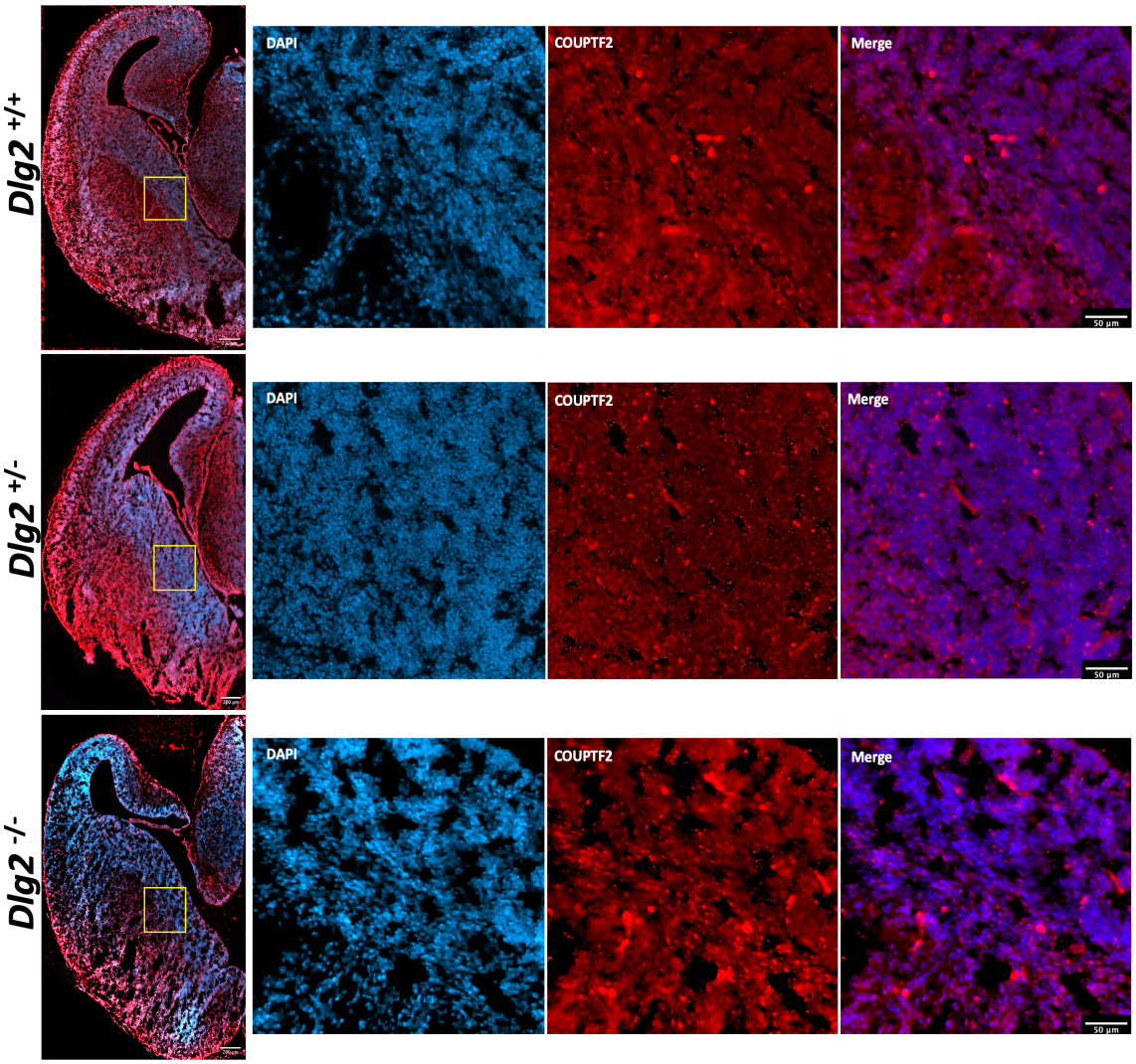


Figure 4.12. Effect of *Dlg2* deficiency on COUPTF2^{+} interneuron progenitors at the medial part of embryonic mouse brain. Immunohistochemistry images of E15.5 embryonic mouse brains of WT (*Dlg2*$^{+/+}$), heterozygous (*Dlg2*$^{+/-}$) and homozygous (*Dlg2*$^{-/-}$). Comparable embryonic mouse brain coronal sections were stained with DAPI (blue) and COUPTF2 (red). COUPTF2 is expressed in CGE-derived progenitors. No differences were observed for COUPTF2^{+} cells in the *DLG2* homozygous (*Dlg2*$^{-/-}$) or heterozygous (*Dlg2*$^{+/-}$) compared to WT (*Dlg2*$^{+/+}$). Images were taken using a slide scanner machine (Zeiss Axioscan Z1, ZEISS, Germany) at 20X magnification. The yellow boxes show the ROI where the counting was performed (the box size is identical across all brains, (width) 601x (height) 628 pixels). The scale bar for the brain images is 200 µm and for the ROI images, it is 50 µm.

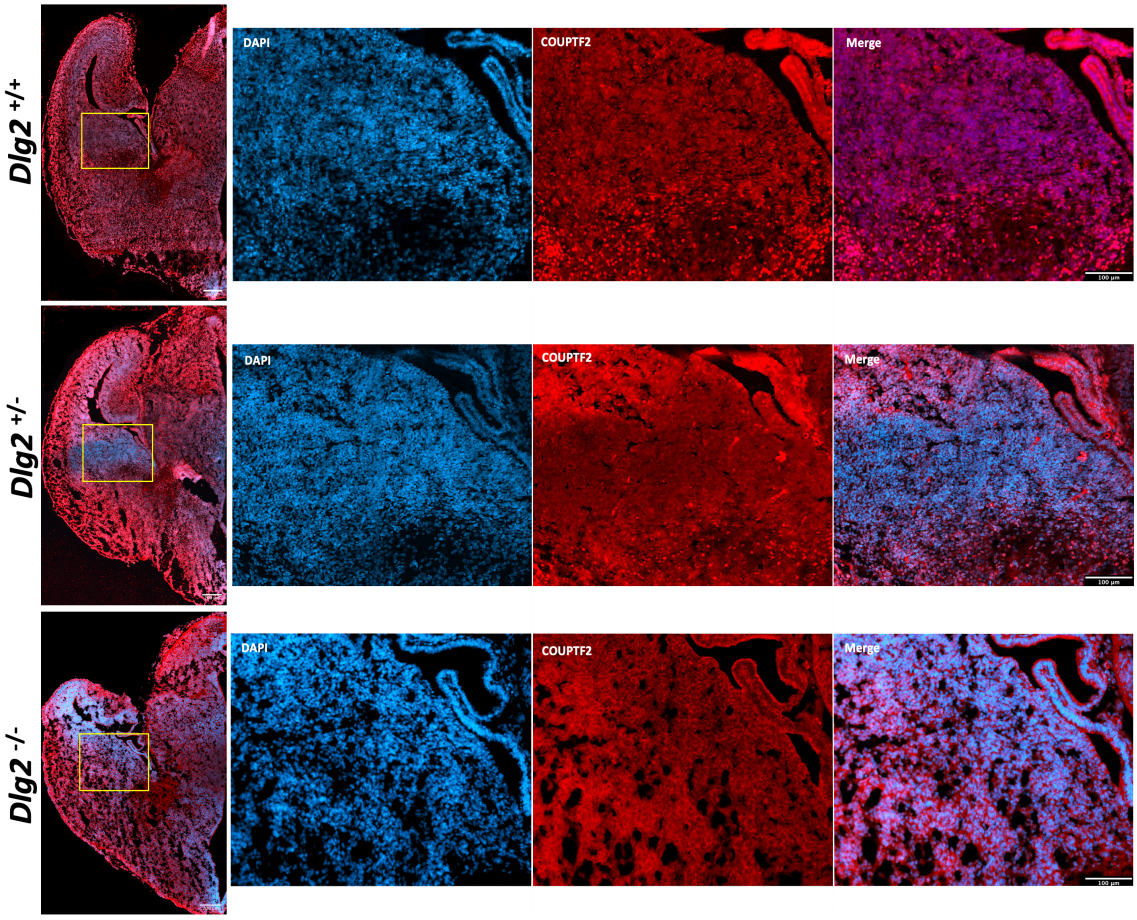


Figure 4.13. Effect of *Dlg2* deficiency on COUPTF2^{+} interneuron progenitors at the caudal-1 level of embryonic mouse brain. Immunohistochemistry images of E15.5 embryonic mouse brains WT (*Dlg2*$^{+/+}$), heterozygous (*Dlg2*$^{+/-}$) and homozygous (*Dlg2*$^{-/-}$). Comparable embryonic mouse brain coronal sections were stained with DAPI (blue) and COUPTF2 (red). COUPTF2 is expressed in CGE-derived progenitors. No differences were observed for COUPTF2^{+} cells in the *DLG2* homozygous (*Dlg2*$^{-/-}$) or heterozygous (*Dlg2*$^{+/-}$) compared to WT (*Dlg2*$^{+/+}$). Images were taken using a slide scanner machine (Zeiss Axioscan Z1, ZEISS, Germany) at 20X magnification. The yellow boxes show the ROI where the counting was performed (the box size is identical across all brains, (width) 985 x (height) 816 pixels). The scale bar for the brain images is 200 µm and for the ROI images, it is 100 µm.

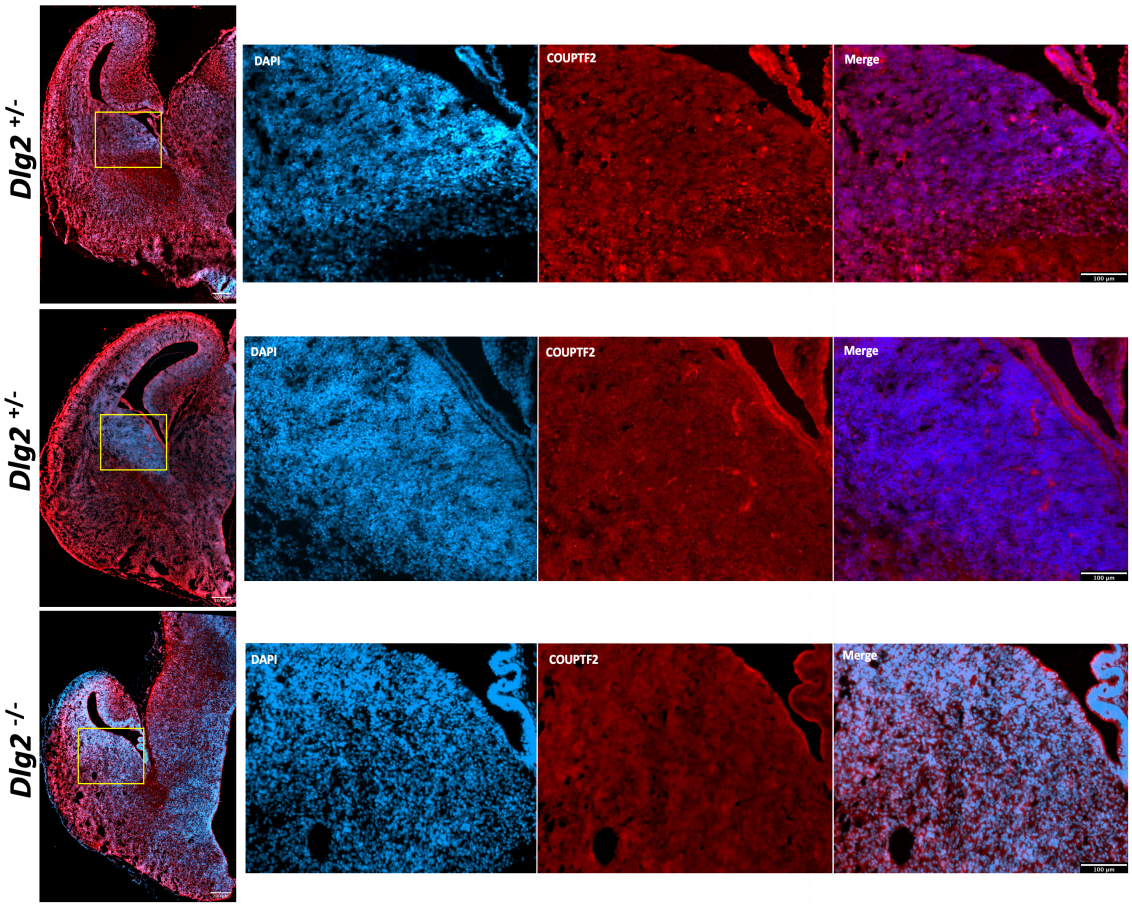


Figure 4.14. Effect of *Dlg2* deficiency on COUPTF2^{+} interneuron progenitors at the caudal-2 level of embryonic mouse brain. Immunohistochemistry images of E15.5 embryonic mouse brains WT (*Dlg2*$^{+/+}$), heterozygous (*Dlg2*$^{+/-}$) and homozygous (*Dlg2*$^{-/-}$). Comparable embryonic mouse brain coronal sections were stained with DAPI (blue) and COUPTF2 (red). COUPTF2 is expressed in CGE-derived progenitors. Fewer COUPTF2^{+} cells were observed in the *Dlg2* homozygous (*Dlg2*$^{-/-}$) and heterozygous (*Dlg2*$^{+/-}$) compared to WT (*Dlg2*$^{+/+}$) mouse brains. Images were taken using a slide scanner machine (Zeiss Axioscan Z1, ZEISS, Germany) at 20X magnification. The yellow boxes show the ROI where the counting was performed (the box size is identical across all brains, (width) 985 x (height) 816 pixels).The scale bar for the brain images is 200 µm and for the ROI images, it is 100 µm.

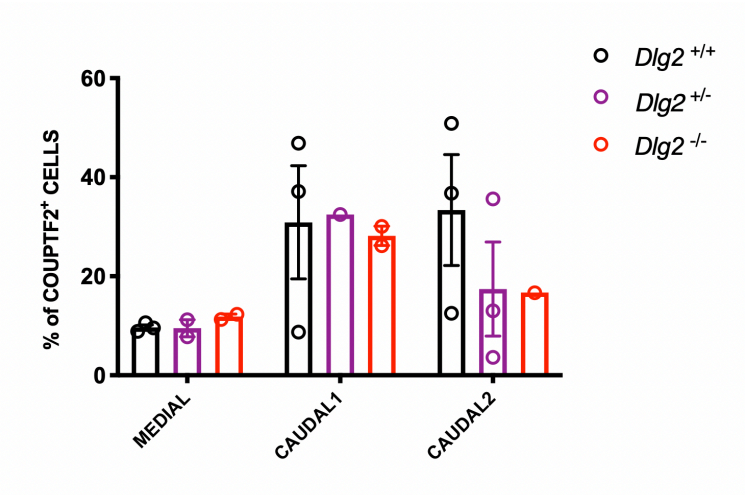


Figure 4.15. Characterisation of COUPTF2^{+} interneuron progenitors at medial and two caudal levels of embryonic mouse brain from *Dlg2*$^{+/+}$, *Dlg2*$^{+/-}$ and *Dlg2*$^{-/-}$ genotypes. Immunohistochemistry on E15.5 embryonic mouse brains of WT (*Dlg2*$^{+/+}$), heterozygous (*Dlg2*$^{+/-}$) and homozygous (*Dlg2*$^{-/-}$) was used to visualise the marker expression. The percentages of positive cells over total cells were analysed using ImageJ (FIJI) and Prism8. According to the two-

way ANOVA test, genotype ($F_{2, 11}$ =0.3443; P=0.7161; n=3) and brain level sections ($F_{2, 11}$=3.164; P=0.0822; n≥1) had no significant effect on the percentage of COUPTF2 positive cells. There was no significant interaction between these two factors (interaction $F_{4, 11}$=0.3966; P=0.8071). Post hoc analysis showed no significant differences between WT vs *Dlg2*$^{+/-}$ and WT vs *Dlg2*$^{-/-}$ in each brain level. Three WT mouse brains were examined for each medial and two-caudal levels. For *Dlg2*$^{+/-}$, two mouse brains for medial level, one brain for caudal1, and three mouse brains for caudal2. For *Dlg2*$^{-/-}$, there were two mouse brains for each medial and caudal1 level, and one brain for caudal2 level. The bar graph presents the mean ± SEM.

4.4. No effect of *Dlg2* deficiency on migrating cells in embryonic mouse brains

In Chapter 3, the role of *DLG2* gene has been investigated by using a cortical interneuron differentiation protocol to characterise WT and *DLG2* deficient cortical interneurons that had differentiated from hESCs. At day 50, a decreasing trend in the expression of GABA in maturing cortical interneurons from *DLG2*$^{-/-}$ hESCs was observed. Thus, a characterisation of *Dlg2* KO embryonic mouse brain was conducted to delineate whether defects in the expression of interneuron markers identified *in vitro* correlate with those observed *in vivo*.

The phenotypes of *Dlg2*$^{+/+}$*, Dlg2*$^{+/-}$ *and Dlg2*$^{-/-}$ embryonic mouse brains (E15.5) (n=4/genotype) were characterised through immunohistochemistry using a range of interneuron markers (SST, PV, NEUN, GAD67 and CALBINDIN). During the staining validation, expression of SST and PV (images not shown) was not detected, even with antigen retrieval treatment. GAD67 did not show a specific binding as well as CALBINDIN and NEUN. There was an issue with the secondary antibody for the latter two markers, as it detected mouse endogenous IgG; therefore, these were not included in this thesis.

NKX2.1 was counted where dispersal can be seen in Figure 4.3 (see Figure 4.16). No significant difference was observed on migrating NKX2.1^{+} cells (see Figure 4.17). However, these expressing NKX2.1 cells are migrating to striatum, and it is known that continued expression of *Nkx2.1* results in striatal interneuron development, whereas discontinued expression results in the neurons following the cortical interneuron fate (Butt et al., 2008; Nobrega-Pereira et al., 2008). The focus of this study is not quantifying striatal interneurons as the continues expression of NKX2.1 during migration is a sign of striatal neurons.

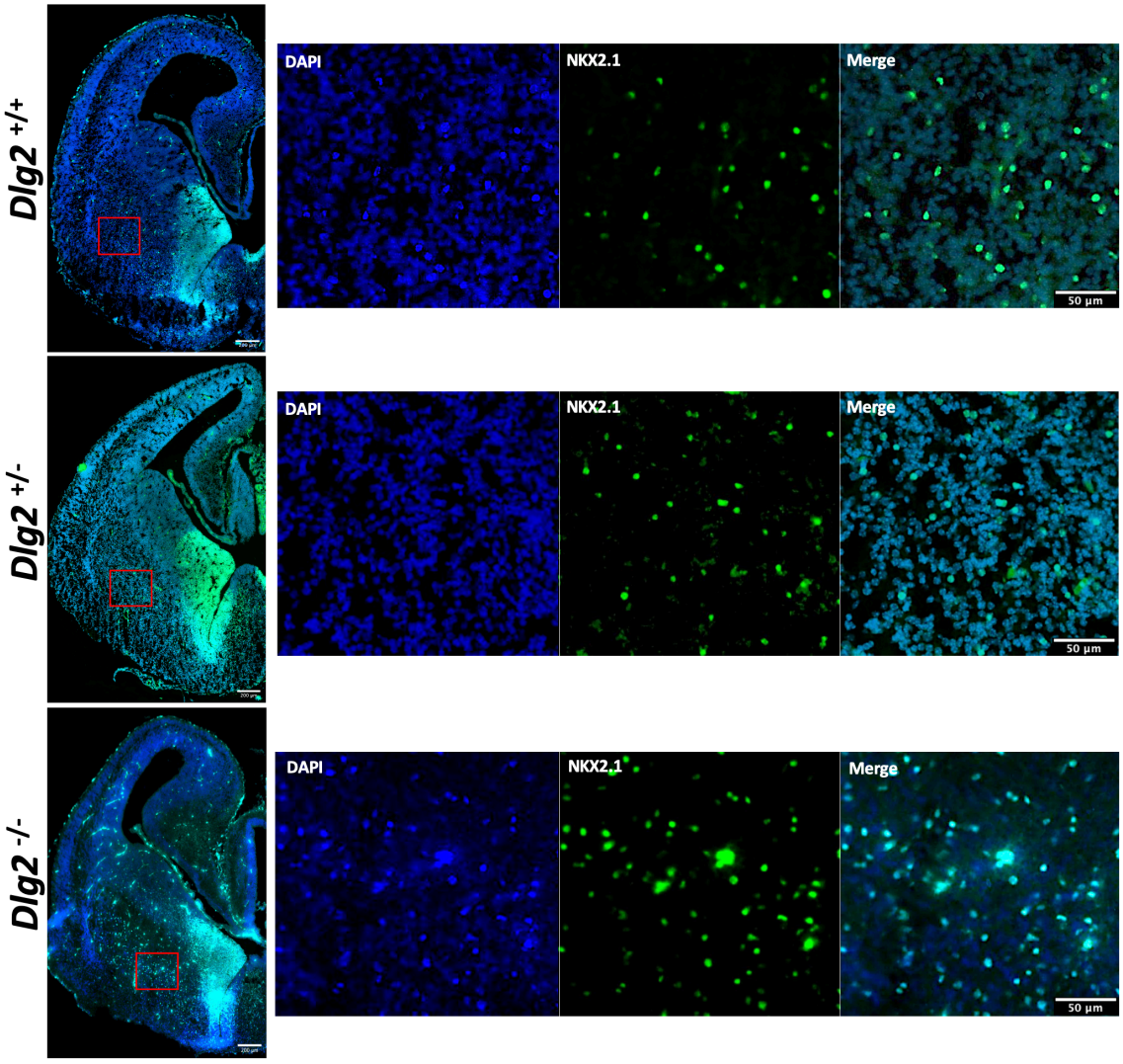


Figure 4.16. Effect of *Dlg2* deficiency on NKX2.1^{+} migrating cells at the intermediate rostral part of embryonic mouse brain. Immunohistochemistry images of E15.5 embryonic mouse brains of WT (*Dlg2*$^{+/+}$), heterozygous (*Dlg2*$^{+/-}$) and homozygous (*Dlg2*$^{-/-}$). Comparable embryonic mouse brain coronal sections were stained with DAPI (blue) and NKX2.1 (green). NKX2.1 marks MGE-derived progenitors. No differences were observed for NKX2.1^{+} cells in the *Dlg2* homozygous (*Dlg2*$^{-/-}$) or heterozygous (*Dlg2*$^{+/-}$) compared to WT (*Dlg2*$^{+/+}$) mouse brains. Images were taken using a slide scanner machine (Zeiss Axioscan Z1, ZEISS, Germany) at 20X magnification. The red boxes show the ROI where the counting was performed (the box size is identical across all brains, (width) 448 x (height) 424 pixels). The scale bar for the brain images is 200 µm and for the ROI images, it is 50 µm.

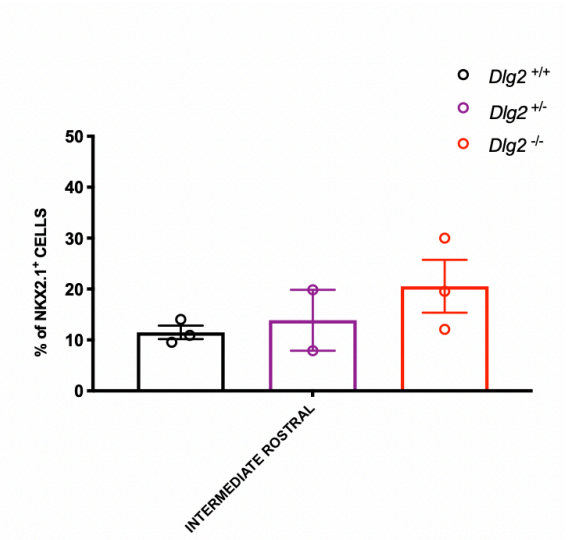


Figure 4.17. Characterisation of NKX2.1$^+$ migrating cells at intermediate rostral level of embryonic mouse brain from *Dlg2*$^{+/+}$, *Dlg2*$^{+/-}$ and *Dlg2*$^{-/-}$ genotypes. Immunohistochemistry on E15.5 embryonic mouse brains of WT (*Dlg2*$^{+/+}$), heterozygous (*Dlg2*$^{+/-}$) and homozygous (*Dlg2*$^{-/-}$) was used to visualise the marker expression. The percentages of positive cells over total cells were analysed using ImageJ (FIJI) and Prism8. According to the one-way ANOVA test, genotype ($F_{2, 5}$=1.331; P=0.3439; n≥2) had no significant effect on the percentage of NKX2.1 positive cells. Three WT mouse brains were examined. For *Dlg2*$^{+/-}$, two mouse brains were examined. For *Dlg2*$^{-/-}$, three mouse brains were examined. The bar graph presents the mean ± SEM.

To sum up, for interneuron progenitors, a declining trend was observed in the number of NKX2.1 positive cells in the homozygous (*Dlg2*$^{-/-}$) frontal rostral and medial embryonic mouse brain sections. Across all genotypes and sections, no difference was observed in the number of OLIG2 positive cells. However, a declining trend was observed in COUPTF2 positive cells in the homozygous (*Dlg2*$^{-/-}$) and heterozygous (*Dlg2*$^{+/-}$) caudal-2 CGE areas. Compared to WT (*Dlg2*$^{+/+}$), no significant difference was observed on migrating NKX2.1$^+$ cells in the intermediate rostral in the homozygous (*Dlg2*$^{-/-}$) or heterozygous (*Dlg2*$^{+/-}$).

4.5. Discussion

In the third chapter, a cortical differentiation protocol, phenotyping of the WT and *DLG2*-deficient neuronal progenitors and mature cortical interneurons was described. The phenotyping analyses performed in Chapter 3 indicates there are significant defects in *DLG2*$^{-/-}$ cells at different developmental stages throughout the differentiation of hESCs into cortical interneurons. These phenotypes include significant differences in the expression of NKX2.1 and OLIG2 in neural precursors at day 20. Also, a reduced trend in GABA expression in interneurons was observed at day 50. Furthermore, neurons lacking DLG2 migrated shorter distances and at reduced speed than neurons that did not lack the protein. Also, *DLG2*-KO neurons displayed fewer secondary neurites than WT neurons. Therefore, heterozygous (*Dlg2*$^{+/-}$) and homozygous (*Dlg2*$^{-/-}$) embryonic mouse brains were characterised to confirm the DLG2 deficiency phenotypes found in a

human culture system (Chapter 3). In general, this chapter can be considered an exploratory study for interneuron developmental phenotypes in *Dlg2* KO mice, as this has not been reported so far.

4.5.1. Choice of the embryonic mouse age

As the involvement of DLG2 in interneuron development was studied in both progenitors and developing interneurons *in vitro*, embryonic mouse brains that have both cell types were used as a comparable system. E15.5 embryonic mouse brain was selected because at this time point there is a strong expression of interneurons progenitor markers (such NKX2.1 and OLIG2) (Brandao and Romcy-Pereira, 2015; Flandin et al., 2010; Gelman et al., 2009); also, interneurons have reached the neocortex (Anderson et al., 1997; Del Rio et al., 2000; Jimenez et al., 2002; Tanaka et al., 2003) at E14.5 (Sahara et al., 2012). The fact that both populations are present in the brain at the same time point makes this a suitable age of embryonic mouse brain for the current study.

4.5.2. Expression of key transcription factors in WT and *Dlg2* mutant embryonic mouse brain

According to the *in vivo* results and the expression pattern and the percentage of markers-positive cells, the deficiency and haplodeficiency of DLG2 in embryonic mouse brain did not prevent the generation of interneurons progenitors overall. Furthermore, the findings shown here conclude that *Dlg2* heterozygosity and homozygosity do not have a strong effect on the development of interneurons in general. However, based on the percentage of cells expressing key transcription factors during interneuron development there are expression trends in homozygous (*Dlg2*$^{-/-}$) MGE and CGE areas and the heterozygous (*Dlg2*$^{-/+}$) CGE area. A declining trend in the percentage of cells expressing NKX2.1 (in *Dlg2*$^{-/-}$only) and COUPTF2 (both heterozygous and homozygous) was observed. This might indicate that the cells' fate has been changed (a shift of identify from MGE to another GE or other, to become a different lineage). Additional staining using different markers, like PAX6 (excitatory progenitors), or FOXP2 (marker for LGE progenitors) could give an insight into the phenotype during this critical developmental stage. Alternatively, the observed decline might be attributable to a delay in expression of these transcription factors (i.e. eventually they will express to the same level as WT). Analysis of the brains at later time points will be able to answer this question.

Lack of statistical significance here may reflect the number of samples analysed; further analysis may bring in significant differences. The confirmation of observed phenotypes and the investigation of the mechanisms underlying these observed phenotypes will

require more experiments with larger sample sizes, staining with more markers and at different developmental time points, including postnatal and adult stages.

4.5.3. Similarities and dissimilarities between human and mouse cellular phenotype

In vitro, hESC-derived cortical interneurons and their progenitors showed signs of the involvement of *DLG2* in interneuron development (fate specification, morphology development and migration) when *DLG2* deficient cell phenotypes were characterised at different time-points during the development of interneurons. Similarly, *in vivo*, (*Dlg2*$^{-/-}$) homozygous brains showed signs of the involvement of *DLG2* in the development of interneurons as well. Therefore, to allow precise comparison, the embryonic mouse brain sections were stained for NKX2.1, OLIG2, COUPTF2, SST, PV and CALBINDIN markers. These markers have also been used to stain the hESCs-derived cortical interneurons. However, *in vivo* no expression of PV or SST was detected in any of the brains. This was not surprising since the early expression of PV was reported postnatally from day 4 onward (Davila et al., 2005). However, these markers were expressed *in vitro*. It possible that the *in vitro* interneurons were more mature than those found *in vivo* at E15.5.

At day 20, there was a significant decrease in interneuron progenitor marker NKX2.1 hESC-derived progenitors; this marks the ventral MGE-derived progenitors. Although not statistically significant, there was a declining trend in the percentage of NKX2.1-positive cells in the frontal rostral and medial sections in the homozygous (*Dlg2*$^{-/-}$) embryonic mouse brains compared to WT (*Dlg2*$^{+/+}$). However, heterozygous (*Dlg2*$^{-/+}$) embryonic mouse brains showed no differences. This means that lack of *Dlg2* affects the expression of NKX2.1 similarly *in vivo* and *in vitro.* As heterozygous *Dlg2*$^{+/-}$ is closer to the human condition than the full knockout, this suggests that MGE-derived NKX2.1^{+} progenitors may not be affected by 11q14.1 deletion in SCZ patients.

The day-20 KO culture showed a significant reduction in OLIG2 expression (expressed in a subpopulation of MGE-derived progenitors). Whereas, *in vivo*, no differences were observed in heterozygous (*Dlg2*$^{-/+}$) or homozygous (*Dlg2*$^{-/-}$) embryonic mouse brains compared to WT (*Dlg2*$^{+/+}$). This could be attributed to a difference in the environment that the cells were exposed to, as KO cells in the embryonic brain receive required signals, unlike hESCs cultured *in vitro*. Also, the expression of OLIG2 *in vitro* was at a much lower level than embryonic mouse brain (the highest *in vitro* expression was 4.7%, whilst the highest *in vivo* expression was 35.7%). Thus, it is either that the percentage of cells expressing certain markers could be different in mice and humans or the differences in model system used. Moreover, a slight increase (not statistically significant) in

COUPTF2-positive cells (expressed in CGE-derived progenitors) was seen *in vitro.* Meanwhile, in the caudal CGE area of heterozygous and homozygous embryonic mouse brains, a decrease trend in the number of COUPTF2-positive cells was detected. Further investigation is required to establish whether DLG2 deficiency affects COUPTF2 expression differently in both humans and mice. Also, the differences observed between human and mouse could be attributed to the possibility that DLG2 loss cause different effect on human and mouse progenitor/interneurons cells or play different function when considering interneuron development as there are differences between human and mouse during early development in specification, migration and final positioning (Laclef and Metin, 2018). All of these findings suggest an involvement of *Dlg2* and *DLG2* in interneuron development and specification.

In general, when considering *in vitro* and *in vivo* results, nothing was going in opposite direction, except for COUPTF2 and OLIG2. However, the percentage of cells expressing certain markers was different *in vitro* and *in vivo*. At day-20, *in vitro* NKX2.1 was 76% in WT, whereas *in vivo*, it was 55.3% in the selected part of MGE; *in vitro* OLIG2 was 4.7%, while *in vivo* it was 35.7% in the selected part of MGE; *in vitro* COUPTF2 was 22%, while *in vivo* it was 33.3% % in the selected part of CGE. These percentages may suggest that the culture contained less developed progenitors and neurons than those in E15.5 brain. However, the percentage can change depending on the brain area that was chosen to investigate; hence, the direct comparison may not be meaningful to draw any conclusion. Perhaps a better comparison can be achieved by investigating the expression of markers after transplanting hESC- and embryonic brain-derived cells into embryonic mouse brain.

4.5.4. Limitation of the study

This study has many limitations. First, the antigen-retrieval protocol that was used resulted in the loss of some areas of sections, as the thin tissues were preheated at 95°C for 10 minutes before performing immunohistochemistry. Developing a better protocol is required for thin tissues to prevent tissue damage during staining process. Also, because damaged sections were removed, it limited the sample size; therefore, the analysis had less power, limiting its ability to detect statistical differences. More samples should be investigated in further studies to validate/consolidate these results. Moreover, during counting, the cells in some sections were packed closely to each other. Were more time and resources available, imaging with a higher magnification lens would increase the accuracy of the quantification. Lastly, some KO brains showed reduced cortical length/areas; this requires more investigation.

4.5.5. Future work

To establish whether there are any differences in the number of mature interneurons, additional staining could be performed using NEUN and GABA antibodies simultaneously. Also, the results suggest the topic of a further investigation of embryonic mouse brain, in which immunohistochemistry is performed using additional markers, such as DLX2, PAX6 and FOXP2. This might reveal the effect of *Dlg2* heterozygosity and homozygosity on fate specification during this developmental period. Also, investigating adult mouse brain might give a clear answer on the persistence of phenotypes upon *Dlg2* deficiency/haplodeficiency. Also, it is important to highlight that staining and imaging whole brain (whether it is embryonic brain or adult) without cutting into sections will enable more detailed/precise quantification of marker expression, structural organisation and connectivity as well as cortex size and volume with much less error. The use of light sheet microscopy is suggested for this purpose.

4.5.6. Summary

The findings reveal that *Dlg2* heterozygosity show decrease trend on the percentage of COUPTF2-positive cells in the CGE area. *Dlg2* homozygosity shows a decreasing trend in NKX2.1 and COUPTF2 positive cells. Some of these differences have been detected in the homozygous (*DLG2*$^{-/-}$) hESCs-derived cortical interneuron culture. All these outcomes suggest *Dlg2* has a role in interneuron development, including cell specification, maturation and migration.

5. Effect of DLG2 deficiency on transcriptomes of human cortical neural progenitors and interneurons and schizophrenia genetic risks

5.1. Introduction

In Chapter 3, through cortical interneurons differentiation, WT and *DLG2*$^{-/-}$ cells were compared at different time-points using immunocytochemistry. This demonstrated a possible DLG2 function in the initial phase of cortical differentiation before neuronal development, implied by a significant decrease in NKX2.1 and OLIG2 and a slight (but non-significant) rise in COUPTF2 expression in *DLG2*$^{-/-}$ neural precursors. Moreover, developing and newborn *DLG2*$^{-/-}$ neurons displayed a slight (but non-significant) decrease in the expression of the interneuron's marker GABA in *DLG2*$^{-/-}$ cells by contrast to WT. Furthermore, it is noteworthy that the actual interneuron generation rate was unaffected by *DLG2* knockout, with the expression of post-mitotic neuron marker (NEUN) being comparable to WT. Additionally, DLG2-deficient interneurons showed defects in migration as they were found to migrate at a slower speed and across a shorter distance. Morphometric analysis provided further evidence for the deferred development of *DLG2*$^{-/-}$ neurons, which appeared structurally less complex compared to WT. Also, in Chapter 4, DLG2-deficient embryonic mouse brains were used to characterise both *Dlg2* homozygous and heterozygous mouse interneurons and investigate whether the *in vivo* and *in vitro* findings were parallel. The results showed that *Dlg2* heterozygosity display decrease trend (not significant) in the percentage of CGE COUPTF2^{+} cells while *Dlg2* homozygosity show a decrease trend (not significant) in NKX2.1 and COUPTF2-positive cells. Some of these differences (*in vitro*, NKX2.1 showed a significant decrease) were observed in the homozygous (*DLG2*$^{-/-}$) hESCs-derived cortical interneurons.

Overall, the above results provided support for a new DLG2 function in cortical interneuron development. The *DLG2*$^{-/-}$ phenotype in neural progenitors and interneurons can be additionally investigated through a wide range of methods. *DLG2*$^{+/-}$ cell lines sourced from patients could be used instead of *DLG2*$^{-/-}$ hESC lines to conduct the work in this thesis. However, homozygous cell lines were best suited for comprehending *DLG2* function in early interneuron development. The recovery of *DLG2*$^{-/-}$ neuron marker expression, migration and morphology to WT levels after a longer time in culture can also be assessed through phenotypic profiling beyond day 54 of interneuron development. However, long-term maintenance of cINs in culture is problematic because it leads to greater expenses and longer periods of differentiation. Also, cINs maturation needs simulation from the surrounding environment. Therefore, examination of *DLG2*$^{-/-}$ phenotypes at later time points is possible by co-culturing cINs with excitatory neurons or transplanting human *DLG2*$^{-/-}$ interneurons into mouse brains for interneuron maintenance and developmental stimulation (Maroof et al., 2013; Nicholas et al., 2013;

Patz et al., 2004). Moreover, using immunocytochemistry or mass-spectrometry, different cortical proteins with extra markers, e.g. OTX2, DLX5, DLX6, LHX6, SOX6, ARX, SP8, GAD65, VIP, Reelin, and NPY (Kessaris et al., 2014) could also be investigated.

Unfortunately, owing to time and resource limitations, further experiments could not be conducted, despite their usefulness. However, it is likely that numerous other biological processes modulated by DLG2 in neural precursors and interneurons are yet to be discovered. Also, *DLG2*$^{-/-}$ phenotype clarification and understanding of the underpinning mechanism necessitates a protocol with higher throughput. Such an approach could be bulk or single-cell RNA transcriptomics (scRNAseq). Dissimilar to bulk RNA sequencing profiling, whereby a sequencing library consists of several heterogeneous cells, scRNAseq technology isolates a single cell and builds a cell-specific sequencing library and marks RNA contents with cell-specific molecular barcodes (Chen et al., 2019). Besides, scRNAseq can identify the differences in transcription in a single cell that would ordinarily be lost and would be undetectable if cell-aggregation techniques are used for analysis. Additionally, this method allows cell types within a particular population to be determined, including rarer cell types (Kumar et al., 2017). This method is useful for studying unique cells, such as early-stage embryonic cells or neurons (Close et al., 2017). Thus, to understand the mechanisms that may trigger DLG2 deficiency phenotypes in human cells (*in vitro*), scRNAseq was performed and gene expression in both *DLG2*$^{-/-}$ and WT cells was measured at several time-points (day 20, 30 and 50) during cortical interneurons differentiation. Investigations sought to determine genes with differential expression in *DLG2*$^{-/-}$ relative to WT lines at all time-points and assess if those identified genes were enriched for certain biological processes or phenotypes. Moreover, one identified phenotype resulting from those investigations was verified via experimental assay (adhesion assay). Lastly, the DEG sets were assessed for the existence of common risk variant enrichment for SCZ.

Analyses were conducted on WT and KOs (KO1 and KO2) cells because the findings for the two *DLG2*$^{-/-}$ lines were equivalent and the study's emphasis was on *DLG2* gene deletion as opposed to particular differences in *DLG2*$^{-/-}$ lines; the data were merged and given as WT versus *DLG2*$^{-/-}$ (combined KO1 and KO2).

5.2. Experimental outline

During cortical interneuron differentiation from hESC (Figure 5.1), 9 samples were collected at 3 time-points (day 20, day 30, and day 50) (Table 5.1). These specific time-points were chosen in order to cover several stages of cortical interneuron development

and different types of cells to better understand and investigate the *DLG2* role former to the PSD formation and synaptic transmission. Also, the protein expression of two time-points (day 20 and day 50) was characterised in Chapter 3. A single-cell preparation for 10x chromium was done to obtain a final cell count of 3000 cells. After single-cell capture, RNA extraction, cDNA, and library preparation, a Chromium Single Cell 3' Reagent Kit v3 (10x Genomics) was used and sequencing was performed using Illumina Hiseq4000. Then, quality control (QC), read alignment on reference transcriptome, and filtering and counting of unique molecular identifier (UMI) were assessed. After that, cells were selected, normalized, and clustered. Then, the differentially expressed genes (DEGs) were identified, and different cell types were identified using WT cells only. A total of 24 clusters were identified representing groups of cell types; then DEGs between each cluster vs the others were identified. Next, DEGs between WT and *DLG2*$^{-/-}$ cells were produced for each cluster. Following that, gene ontology analysis (GO) was performed on these sets of DEGs (comparing KO vs WT) utilising gene ontology data to identify specific biological or functional processes that were significantly enriched or depleted in *DLG2*$^{-/-}$. An adhesion experimental assay was performed to investigate whether this biological process—recognised from the enrichment analysis—displayed dysregulation evidence in *DLG2*$^{-/-}$ cultured cells, thus providing validation of the bioinformatic analyses outcomes.

Finally, after the DEGs and GO enrichment analysis were identified and characterised between *DLG2*$^{-/-}$ and WT, these DEG sets were assessed for the presence of common risk variant enrichment for schizophrenia (SCZ). SCZ-related GWAS summary statistics were extracted from The Schizophrenia Working Group of the Psychiatric Genomics Consortium *et al.,* (Pardinas et al., 2018; Trubetskoy et al., 2022). A MAGMA gene set analysis was carried out on upregulated and down-regulated genes in the clusters and groups. The MAGMA analysis was used to identify if any over-represented gene set was involved in SCZ and if so, to what extent. The gene sets employed for the work documented in the current chapter were extracted from the merged *DLG2*$^{-/-}$ versus WT scRNAseq data presented in the same chapter. The experiments performed for this chapter are described in detail in Chapter 2.

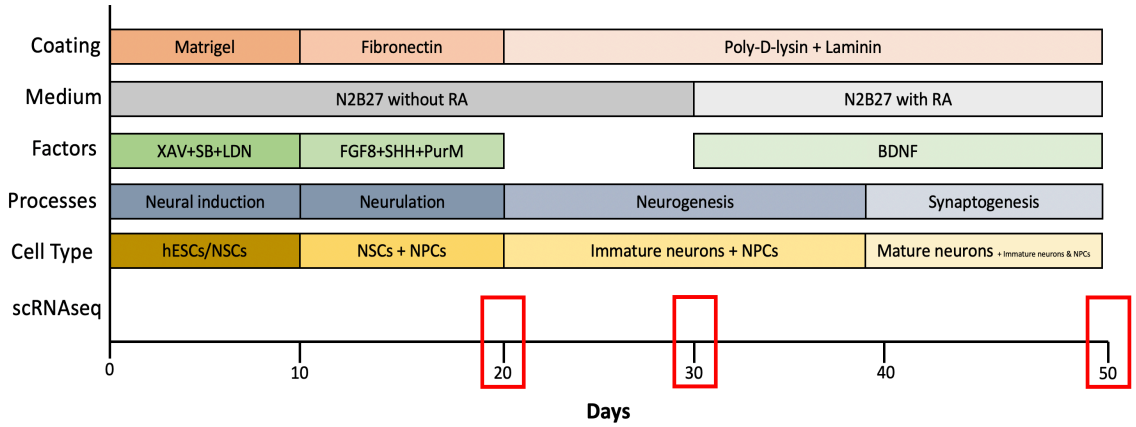


Figure 5.1. Outline of the differentiation protocol. Summary of the differentiation protocol used including timing, cell types, and developmental process. Selected time points for scRNAseq are shown in red boxes. The plate coatings, cell mediums, and molecule factors used throughout the cortical interneuron differentiation process are explained in detail in the Chapter 2.

Samples	Cell Viability
Sample 1 (Day 20 - WT$^{DLG2\ +/+}$)	90%
Sample 2 (Day 20 – KO1$^{DLG2\ -/-}$)	92%
Sample 3 (Day 20 – KO2$^{DLG2\ -/-}$)	94%
Sample 4 (Day 30 - WT$^{DLG2\ +/+}$)	89.9%
Sample 5 (Day 30 – KO1$^{DLG2\ -/-}$)	90.4%
Sample 6 (Day 30 – KO2$^{DLG2\ -/-}$)	87%
Sample 7 (Day 50 - WT$^{DLG2\ +/+}$)	82.77%
Sample 8 (Day 50 – KO1$^{DLG2\ -/-}$)	85%
Sample 9 (Day 50 – KO2$^{DLG2\ -/-}$)	81.26%

Table 5.1. Samples' cell viability. Shows the percentage of cell viability in each sample. Cell counting was done manually in order to achieve 120,000 cells in 100µl. The cell calculation method is fully described in Chapter 2.

5.3. Quality checks and filtering

RNA was obtained from WT and two *DLG2*$^{-/-}$ KOs (KO1 and KO2) cell lines (on days 20, 30, and 50, respectively) during the process of cortical interneuron differentiation using 10x. The cDNA was synthesised, and the concentration of cDNA was measured for all 9 samples. Table 5.2 confirms the concentration of cDNA for all samples. Also, the shape of the cDNA bands on the gel was equivalent between samples (e.g. comparable size) (Figure 5.2). The cDNA concentration gives an indication of what might be anticipated; among samples from the same time-points there were similar amounts of cDNA (Table 5.2). The cDNA concentrations for the samples on day 20 (1, 2, and 3) were 24.4, 20.8, and 25.8, respectively; for the samples on day 30 (4, 5, and 6) the cDNA concentrations were 19.75, 17.9, and 16.6, respectively, while for the samples on day 50 (7, 8, and 9) the cDNA concentrations were 24.6, 24.2, and 21, respectively (Table 5.2). Next, the cDNA library was created through PCR amplification. After quantification and authentication of the libraries via Tape-Station (to ensure that adequate cDNA was

present for the efficacious sequencing of the specimens) (Figure 5.3), these were later sequenced by Joanne Morgan.

Samples	cDNA Qubit concentration (ng/µl)
1	24.4
2	20.8
3	25.8
4	19.75
5	17.9
6	16.6
7	24.6
8	24.2
9	21

Table 5.2. Sample's cDNA concentration. All samples were checked for cDNA quantity. Among samples from the same time-points there were similar amounts of cDNA. Sample 1 (Day 20 - WT$^{DLG2\ +/+}$), sample 2 (Day 20 – KO1$^{DLG2\ -/-}$), sample 3 (Day 20 – KO2$^{DLG2\ -/-}$), sample 4 (Day 30 - WT$^{DLG2\ +/+}$), sample 5 (Day 30 – KO1$^{DLG2\ -/-}$), sample 6 (Day 30 – KO2$^{DLG2\ -/-}$), sample 7 (Day 50 - WT$^{DLG2\ +/+}$), sample 8 (Day 50 – KO1$^{DLG2\ -/-}$), and sample 9 (Day 50 – KO2$^{DLG2\ -/-}$).

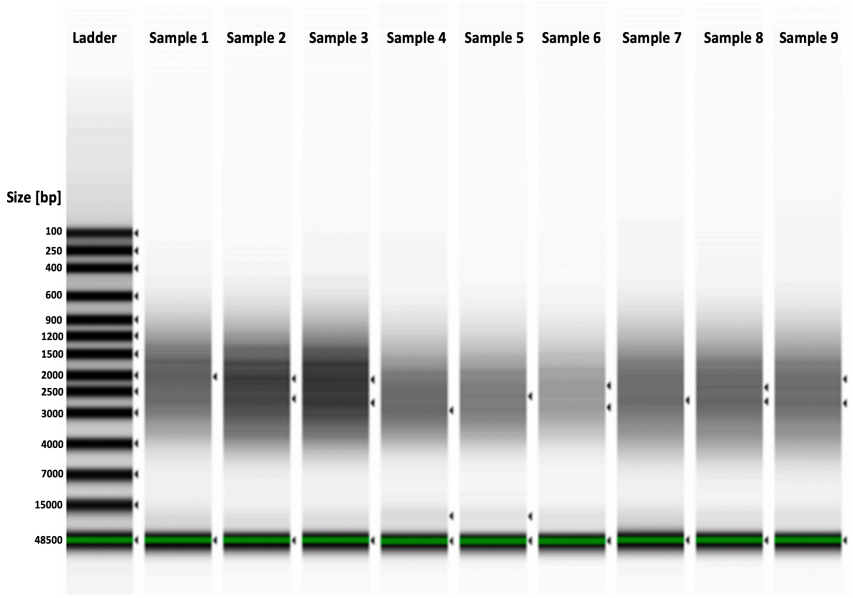


Figure 5.2. Extracted cDNA concentration check. The cDNA quantification in each sample (results of samples 1– 9 respectively) using Tape-Station 2200. All samples' cDNA appeared quite similar in size and distribution on the equivalent of gel. The first 3 samples (samples 1, 2, and 3) were more concentrated, which matched the Qubit, although this was slightly less in the last 3 samples (samples 7, 8, and 9). And day 30 samples (samples 4, 5 and 6) appeared similar considering the distribution on the gel. Sample 1 (Day 20 - WT$^{DLG2\ +/+}$), sample 2 (Day 20 – KO1$^{DLG2\ -/-}$), sample 3 (Day 20 – KO2$^{DLG2\ -/-}$), sample 4 (Day 30 - WT$^{DLG2\ +/+}$), sample 5 (Day 30 – KO1$^{DLG2\ -/-}$), sample 6 (Day 30 – KO2$^{DLG2\ -/-}$), sample 7 (Day 50 - WT$^{DLG2\ +/+}$), sample 8 (Day 50 – KO1$^{DLG2\ -/-}$), and sample 9 (Day 50 – KO2$^{DLG2\ -/-}$).

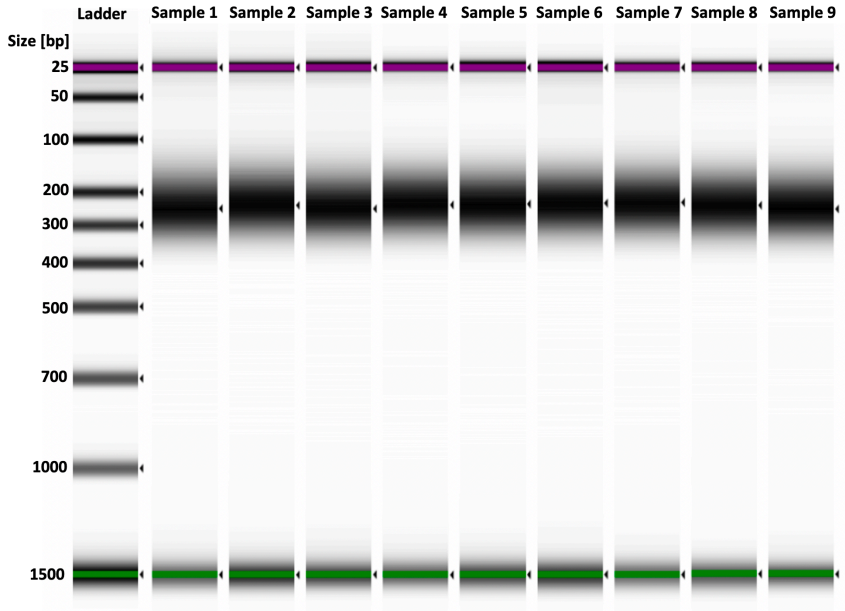


Figure 5.3. Library quality of cDNA. The gel from Tape-Station 2200 was utilised to check the quality of the cDNA libraries. cDNA libraries fragment size quantification was good for the 9 samples. The samples appeared very similar in size with one clear band considering the size and shape of the library. All samples' libraries were good; there were even thick black smeared bands between 200–300bp, which is the sequencing's optimal range. Sample 1 (Day 20 - WT$^{DLG2\ +/+}$), sample 2 (Day 20 – KO1$^{DLG2\ -/-}$), sample 3 (Day 20 – KO2$^{DLG2\ -/-}$), sample 4 (Day 30 - WT$^{DLG2\ +/+}$), sample 5 (Day 30 – KO1$^{DLG2\ -/-}$), sample 6 (Day 30 – KO2$^{DLG2\ -/-}$), sample 7 (Day 50 - WT$^{DLG2\ +/+}$), sample 8 (Day 50 – KO1$^{DLG2\ -/-}$), and sample 9 (Day 50 – KO2$^{DLG2\ -/-}$).

The Cell Ranger Single-Cell Software Suite (version 3.1.0) was used to perform sample demultiplexing, barcode processing and single-cell 3′ gene counting (https://software.10xgenomics.com/single-cell-gene-expression). Briefly, the *cellranger mkfastq* command was used to demultiplex the Illumina sequencer's base call files (.bcl) and generate FASTQ files (.fastqc) for the Read1 and Read2 paired-end reads containing the barcode and read data. The first read of the pair, containing the cell part of the primer, was divided into barcode + UMI. In the demultiplexing step, the low-quality sequences according the PHRED score (Wilkins et al., 2021) wear also filtered out. A further quality check was performed by using the FastQC tool (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) to confirm the good quality of the sequencing, with a mean per base sequence quality above 30 PHRED for all the samples. The reads were successively processed by the *cellranger count* pipeline with standard parameters that performed the filtering, the alignment on reference transcriptome (GRCh38-3.0.0), and UMI counting.

Sample	**Millions reads in total**	**Estimated number of cells**	**Mean reads per cell**	**Median genes per cells**	**Reads mapped to genome**
Sample 1 (Day 20 – WT$^{DLG2+/+}$)	216,141,944	3269	66119	5023	92.8%

Sample 2 (Day 20 – KO1$^{DLG2-/-}$)	179,034,637	13681	13086	1240	92.2%
Sample 3 (Day 20 – KO2$^{DLG2-/-}$)	177,519,180	3709	47862	4326	92.8%
Sample 4 (Day 30 - WT$^{DLG2+/+}$)	189,409,427	2763	68552	4379	92.3%
Sample 5 (Day 30 – KO1$^{DLG2-/-}$)	214,526,101	2990	71748	4482	92.0%
Sample 6 (Day 30 – KO2$^{DLG2-/-}$)	205,225,823	2199	93327	4847	92.1%
Sample 7 (Day 50 - WT$^{DLG2+/+}$)	178,576,973	4931	36245	2528	91.4%
Sample 8 (Day 50 – KO1$^{DLG2-/-}$)	183,826,258	4732	38847	2778	91.6%
Sample 9 (Day 50 – KO2$^{DLG2-/-}$)	205,669,402	4800	42848	2912	92.0%

Table 5.3. Summary of sample cells' mean reads per cell number, median genes per cell and percentage reads mapped to genome from 10x Genomics Cell Ranger report.

Overall, considering the quality of the libraries and sequencing quality of the reads, 8 samples had comparable high standard values with ~3000 cells per sample containing ~58000 reads/cell (mean value among the samples), ~3800 genes/cell (median value among the samples), and less than 20% of mitochondrial genes detected (Figure 5.4 and Table 5.3). However, sample 2 showed different characteristics compared to the other samples, with ~13500 cells, but only ~13000 reads/cells and ~1200 genes/cells (Figure 5.4 and Table 5.3). Sample 2 showed different characteristics compared to other samples from the same time-point (sample 1 and 3), because it had a very high number of cells (13681 estimated number of cells), which was four times more than sample 1 (estimated number of cells: 3269) and sample 3 (estimated number of cells: 3709) and less number of genes per cells (1240 median genes per cells) compared to sample 1 (5023 median genes per cells) and sample 3 (4326 median genes per cells) (Table 5.3). This may be due to incorrect cell identification or an issue in barcoding possibly related to technical issues with sample 2. Therefore, sample 2 was excluded from further analysis.

Also, more QC was performed on the remaining samples. The cells were selected considering the number of counts per barcode (count depth), the number of genes per barcode, and the fraction of counts from mitochondrial genes per barcode. Firstly, the cells that had a mitochondrial read rate < 15% were selected. Then the low-quality cells were filtered out, keeping only the cells with at least 5% of expressed genes and a

number of counts > 5% with respect to the entire cell population. Furthermore, potential doublets were filtered out by removing cells with > 95% of expressed genes and a number of counts > 95% with respect to the entire population. Only protein-coding genes were included for downstream analysis.

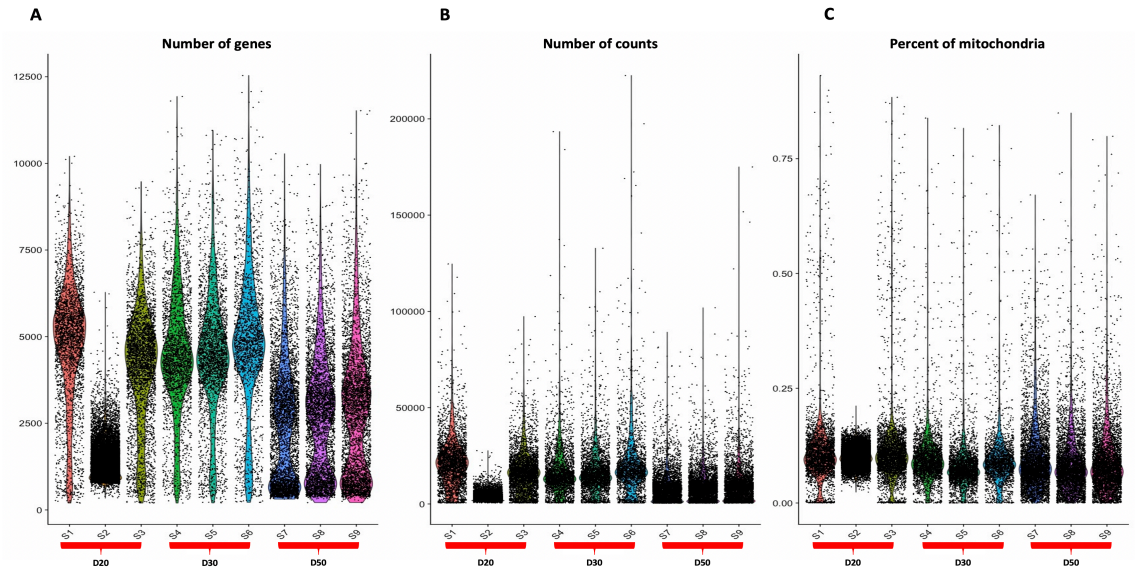


Figure 5.4. Single-cell quality checks before filtering. A) Violin plot showing the number of genes for all samples. **B)** Violin plot showing the number of counts for all samples. **C)** Violin plot showing the percentage of mitochondria genes for all samples. All plots display similar distributions apart from S2, with fewer genes and a smaller number of counts. The average number of genes for all samples was 5000, the number of counts was 50000 and a 15% mitochondria rate. Each dot represents one cell. S1: sample 1 (Day 20 - $WT^{DLG2\ +/+}$), S2: sample 2 (Day 20 – $KO1^{DLG2\ -/-}$), S3: sample 3 (Day 20 – $KO2^{DLG2\ -/-}$), S4: sample 4 (Day 30 - $WT^{DLG2\ +/+}$), S5: sample 5 (Day 30 – $KO1^{DLG2\ -/-}$), S6: sample 6 (Day 30 – $KO2^{DLG2\ -/-}$), S7: sample 7 (Day 50 - $WT^{DLG2\ +/+}$), S8: sample 8 (Day 50 – $KO1^{DLG2\ -/-}$), and S9: sample 9 (Day 50 – $KO2^{DLG2\ -/-}$).

In total, 22736 cells passed QC and were analysed using Seurat in order to investigate the presence of systematic alterations amongst the cell lines, e.g. to identify DEGs between WT, and $DLG2^{-/-}$ vs WT cell types. The row counts for each gene obtained from the sequencing were log-normalised to library size and scaled to 10,000. After that, dimensionality reduction and clustering were applied. The former was performed via principal component analysis (PCA) using the *RunPCA* function and the statistically significant principal components were identified using both *jackstraw* approach and Elbow method, using *JackStraw*, *ScoreJackStraw* and *ElbowPlot* functions. Louvain clustering was performed with a resolution parameter of 1; t-SNE was used in order to present data in two-dimensional coordinates (Van der Maaten and Hinton, 2008) using the *FindNeighbors*, *FindClusters* and *RunTSNE* functions. All the above-mentioned functions were implemented in the Seurat package. The entire analysis produced a total of 24 clusters (Figure 5.5A). The cells were plotted (clustered) from the WT and KO samples from all time-points (day 20, day 30, and day 50) (Figure 5.5A & B). All clusters

were composed of the two genotypes, indicating no emergence of KO-specific cell types (Figure 5.5).

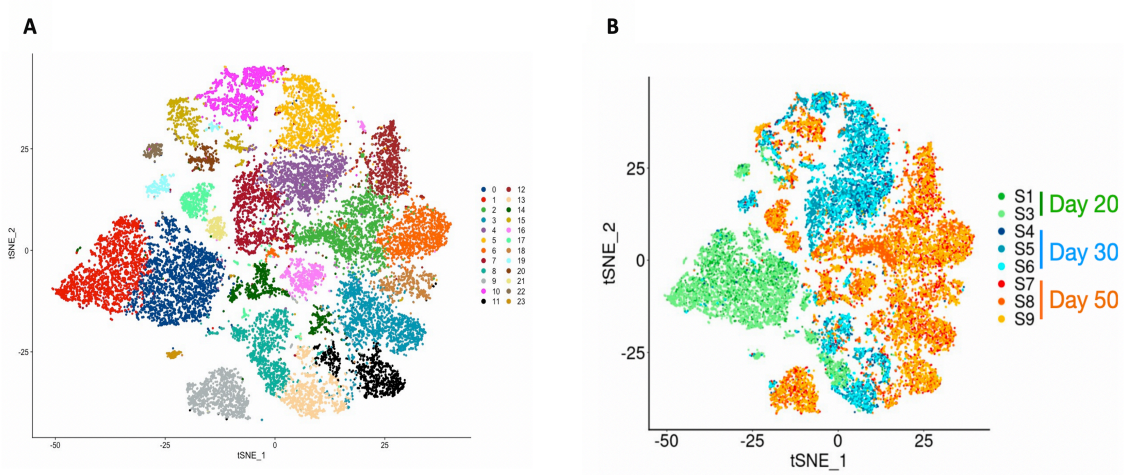


Figure 5.5. TSNE plots for sample's clusters in two-dimensional coordinates tSNE1 and tSNE2. A) The figure shows the TSNE cluster obtained considering WT and KOs cells from 8 samples (sample number 2 was excluded). The resolution parameter was set to 1, perplexity to 30 (default) and the number of principal components was 20. **B)** Shows TSNE plot for all samples (WT and KO) clustered together (S1=WT day 20, S3=KO2 day 20, S4=WT day 30, S5=KO1 day 30, S6=KO2 day 30, S7=WT day 50, S8=KO1 day 50, and S9=KO2 day 50). No domination of one genotype was evident in each cluster. Resolution parameter set to 1.

5.4. Cluster identity

Nowakowski and colleagues succeeded in obtaining single-cell transcriptomes of progenitor cells to post-mitotic neurons of various stages of development from human embryonic cortical and MGE samples (48 human tissue samples aged from 5.85 to 37 PCW, covering different phases of peak neurogenesis) using scRNAseq (Nowakowski et al., 2017). In this thesis, hESCs were differentiated into cortical interneurons and single-cell samples were collected at different stages of differentiation. Due to the similarity of cell types and developmental stages, the Nowakowski dataset was used as a reference to define the clusters' identity together with gene ontology terms enriched in cluster-specific differentially expressed genes and well-known markers for cell types and lineages.

First, single-cell RNA-sequencing gene expression data from Nowakowski *et al.* were downloaded [https://cells.ucsc.edu/?ds=cortex-dev]. Cells corresponding to distinct inhibitory cell types only—radial glia, intermediate progenitor cells (IPCs), dividing, newborn, mature MGE (all MGE derived population), and mature CGE—were identified and extracted, collating all cells from the corresponding *in vivo* cell clusters in 6 cell-types as in Table 5.4.

MGE Progenitors	Cortical inhibitory neurons
MGE RG: "MGE-RG1", "MGE-RG2"	**MGE Newborn:** “nIN1”, “nIN2”, “nIN3”, “nIN4”, “nIN5”
MGE IPC: "MGE-IPC1", "MGE-IPC2", "MGE-IPC3"	**MGE mature:** "IN-CTX-MGE1”, “IN-CTX-MGE2"
MGE DIV: "MGE-div"	**CGE mature:** "IN-CTX-CGE1”, “IN-CTX-CGE2"

Table 5.4. Cells from Nowakowski *et al*. Cells selected from Nowakowski *et al* corresponding to different inhibitory cell types.

Class	UP	DN
IN-MGE_vs_others	222	243
IN-CGE_vs_others	198	295
MGE-RG_vs_others	1170	177
MGE-Newborn_vs_others	124	628
MGE-IPC_vs_others	680	97
MGE-DIV_vs_others	627	75

Table 5.5. Differentially expressed genes on Nowakowski *et al*. selected cells. The numbers of up (UP) and down (DN) regulated genes for each comparison.

Differential expressed genes (upregulated and downregulated) between each specific cluster and the others were identified using the Seurat (v3.1.1) function *FindMarkers.* Differentially expressed genes were considered significant if their p-value after Bonferroni correction was < 0.05 (Table 5.5) (Supplementary Table1). The DEG analysis on Nowakowski data set was performed by Dr Daniel D'Andrea (MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University). All subsequent analyses were performed by me.

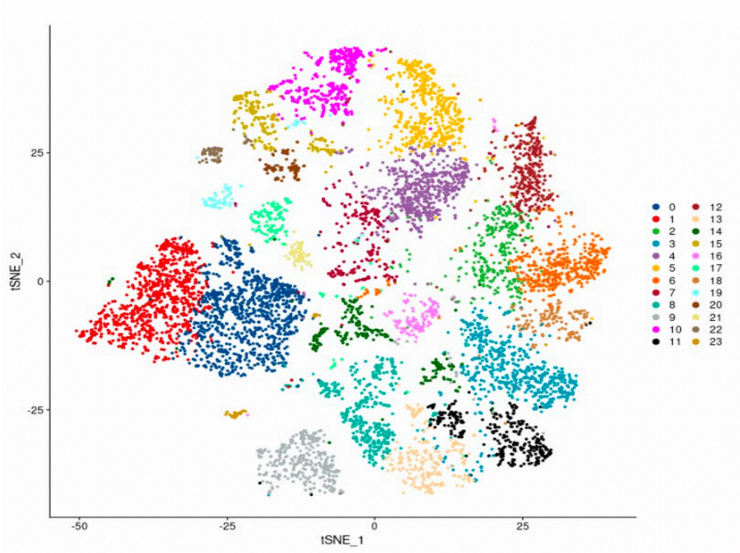


Figure 5.6. TSNE plots for WT sample clusters in two-dimensional coordinates tSNE1 and tSNE2. Displays TSNE plot of WT sample clusters (S1=WT day 20, S4=WT day 30, and S7=WT day 50) (KO cells were removed). Resolution parameter set to 1.

Once clusters were obtained and the distributions of samples in the clusters were checked, the KO samples were removed (Figure 5.6). Then, using data from WT cells only (due to the possibility of altered cell status in KO cells in comparison to WT cells in each cluster, which would potentially hinder the accurate determination of the cluster identity) in each cluster (S1=WT day 20, S4=WT day 30, and S7=WT day 50) (Figure 5.6), DEG analysis on the protein-coding genes was performed testing a total of 16742 genes. Once the data were log-normalised and scaled, the differentially expressed genes (upregulated and downregulated) between each specific cluster and the other clusters were identified using the Seurat function *FindMarkers* (Supplementary Table 2). DEGs were considered significant if their p-value after Bonferroni correction was < 0.05. Afterwards, the sets of upregulated genes in each cluster were tested for enrichment against those in each cell-type group obtained from the Nowakowski *et al*.'s study. The overlap degree was calculated using Fisher's exact test and all genes from hESCs derived cortical interneurons data and in Nowakowski's dataset were used as the background set. Enrichment was deemed significant if the p-value after Bonferroni correction was < 0.05.

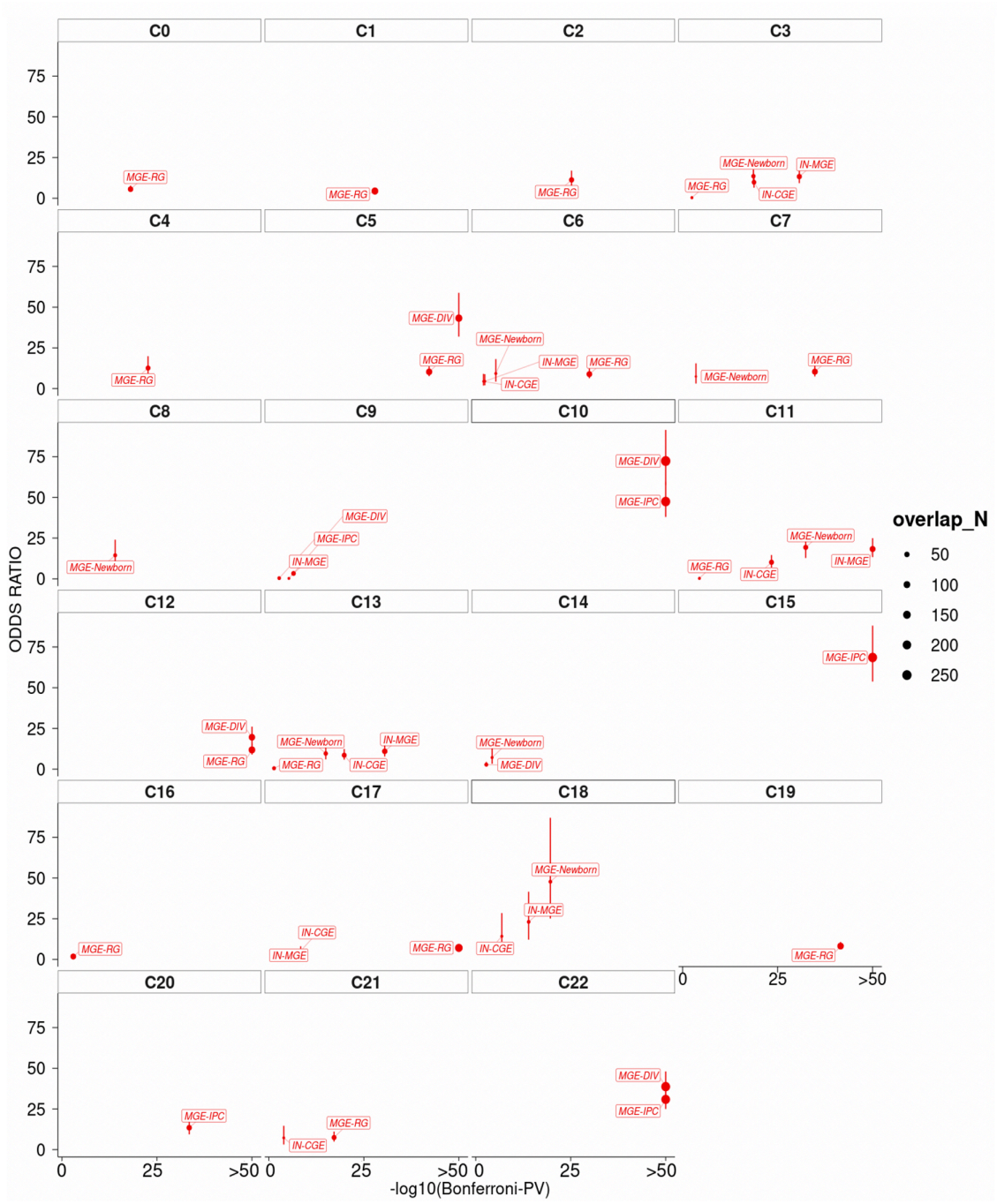


Figure 5.7. Scatter plots showing the results from enrichment analysis between DEGs from Nowakowski and DEGs from WT clusters. Odds Ratios and -log$_{10}$(Bonferroni-p-value) for each cluster. The size of the points shows the number of overlapping genes between DEGs from Nowakowski and DEGs from WT clusters. The lines show the lower and higher odds ratio values. Only enrichments with Bonferroni P<0.05 are shown.

The DEGs of all clusters apart from the smallest cluster (cluster 23, data not shown) indicated strong enrichments in line with Nowakowski et al.'s well-defined clusters, thus enabling the clusters to be identified (Figure 5.7). By comparing the upregulated genes in WT clusters to the upregulated genes in Nowakowski's clusters, each of the 24 clusters was associated with one of the Nowakowski classes:

- **Radial glia cells** (RG) (C0, C1, C2, C4, C16, C19): enrichment in Nowakowski RG only;

- **Dividing radial glia cells** (Div RG) (C5, C12): enrichment in Nowakowski RG and Nowakowski dividing RG;
- **Cells in the process of transitioning from progenitors into neurons** (Tran RG) (C6, C7, C17, C21): enrichment in Nowakowski RG and Nowakowski neurons (newborn/developing), where RG and neuron enrichments are similar (according to the odds-ratio);
- **Intermediate progenitor cells** (IPC) (C15, C20): enrichment in Nowakowski IPC only;
- **Dividing intermediate progenitor cells** (Div IPC) (C10, C22): enrichment in Nowakowski IPC and Nowakowski dividing IPC;
- **Neurons derived from intermediate progenitor cells** (IPC-derived neuron – IPC DN) (C9): enrichment in Nowakowski IPC and Nowakowski neurons (newborn/developing), where IPC and neuron enrichments are similar (according to the odds-ratio);
- **Newborn neuron** (Newb N) (C8, C14, C18): enrichment in Nowakowski newborn notably higher than that for other classes (according to the odds-ratio); and
- **Developing neuron** (Dev N) (C3, C11, C13): enrichment in Nowakowski neuronal (MGE/CGE) and newborn are similar.

It is important to note that cluster 9 looks a bit different from other neuronal clusters as it has no newborn enrichment. One hypothesis is that this difference may be due to it being derived from IPCs rather than RG (as it has some MGE-IPC and MGE-DIV enrichment, and no RG enrichment). In order to better characterise this cluster, further validations could be drawn from GO and marker analyses. Cluster 23 did not show any enrichment, so it was left as unknown.
It was interesting that the two big Tran RG clusters (C6 and C7) showed enrichment in upregulated genes in newborn clusters while the smaller Tran RG clusters (C17 and C21) had enrichment in genes upregulated in later-stage developing interneurons. This suggests two possibilities: (i) the developmental stage of C6 and C7 is earlier than that of C17 and C21; and/or (ii) they are different sub-types of RG, hence their different internal state.

As the first cell identity of clusters was driven from Nowakowski datasets, and in order to better investigate the biological functions associated with the up- and down-regulated DEGs from each cluster and identify which functional classes were changing, GO term enrichment for the sets of up- and down-regulated genes was performed in each of the clusters. The R script for this process was generously scribed and offered by Dr. Andrew Pocklington (MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University) (Supplementary File 1). The enrichment analysis produced a list of gene ontology classes enriched for up- and down-regulated genes (Supplementary Table 3).

The most informative GO classes, i.e. those that provided insight into cluster identity, are summarised in Figure 5.8.

The GO enrichment analysis confirmed that IPC-derived neuron and Tran RG clusters were truly transitioning, as the GO outcome shows downregulation of cell-cycle-related terms and upregulation of neurogenesis-related terms (Figure 5.8). Regarding the transitioning RGs class (Tran RG), clusters C6, C17, and C21 were clearly transitioning to neurons since the GO enrichment illustrates upregulation in nervous system development, receptor binding and neuron differentiation, and downregulation of cell-cycle-related processes. In contrast, cluster C7 showed upregulation in neuron differentiation and developmental process genes and downregulation in neuron part and neuron projection. Therefore, this cluster was interpreted as transitioning into neurons. The IPC-derived neuron cluster (C9) appeared to be transforming into neurons because it demonstrated upregulation of terms such as neurogenesis and neuron part and downregulation of cell-cycle related terms.

For developing and newborn neurons, GO terms were in line with the defined identity (based on downregulation in cell-cycle terms and upregulation in neuron development, cell morphogenesis, and axonogenesis) (Figure 5.8). Interestingly, the RG, Div RG, IPC, and Div IPC GO enrichments were in line with their defined identities. The GO terms showed upregulation in cell-cycle regulation, cell adhesion, and RNA regulation, and downregulation in neurogenesis and neuron development for the RG group (C0, C1, C16, C19, C2, and C4). Meanwhile, Div RG, IPC, and Div IPC all showed upregulation in DNA replication, RNA processing, and downregulation in neurogenesis, except for Div IPC, which showed downregulation in the apoptotic process.

Higher Cluster	C	GO Term
RG (C0, C1, C16, C19, C2, C4)	C0	RNA processing, regulation of cell cycle and regulation of apoptotic process
	C1	Cell adhesion and cytoskeleton regulation
	C16	Endoplasmic reticulum membrane process and cellular protein localization
	C19	Wnt protein binding, glycolysis and cell adhesion
	C2	RNA processing, cell adhesion regulation and ribosome biogenesis
	C4	RNA transulation regulation and ribosome biogenesis
Tran RG (C17, C21, C6, C7)	C17	Neurogenesis, nervous system development and cell differentiation and growth
	C21	Developmental process regulation, receptor binding and cell differentation
	C6	Cell structure morphogenesis, nervous system development and neuron differentiation
	C7	Extracellular matrix activity, neuron differentiation, cell adhesion and developmental process
Div RG (C12, C5)	C12	Ribosome and RNA processing, cell cycle regulation and DNA replication
	C5	DNA replication, RNA binding and cell cycle regulation
IPC (C15, C20)	C15	Chromosome segregation, cellular localization, microtubule cytoskeleton organization and nuclear division regulation
	C20	RNA processing, intracellular transport, cell cycle phase transition and DNA replication
IPC DN (C9)	C9	Neurogenesis and neuron part
Div IPC (C10, C22)	C10	DNA and cell cycle replication, cell cycle transion, RNA process and microtubule regulation
	C22	Cell cycle phase transition, DNA replication and repair, intracellular signal transduction
Newb N (C14, C18, C8)	C14	Ribosome and RNA processing, cellular and mitochondrial proces
	C8 & C18	Neuron differentation, neuron development and nervous system development (RNA and DNA binding C8 only)
Dev N (C11, C13, C3)	C11	Cytoskeleton, microtubule and neural cell body orgnization, neurogenesis and axonogenesis
	C13	Neurogenesis regulation, cell signaling, cell morphogenesis, axonogenesis, microtubule cytoskeleton regulation
	C3	Neuron differentation, neuron development and nervous system development

Higher Cluster	C	GO Term
RG (C0, C1, C16, C19, C2, C4)	C0	Neurogenesis, differentation and nervous system development
	C1	Neurogenesis, differentation and nervous system development
	C16	Ribosome and RNA processing, cell cycle phase transition and regulation of mitochondria
	C19	Neuron development and nervous system development
	C2	Celular and molecular component regulation and cell cycle processing
	C4	Neuron, cytoskeleton and nervous system development
Tran RG (C17, C21, C6, C7)	C17	Ribosome, RNA processing and cellular processes
	C21	Ribosome and RNA processing, translation and mitochondrial regulation
	C6	DNA packaging and cell cycle regulation
	C7	Neuron part and neuron projection
Div RG (C12, C5)	C12	Neurogenesis, cell differentation and cell projection
	C5	Neuron differentiation, neuron development and cell morphogenesis
IPC (C15, C20)	C15	Transferrin transport
	C20	Cellular localization and neurogenesis
IPC DN (C9)	C9	RNA binding, cell cycle transition, protien regulation and mitochondrion regulation
Div IPC (C10, C22)	C10	Apoptotic process
	C22	
Newb N (C14, C18, C8)	C14	Neuron differentiation, neuron development, RNA splicing and regulation of chromatin organization
	C18	Extracellular matrix regulation
	C8	Collagen processing, focal adhesion, DNA packaging and cell cycle process
Dev N (C11, C13, C3)	C11	Ribosome and RNA processing, cell cycle phase transition and cell adhesion
	C13	Ribosome and RNA processing, cell cycle regulation, apoptotic process and cell adhesion
	C3	Cell cycle arrest regulation, ribosomal biogenesis and RNA binding

Figure 5.8. GO terms enriched for upregulated- (green) and downregulated (orange) genes for each cluster. The GO terms were selected and grouped together. The terms were highlighted

manually considering the ones that were relevant and helped to identify the clusters' functional clarification for up- and down-regulated genes. C is the individual cluster, RG (radial glia), Div RG (dividing RG), Tran RG (transitioning RG), IPC (intermediate progenitor cells), Div IPC (dividing IPC), IPC DN (IPC-derived neurons), Newb N (newborn neurons), and Dev N (Developing neurons).

To obtain further support for accurate cluster identity defined using the data of Nowakowski and GO terms, further investigation of the expression of well-known markers of RG, Div RG, Tran RG, IPC, Div IPC, IPC DN, Newb N, and Dev N cell type was undertaken. Progenitor markers (Bansod et al., 2017; Bernal and Arranz, 2018; Ferri et al., 2013; Hatakeyama et al., 2004) were generally expressed in all progenitor clusters apart from *HES5* (Figure 5.9). *HES5* expression level was low in C0 and C1 compared to other markers such as *SOX2* and *NES*. *Hes5* expression is stronger once neural epithelial cells/neural stem cells transit into radial glial cells (neural progenitor cells) when they start to receive strong intercellular Delta-Notch signals from differentiating and maturing neurons (Bansod et al., 2017; Hatakeyama et al., 2004; Kageyama and Ohtsuka, 1999). Possibly, C0 and C1 represented a less-developed group of RGs; this may be indicated by their differentiation age (day 20), and the lack of notch signalling produced by neurons (Kageyama and Ohtsuka, 1999) may account for the lack of *HES5* in these clusters.

Further progenitor-related genes were identified by dividing them into specific stages of the cell cycle. For the mitotic phase of the cell cycle (D'Avino, 2017; Ding et al., 2017; Jackman et al., 1995; Sun and Kaufman, 2018; Whitfield et al., 2002), these clusters (Dividing IPC C10, IPC C15, IPC C20 and IPC C22) were expressing mitotic markers strongly, which was in line with their identity. Although C5 and C12 were predicted to be dividing RG, they were not expressing mitotic genes strongly (Figure 5.9); however, they expressed S phase (DNA replication) genes strongly (Morris and Mathews, 1989; Whitfield et al., 2002; You et al., 1999), demonstrating their identity in cell cycle. In addition, dividing IPC C10, RG C19 and Div IPC C22 were expressing DNA replication genes (*MCM6*, *MCM2*, *PCNA* and *MCM5*) (Morris and Mathews, 1989; Whitfield et al., 2002; You et al., 1999), thus confirming their progenitor identity. For the cell-growth phase (G1 and G2) (Astarita et al., 2012; Brekken and Sage, 2000; Wianny et al., 1998), most clusters were expressing these genes (*SPARC*, *CCND1* and *CCND2*), mostly devoid of neuronal clusters (both newborn and developing) and cluster C9 IPC-derived neurons (Figure 5.9).

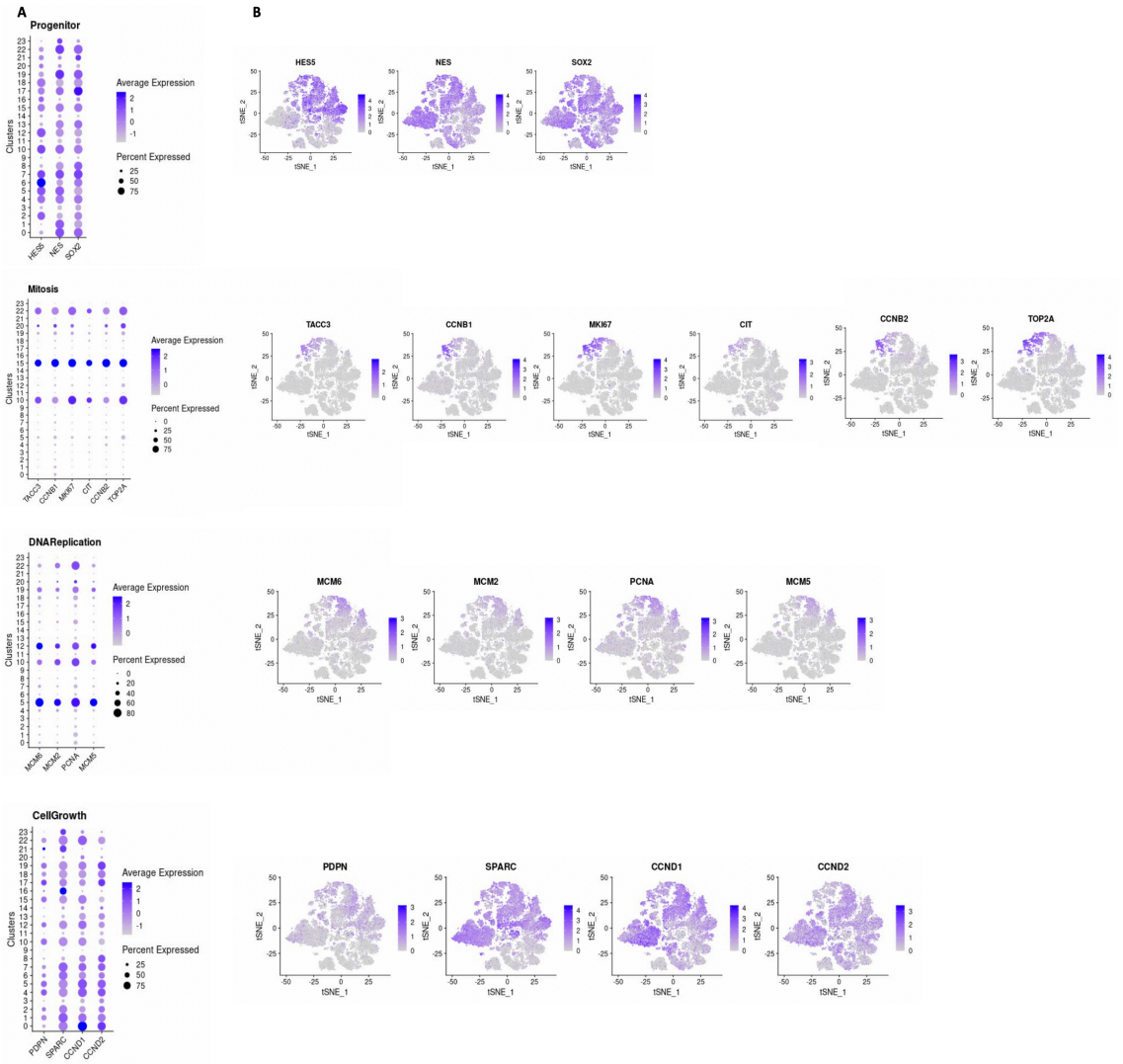


Figure 5.9. Results containing the plots with the expression of the markers in clusters (progenitors and in-cell cycle genes). A) Dot plot visualization: Dot size encodes the percentage of cells within a class, while the colour encodes the AverageExpression level across all cells within a class (blue indicates a high level of expression). **B)** Expression plot visualisation: The cells are coloured on the t-SNE plot according to their gene expression (blue indicates a high level of expression).

Then, genes expressed in neurons were investigated in relation to their function (growth cone, cytoskeleton, neuronal, neurotransmission, and migration). Among the growth-cone-related genes (*CAP1*, *GAP43*, *CYFIP1* and *SNAP25*) (Fishman, 1996; Kimura et al., 2003; Nozumi et al., 2009; Osen-Sand et al., 1993; Schneider et al., 2021) *GAP43* was most strongly and abundantly expressed in neuronal clusters (C13, C18, C11) and RG clusters such as C1 (Figure 5.10). This may be because GAP43 is upregulated in differentiating glia (Mani et al., 2001) and possibly because RGs possess bidirectional radial fibres with growth cone-like endfeet (Rigby et al., 2020).

Then, neuronal cytoskeleton-related genes (*STMN1*, *MAP2*, *TUBB*, *NEFM*, *STMN2*, *TUBB3*, *MAPT*, *NEFH* and *DCX*) (Benarroch, 2021; Himi et al., 1994; Lariviere and Julien, 2004; Lasser et al., 2018; Riederer et al., 1997) were investigated. Strong

expression occurred in neuronal clusters (C3, C13, C8, C18, and C11) (Figure 5.10). Following this, synaptic genes (*NLGN2*, *GPHN* and *SYP*) (Chubykin et al., 2007; Eastwood et al., 2000; Fritschy et al., 2008; Liebau et al., 2007; Tretter et al., 2012) were explored. *NLGN2* was poorly expressed but the synaptophysin gene (*SYP*) was moderately expressed (although a lower proportion of cells expressed the gene than in other clusters) in cluster C9, which was an IPC-derived neuron cluster (synaptophysin gene) (Figure 5.10). However, the canonical neuronal gene NEUN (*RBFOX3*) (Duan et al., 2016) showed expression in neuronal clusters (C3, C13, C8, C18 & C11) rather than progenitors' clusters, thus supporting the neuronal identity of these clusters.

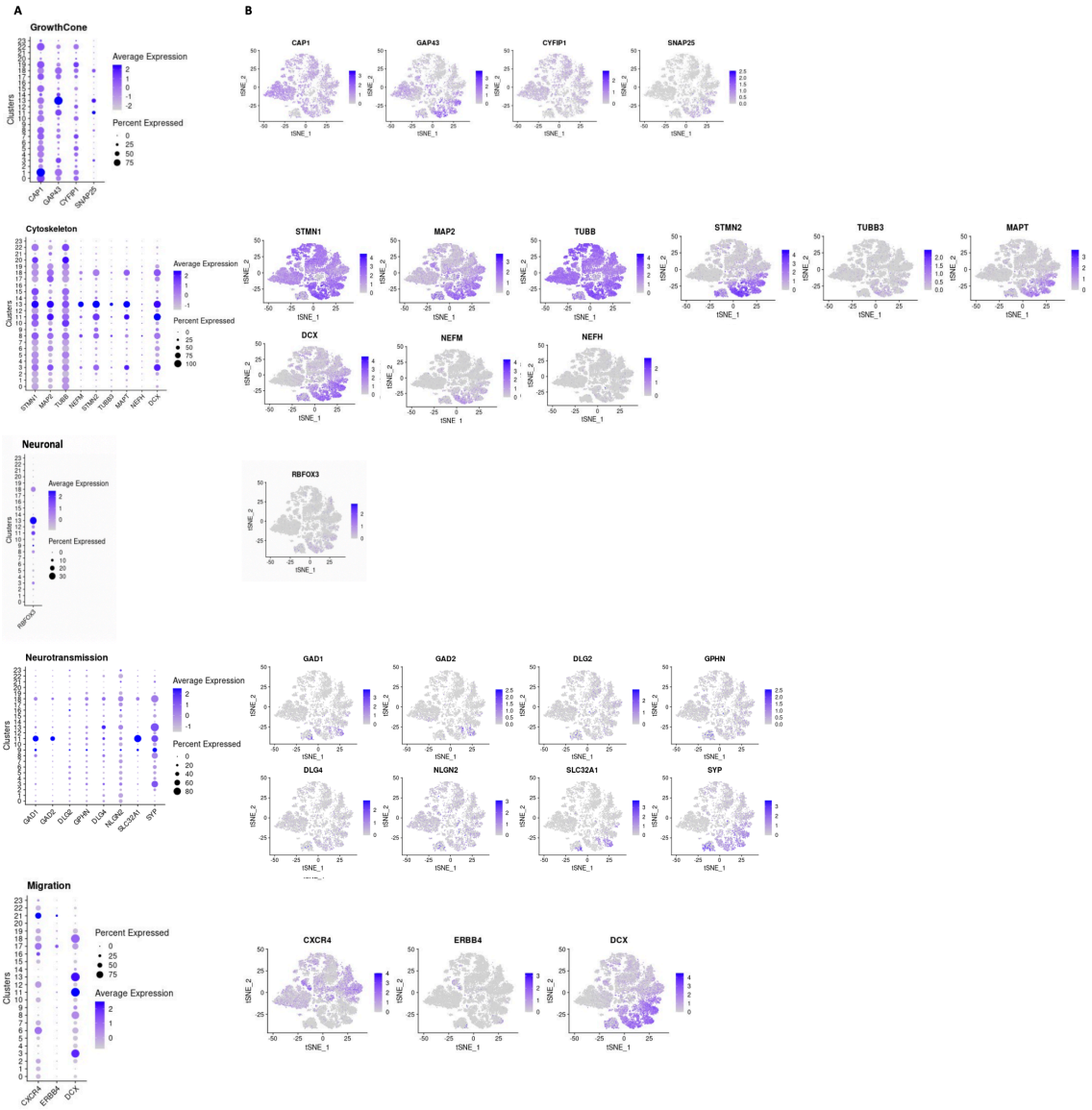


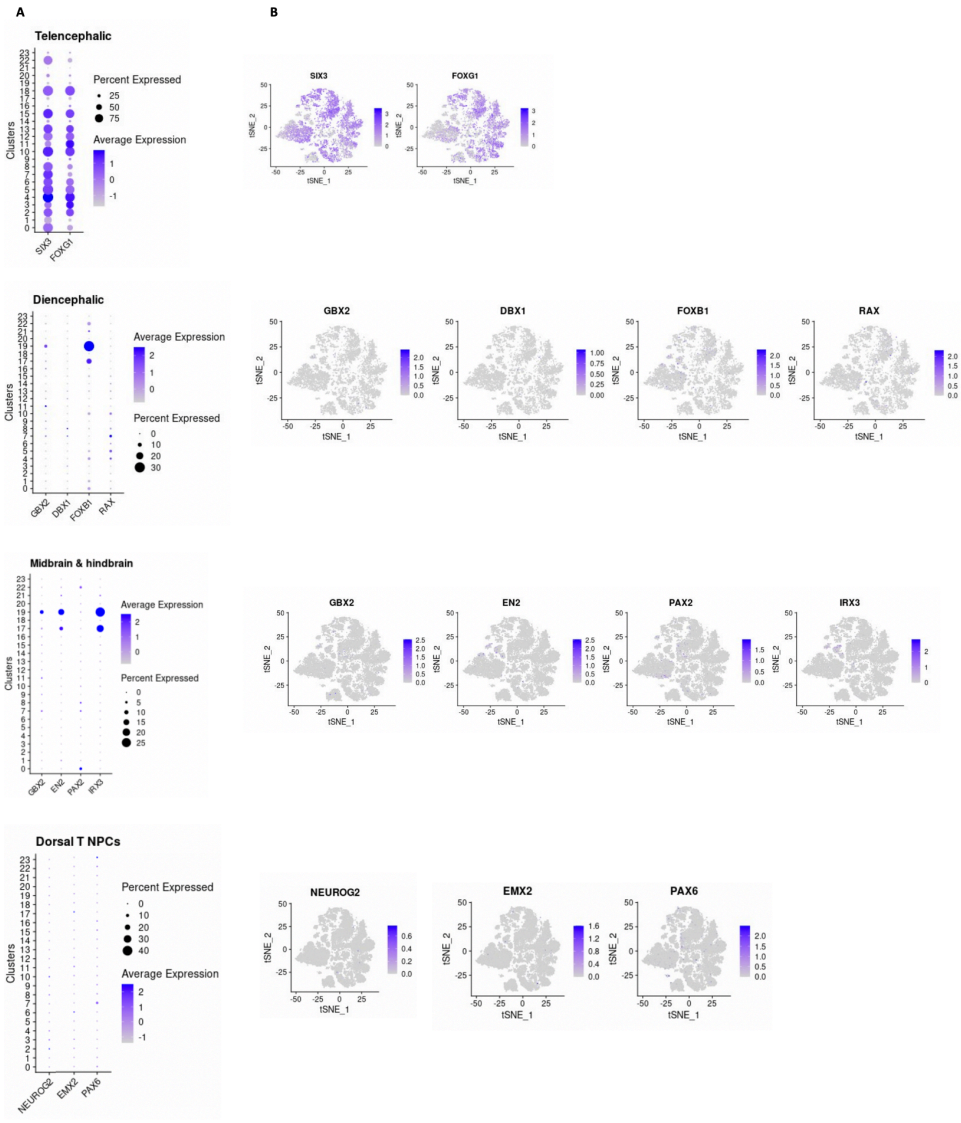
Figure 5.10. Results containing the plots with the expression of the markers in clusters (neurons genes). A) Dot plot visualization: Dot size encodes the percentage of cells within a class, while the colour encodes the AverageExpression level across all cells within a class (blue indicates a high level of expression). **B)** Expression plot visualisation: The cells are coloured on the t-SNE plot according to their gene expression (blue indicates a high level of expression).

GABAergic neurotransmission genes (*GAD1*, *GAD2* and *SLC3A1*) (Asada et al., 1996; Asada et al., 1997; Gasnier, 2004) were likely picking up more mature neurons with relatively well-developed pre-synaptic terminals (Figure 5.10). Again, they showed expression in a few clusters with the most abundant expression in C11. C8 and C18 had no or very low expression of these markers; however, they expressed SST (MGE-derivate, Figure 5.11), indicating the mature developmental status of these clusters. From day 50, some cells in the IPC-derived neuron cluster C9 also showed some expression of *GAD1* (glutamate decarboxylase 1) and *SLC32A1* (vesicular inhibitory amino acid transporter), suggesting the mixture of more developed neurons and less developed cells in the same cluster and possible finer clustering (Figure 5.10).

Regarding migration-related genes (Fazzari et al., 2010; Flames et al., 2004; Friocourt et al., 2007; Stumm et al., 2003; Wang et al., 2011), *CXCR4* and *DCX* showed complimentary expression patterns, with the expression of the former mainly in progenitor clusters and the latter predominantly in neuronal clusters. *ERBB4* expression was much rarer, with some expression in C17 and C21 (Figure 5.10). *DCX* is expressed in newborn neurons, and its function is related to migration (Friocourt et al., 2007). High expression levels were seen in the neuron clusters (C3, C8, C11, C13 and C18) (Figure 5.10), indicating that neurons in this dataset were relatively immature. Most of the clusters were of telencephalic fate, with no obvious expression of markers indicating other fates (diencephalon either thalamic or hypothalamic, mid/hindbrain) (Figure 5.11). *SIX3* and *FOXG1* (Gestri et al., 2005; Hebert and Fishell, 2008; Sousa and Fishell, 2010) were expressed in many clusters, but it was interesting that some clusters expressed no or very low level of *FOXG1* or *SIX3* (like C9, C14, C16, C17, C19 ,C21 and C23) and the same was found for the *NKX2.1* expression pattern (Xu et al., 2010a; Xu et al., 2008), which was less evident in C9, C14, C16, C17, C19, C20, C21 and C23. Thus, it might be that these small clusters were of another origin (e.g. non-telencephalic cells) (Figure 5.11). However, the markers for diencephalic (Orquera et al., 2016; Szabo et al., 2009a; Szabo et al., 2009b), mid/hindbrain expression (Anselme et al., 2007; Close et al., 2017; Rhinn and Brand, 2001) showed that they were not caudally fated except for C17 and C19, which showed expression of caudal genes such as *IRX3* and *EN2*.

Clusters showed no expression for *NEUROG2*, *EMX2* and *PAX6*, which are expressed in the dorsal telencephalic/cortical NPCs (excitatory NPCs) (Englund et al., 2005; Fode et al., 2000; Muzio et al., 2002). Moving on to evaluate the ventral telencephalic genes (*DLX1*, *DLX2*, *DLX5*, *DLX6* and *ASCL1*) (Anderson et al., 1997; Casarosa et al., 1999; Eisenstat et al., 1999; Hansen et al., 2013; Long et al., 2009; Sultan et al., 2013), cells do not express much of these genes in hESCs-derived cortical interneurons dataset in

general but *in-vivo* staining studies on the embryonic mouse brain show strong expression of these genes in ventral telencephalon (Long et al., 2009). *ASCL1* was found to be more abundantly expressed in clusters than *DLX* genes; however, cluster C11 and C18 displayed expression of all 5 genes. This perhaps indicates a more committed/developed stage of differentiation. In addition to the developmental stage, it is plausible that these genes were expressed at such low levels that the current scRNAseq could not reliably detect them due to technical limitations (Chen et al., 2019). The canonical CGE markers *NR2F1* and *NR2F2* (Alzu'bi et al., 2017b; Kanatani et al., 2008) showed expression in certain clusters like C17 and C19. Also, C6 expressed the highest amount of *NR2F1*, which is a transitioning RG (Figure 5.11). These did not overlap with the clusters expressing *DLX* genes, potentially indicating the CGE origin of clusters C17 and C19 as they also showed different expression patterns for telencephalic genes like *FOXG1* and *SIX3* and mid/hindbrain genes like *IRX3* and *EN2*. There are mainly three types of neurons derived from CGE (*LAMP5*, *PAX6* and *VIP*) (Hodge et al., 2019), which showed no expression (Figure 5.11).



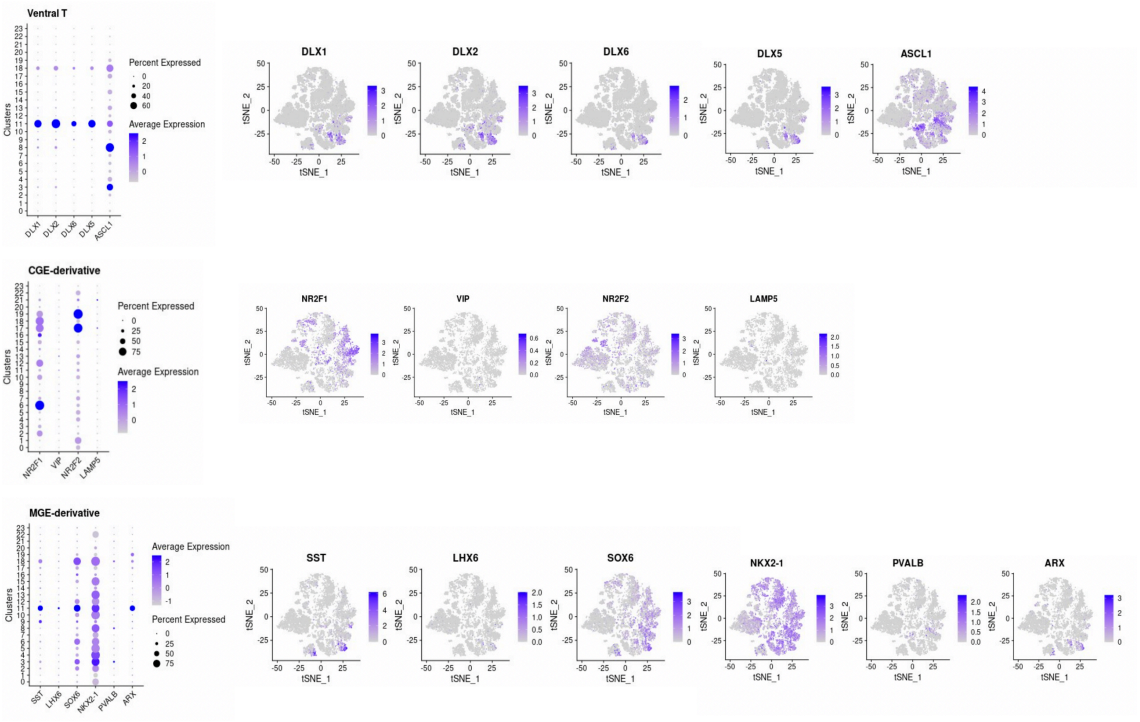


Figure 5.11. Results containing the plots with the expression of the markers in clusters (region identity genes). A) Dot plot visualization: Dot size encodes the percentage of cells within a class, while the colour encodes the AverageExpression level across all cells within a class (blue indicates a high level of expression). **B)** Expression plot visualisation: The cells are coloured on the t-SNE plot according to their gene expression (blue indicates a high level of expression). Dorsal T NPCs is dorsal telencephalic neural precursor cells, Ventral T is ventral telencephalon, CGE- derivative is caudal ganglionic eminence derivative, MGE-derivative is medial ganglionic eminence derivative.

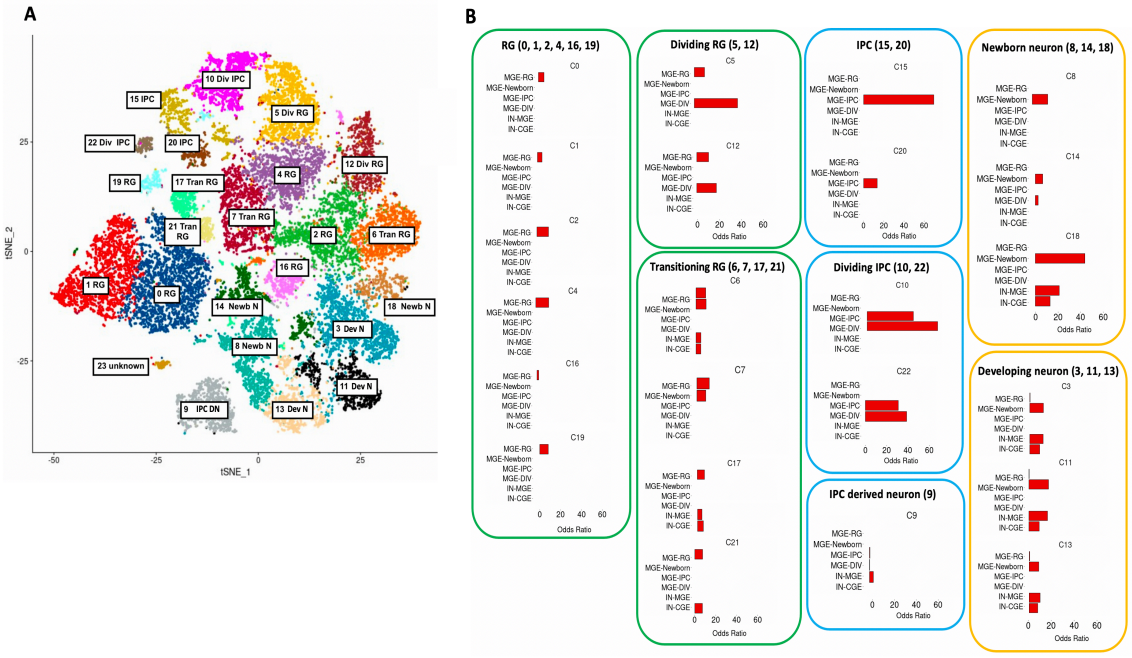


Figure 5.12. Final classification of the clusters. A) Labelled clusters representing the final classes of cells. **B)** The odds ratio from Nowakowski's dataset obtained from the analysis. Eight different broad clusters were identified: RG (radial glia), Div RG (dividing RG), Tran RG (transitioning RG), IPC (intermediate progenitor cells), Div IPC (dividing IPC), IPC DN (IPC-derived neurons), Newb N (newborn neurons), Dev N (Developing neurons).

In conclusion, cluster identities determined by/through the Nowakowski gene set analysis were supported by the GO enrichment and expression of well-known markers. As a result, the clusters were classified into eight higher/broader clusters (Figure 5.12) as follows: RG (C0, C1, C2, C4, C16 and C19), dividing RG (C5 and C12) and transitioning RG (C6, C7, C17 and C21). Also, there were IPC (C15 and C20), dividing IPC (C10 and C22) and IPC-derived neuron (C9), and newborn neuron (C8, C14, C18), developing neurons (C3, C11, C13).

Cluster	Sample 1	Sample 3	Sample 4	Sample 5	Sample 6	Sample 7	Sample 8	Sample 9	Total
0	1052	1250	9	17	14	1	2	0	2345
1	838	1055	8	9	7	0	1	0	1918
2	0	0	16	8	3	319	906	424	1676
3	1	0	11	6	7	593	475	544	1637
4	6	1	634	434	394	11	6	20	1506
5	0	0	498	448	375	12	20	12	1365
6	0	0	1	1	0	467	289	507	1265
7	0	0	118	567	155	99	198	119	1256
8	262	325	150	238	149	30	36	20	1210
9	3	2	21	19	21	316	321	337	1040
10	1	1	214	188	96	168	133	168	969
11	6	5	87	31	56	236	149	336	906
12	0	0	2	1	2	382	172	302	861
13	6	11	212	297	177	31	76	43	853
14	78	74	22	36	32	161	172	153	728
15	38	21	116	112	82	83	83	96	631
16	4	4	10	29	31	133	171	164	546
17	1	0	7	2	6	162	237	125	540
18	0	0	3	0	1	158	107	175	444
19	1	0	87	80	60	11	34	5	278
20	29	21	13	18	17	62	51	51	262
21	0	0	4	0	3	68	94	63	232
22	69	98	1	0	2	1	0	3	174
23	27	53	1	4	3	4	1	1	94

Table 5.6. Number of cells after clustering. Shows number of cells in each cluster for each sample. Sample 1 (Day 20 – $WT^{DLG2+/+}$), Sample 2 (Day 20 – $KO1^{DLG2-/-}$), Sample 3 (Day 20 – $KO2^{DLG2-/-}$), Sample 4 (Day 30 - $WT^{DLG2+/+}$), Sample 5 (Day 30 – $KO1^{DLG2-/-}$), Sample 6 (Day 30 – $KO2^{DLG2-/-}$), Sample 7 (Day 50 - $WT^{DLG2+/+}$), Sample 8 (Day 50 – $KO1^{DLG2-/-}$), Sample 9 (Day 50 – $KO2^{DLG2-/-}$).

For each sample, the total number of cells in each cluster was obtained (Table 5.6). For clusters, it was possible to identify the sample from which the cells originated for WT and KO. This helped to identify the characteristics of the clusters; for example, C0 and C1 were mainly from samples 1 and 3; with 1052, 1250 for C0 and 838, and 1055 for C1 for samples 1 and 3, respectively. These cells were from day 20 samples, and they were RG cells (progenitors).

5.5. DEGs from *DLG2*$^{-/-}$ NSCs/NPCs suggest neurodevelopmental deficits

To investigate the effect of DLG2 KO on human interneuron development, two comparisons between KO and WT were performed. First, a within-cluster comparison was performed for each individual cluster (KO vs WT). As for the previous analyses, the small cluster C23 whose identity was unclear was removed. Second, a comparison was performed on each higher/broader cluster (KO vs WT) (RG, IPC, Dividing RG, dividing IPC, transitioning RG, IPC-derived neuron, newborn neurons, and developing neurons). The latter analysis was performed by Dr Daniel D'Andrea.

The DEGs for each individual cluster (KO vs WT within-cluster) were produced. While some clusters had an acceptable number of DEGs (e.g. C0, C1, and C2) other clusters did not have many DEGs (e.g. C9) (5.13 A). This may be because the cells were very variable (heterogeneous cells) or maybe because as they clustered together, the cells must have been reasonably homogeneous. However, a more obvious explanation is that DLG2 KO did not have such a strong effect on expression within the cell type/cluster.

A

Cluster	Total DEGs in cluster	UP	DN
C0	354	80	274
C1	439	133	306
C2	218	175	43
C3	33	21	12
C4	208	143	65
C5	194	146	48
C6	76	55	21
C7	71	58	13
C8	46	21	25
C9	5	3	2
C10	116	97	19
C11	7	3	4
C12	26	17	9
C13	29	24	5
C14	0	0	0
C15	12	9	3
C16	2	1	1
C17	18	8	10
C18	7	3	4
C19	2	0	2
C20	0	0	0
C21	0	0	0
C22	5	3	2
C23	0	0	0

RG
DivRG
TranRG
IPC
DivIPC
IPC DN
Newb N
Dev N

B

Cluster	Bonferroni	Gene	Avg_LogFC	Regulation
C0	5.42×10^{-72}	PRAF2	-0.300020123	DN
	2.64×10^{-40}	TMLHE	0.122492444	UP
	1.94×10^{-37}	TCEAL5	0.125301801	UP
	2.87×10^{-31}	TIMM17B	0.234815854	UP
	1.65×10^{-23}	RPL39	0.120841207	UP
	3.49×10^{-23}	CSRP1	0.088842909	UP
C1	5.67×10^{-65}	PRAF2	-0.305134271	DN
	3.70×10^{-45}	TIMM17B	0.288371821	UP
	7.46×10^{-40}	TCEAL5	0.148331625	UP
	3.59×10^{-38}	CSRP1	0.175728555	UP
	2.47×10^{-35}	TMLHE	0.103892431	UP
	2.88×10^{-24}	AC004086.1	-0.141150242	DN
C2	6.31×10^{-30}	NR2F1	-0.650018344	DN
	5.66×10^{-25}	S100A10	0.726983152	UP
	1.46×10^{-17}	HES1	0.47540539	UP
	2.20×10^{-16}	FTH1	0.267344355	UP
	5.98×10^{-15}	IGFBP7	0.444754715	UP
	8.19×10^{-15}	SPARCL1	0.497114301	UP
C3	6.32×10^{-12}	SOX2	0.419429752	UP
	2.03×10^{-11}	TCEAL5	0.168505211	UP
	5.85×10^{-11}	NR2F1	-0.367272708	DN
	3.67×10^{-10}	PCDHB5	-0.141229652	DN
	4.14×10^{-10}	ZNF439	0.209239719	UP
	2.42×10^{-09}	TIMM17B	0.262288183	UP
C4	6.53×10^{-64}	GNAS	0.295989669	UP
	1.11×10^{-32}	TIMM17B	0.313860441	UP
	2.01×10^{-32}	TUBA1A	0.362521825	UP
	2.27×10^{-24}	TMSB4X	0.283608748	UP
	6.23×10^{-22}	SPON1	-0.330643144	DN
	1.36×10^{-19}	TMLHE	0.111410887	UP
C5	1.81×10^{-77}	GNAS	0.318174644	UP
	1.77×10^{-36}	TMSB4X	0.42624819	UP
	3.05×10^{-32}	TIMM17B	0.317604041	UP
	1.68×10^{-27}	TUBA1B	0.332075172	UP
	9.08×10^{-26}	FTH1	0.224186847	UP
	2.79×10^{-22}	TUBA1A	0.339882918	UP
C6	5.11×10^{-28}	NR2F1	-0.539571292	DN
	1.27×10^{-16}	PCDHB5	-0.18456978	DN
	2.72×10^{-15}	IGFBP7	0.417462923	UP
	5.63×10^{-14}	FABP7	0.395937124	UP
	7.66×10^{-14}	TCEAL5	0.157923118	UP
	9.93×10^{-13}	GPC3	0.430551077	UP
C7	4.05×10^{-21}	ARMCX6	-0.293591418	DN
	2.46×10^{-14}	TIMM17B	0.315505514	UP
	2.05×10^{-13}	FBLN1	0.487869196	UP
	1.37×10^{-12}	PCDHB5	-0.163758839	DN
	1.29×10^{-10}	SHH	-0.568057253	DN
	6.21×10^{-10}	SLC2A1	0.489854593	UP
C8	6.63×10^{-14}	TIMM17B	0.284969101	UP
	1.66×10^{-11}	PCDHB5	-0.12588515	DN
	4.86×10^{-09}	TMLHE	0.113651337	UP
	1.36×10^{-08}	MT-CO1	-0.158033584	DN
	2.29×10^{-07}	TCEAL5	0.098200108	UP
	2.79×10^{-07}	PRAF2	-0.217915945	DN

Cluster	Bonferroni	Gene	Avg_LogFC	Regulation
C9	3.92×10^{-05}	RPS3A	-0.382336989	DN
	6.97×10^{-04}	INPP5F	0.508376803	UP
	2.76×10^{-03}	ZNF439	0.454419291	UP
	6.69×10^{-03}	VWC2	-0.354407269	DN
	2.68×10^{-02}	ISLR2	0.426291124	UP
C10	1.47×10^{-15}	INPP5F	0.231989224	UP
	7.94×10^{-15}	TIMM17B	0.256080525	UP
	1.22×10^{-14}	FABP7	0.590817594	UP
	1.60×10^{-13}	TUBA1B	0.225486973	UP
	4.97×10^{-12}	VIM	0.306603737	UP
	1.02×10^{-11}	SLC16A3	0.176702857	UP
C11	2.74×10^{-11}	PCDHB5	-0.154380464	DN
	4.13×10^{-10}	ZNF439	0.282796731	UP
	8.01×10^{-10}	TCEAL5	0.283338171	UP
	2.32×10^{-03}	LHX8	-0.307422898	DN
	2.33×10^{-03}	ATP11C	-0.238403524	DN
	4.94×10^{-03}	TIMM17B	0.282781809	UP
C12	8.89×10^{-15}	AMOT	-0.521541139	DN
	2.17×10^{-10}	NR2F1	-0.486599614	DN
	3.19×10^{-09}	INPP5F	0.229953599	UP
	5.57×10^{-08}	MT-ND1	-0.193952059	DN
	1.02×10^{-06}	TMLHE	0.088051308	UP
	2.05×10^{-06}	MT-ND6	-0.221791213	DN
C13	6.15×10^{-11}	ARMCX6	-0.198554243	DN
	5.70×10^{-09}	PCDHB2	0.237606649	UP
	2.07×10^{-08}	PCDHB5	-0.134037711	DN
	2.54×10^{-07}	GNAS	0.184675131	UP
	3.25×10^{-07}	TIMM17B	0.242830537	UP
	2.85×10^{-06}	ZNF439	0.178351249	UP
C15	7.86×10^{-05}	FABP7	0.404913021	UP
	5.65×10^{-04}	VIM	0.303049777	UP
	2.15×10^{-03}	ETV4	0.142584696	UP
	2.46×10^{-03}	IGDCC3	-0.268599819	DN
	2.53×10^{-03}	TUBA1B	0.176600208	UP
	7.43×10^{-03}	TMSB4X	0.262149725	UP
C16	0.01067823	RPL35	-0.445004183	DN
	0.01739517	SIX3	0.662502845	UP
C17	3.48×10^{-09}	PCDHB5	-0.218627484	DN
	1.45×10^{-07}	RHOBTB3	0.620331519	UP
	1.69×10^{-07}	ARMCX6	-0.250232827	DN
	9.72×10^{-07}	CDKN1A	-0.407135041	DN
	2.48×10^{-04}	NRP1	0.315502672	UP
	2.34×10^{-03}	UPF2	0.239525454	UP
C18	1.33×10^{-10}	NR2F1	-0.588617879	DN
	4.52×10^{-08}	TCEAL5	0.195490916	UP
	7.23×10^{-05}	PCDHB5	-0.14224219	DN
	7.08×10^{-04}	MT-ND1	-0.216496567	DN
	3.92×10^{-03}	MARCKS	-0.159648358	DN
	4.76×10^{-03}	PQBP1	0.196122634	UP
C19	0.000128312	PCDHB5	-0.16970306	DN
	0.025340742	ANXA1	-0.257465061	DN
C22	0.0016194	FHL1	0.321646529	UP
	0.00556639	UBE2C	-0.427297108	DN
	0.00981725	TIMM17B	0.270250509	UP
	0.02472089	TCEAL5	0.210230104	UP
	0.02608057	PRAF2	-0.265906587	DN

Figure 5.13. Genes differentially expressed between KO (*DLG2*$^{-/-}$) and WT. A) Shows number of up- and down-regulated genes in each cluster (KO1 & KO2 combined) compared to all the others. Some clusters showed a high number of DEGs while others had a low number of DEGs or no DEGs expressed. **B)** Displays the top six significantly DEGs in each cluster (Bonferroni p < 0.05). C14, C20, C21, and C23 did not show any DEGs. DEGs are the differentially expressed genes, UP: up-regulated genes, DN: down-regulated genes. RG (radial glia), DivRG (dividing RG), TranRG (transitioning RG), IPC (intermediate progenitor cells), DivIPC (dividing IPC), IPC DN (IPC-derived neurons), Newb N (newborn neurons), and Dev N (Developing neurons).

First, the analysis was performed considering each cluster (all 22 clusters) (Figure 5.13 A). In contrast to the WT cells, within the KO, numerous genes demonstrated notable

differential expression within the clusters from days 20, 30, and 50 (Bonferroni $p < 0.05$), with the largest DEGs on day 20 (C0 354 DEGs, C1 439 DEGs) (Figure 5.13A). This encompassed DEGs that exhibited both upregulation and downregulation. These results offer clear evidence of gene expression dysregulation in the *DLG2* KO samples over the duration of hESC differentiation. This encompasses all the developmental stages examined, not simply the latter stages involving more advanced and electrically functional neurons, but also the preceding stages comprising immature neurons and neural progenitor cells. Also, the most significant impact of DLG2 KO was on progenitor cells (day 20) compared to neurons (day 50), which demonstrates some sort of recovery––at least at the RNA level—took place at later developmental stages.

The top DEGs within KO cells (Figure 5.13 B) encode proteins that are major actors in/for cortical development regulatory processes, differentiation, and specification (*LHX8*, *SOX2* and *SIX3*) (Appolloni et al., 2008; Ferri et al., 2013; Flandin et al., 2011), mitochondrial genes (*MT-ND1*, *MT-ND6* and *MT-CO1*) (Carelli and Chan, 2014), transcriptional regulators (*NR2F1*, *HES1*, *SOX2* and *ETV4*) (Alzu'bi et al., 2017b; Fontanet et al., 2018; Nakamura et al., 2000; Zhang and Cui, 2014), cell cycle progression (*CDKN1A*) (Kippin et al., 2005) and microtubules of the cytoskeleton functioning (*TUBA1A*) (Lasser et al., 2018) (Figure 5.13). The above findings give an indication of the type/extent of cellular processes regulated by *DLG2* during early brain development (such as regulation of mitochondrial genes, differentiation, and specification). Also, the above findings confirm the role of *DLG2* during development as DEGs found are known for their role in processes during the development and specification of cortical interneurons. Moreover, key genes related to Chapter 3 and 4 were checked for expression in WT and KO clusters (Figure 5.14)

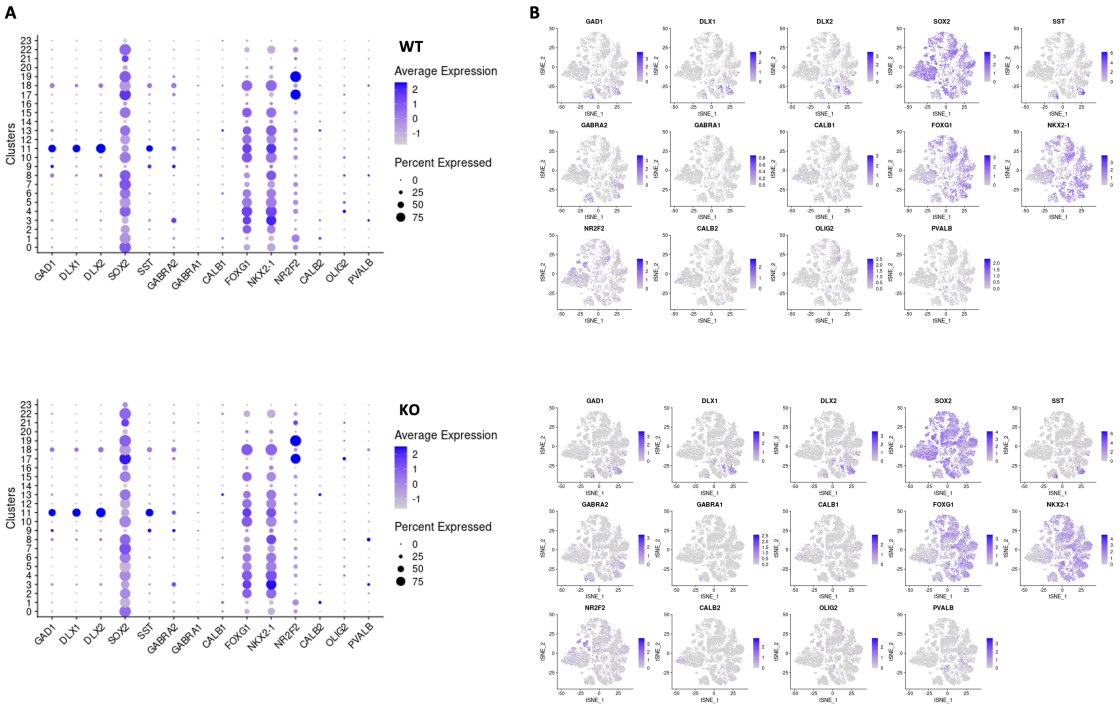


Figure 5.14. Results containing the plots with the expression of the genes in WT and KO clusters (Chapter 3 and 4 key genes). A) Dot plot visualization: Dot size encodes the percentage of cells within a class, while the colour encodes the AverageExpression level across all cells within a class (blue indicates a high level of expression). **B)** Expression plot visualisation: The cells are coloured on the t-SNE plot according to their gene expression (blue indicates a high level of expression).

A

Group	DEGs	UP	DN
RG	3253	221	3032
Dividing RG	223	180	43
Transitioning RG	470	393	77
IPC	13	7	6
Dividing IPC	110	94	16
IPC derived neuron	1	0	1
Newborn Neuron	48	15	33
Developing Neuron	72	49	23

B

Cluster	Bonferroni	Gene	Avg_LogFC	Regulation
RG	$1.93x10^{-117}$	PRAF2	-0.3148107	DN
	$2.59x10^{-96}$	TMLHE	0.1533798	UP
	$3.49x10^{-89}$	PCDHB5	-0.1265438	DN
	$1.28x10^{-80}$	TCEAL5	0.1446460	UP
	$1.29x10^{-72}$	TIMM17B	0.3198787	UP
	$6.63x10^{-60}$	IGDCC3	-0.3584225	DN
Dividing RG	$9.46x10^{-32}$	TIMM17B	0.3846841	UP
	$1.94x10^{-30}$	TMSB4X	0.4951220	UP
	$2.38x10^{-30}$	INPP5F	0.3189534	UP
	$6.79x10^{-30}$	GNAS	0.2276952	UP
	$8.27x10^{-28}$	AMOT	-0.5509897	DN
	$1.81x10^{-24}$	TMLHE	0.1288819	UP
Transitioning RG	$2.29x10^{-52}$	NR2F1	-0.9607222	DN
	$1.47x10^{-46}$	PCDHB5	-0.2592927	DN
	$1.33x10^{-30}$	S100A10	0.7241220	UP
	$1.24x10^{-29}$	CHST7	0.5266651	UP
	$2.73x10^{-29}$	IGFBP7	0.5788681	UP
	$1.77x10^{-27}$	MYL9	0.4912560	UP
IPC	0.000122	FABP7	0.7104334	UP
	0.001427	IGDCC3	-0.3631768	DN
	0.001588	PCDHB5	-0.1544961	DN
	0.001869	TMLHE	0.1360956	UP
	0.006026	TIMM17B	0.3455725	UP
	0.007150	CRABP2	-0.3588974	DN
Dividing IPC	$3.23x10^{-20}$	TIMM17B	0.3727836	UP
	$8.46x10^{-14}$	FABP7	0.7775685	UP
	$1.71x10^{-13}$	TMLHE	0.1181637	UP
	$3.38x10^{-13}$	SLC16A3	0.2663035	UP
	$1.32x10^{-12}$	INPP5F	0.2810023	UP
	$3.16x10^{-12}$	PCDHB5	-0.1425812	DN
IPC derived neuron	0.009448	BCL7B	-0.6095624	DN
Newborn Neuron	$2.65x10^{-20}$	PCDHB5	-0.1691659	DN
	$1.98x10^{-18}$	TCEAL5	0.2029922	UP
	$2.76x10^{-14}$	TIMM17B	0.3487274	UP
	$3.52x10^{-13}$	TMLHE	0.1332458	UP
	$1.19x10^{-10}$	PRAF2	-0.2635217	DN
	$2.63x10^{-08}$	ZNF439	0.1370155	UP
Developing Neuron	$2.97x10^{-34}$	PCDHB5	-0.2134861	DN
	$2.42x10^{-32}$	ZNF439	0.3192333	UP
	$1.29x10^{-30}$	TCEAL5	0.2709887	UP
	$5.81x10^{-27}$	TIMM17B	0.3943787	UP
	$5.78x10^{-23}$	PQBP1	0.3406668	UP
	$3.99x10^{-21}$	LRRC61	0.1635524	UP

Figure 5.15. Genes differentially expressed between KO (*DLG2*$^{-/-}$) and WT in Groups. A) Shows the number of up- and down-regulated genes in each group (KO1 and KO2 combined) compared to all the others. **B)** Displays the top six most significant DEGs in each group (Bonferroni $p < 0.05$). DEGs: differentially expressed genes, UP: up-regulated genes, DN: down-regulated genes, RG (radial glia), Dividing RG (dividing radial glia), Transitioning RG (transitioning radial glia), IPC (intermediate progenitor cells), Dividing IPC (dividing intermediate progenitor cells), IPC derived neuron (intermediate progenitor cell-derived neurons).

Then, the clusters were grouped together in eight higher clusters (RG, Dividing RG, Transitioning RG, IPC, Dividing IPC, IPC derived neuron, Developing neuron, and Newborn neuron) and differential expression analysis was performed between WT and KO cells for each group (Figure 5.15 A and B). At the group level, approximately 4190 genes demonstrated differential expression in the KO samples in contrast to the WT, which was similar to the cluster-level results, repeatedly demonstrating that DLG2 participates in the governance of cortical neural development. Also, like the cluster-level DEGs, the RG group demonstrated the highest number of DEGs (3253 DEGs) confirming the role of DLG2 during progenitor development or very early stages of brain development. The top six DEGs within the KO cells in these groups were similar to those observed from specific cluster analysis, e.g. *NR2F1* encodes a major neurodevelopmental transcriptional regulator (Alzu'bi et al., 2017b) (Figure 5.15 B).

However, some genes were present in group DEGs that were not observed in cluster-level DEGs i.e. *BCL7B*.

In order to understand the gene expression data in more detail and to elucidate the part played by *DLG2* in cortical interneuron development, the DEGs for each stage were analysed using Fisher's exact tests for functional (GO term) enrichment. Such work facilitated specific characteristics to be allocated to the DEGs through the recognition of functional gene categories, i.e. those linked with a particular molecular function (MF) or cellular component (CC) or biological process (BP) that were enriched within DEGs for each group. For the clusters evaluated, DEGs with a Bonferroni $p < 0.05$ were separated into two gene sets, i.e. those that were upregulated and those that were downregulated. Functional enrichment analysis employing terms from GO sources was carried out on each set. For the purposes of the current analysis, the background gene set comprised all genes that were expressed in one or more clusters in both KO and WT samples. Considering Bonferroni $p < 0.05$ as the threshold, the outcome results detected significantly enriched classes in numerous GO terms up- and down-regulated genes in the clusters. A refinement method was performed as mentioned in Chapter 2 (Section 2.12).

At the cluster level, when all clusters were tested for GO enrichment, enriched GO terms (Bonferroni $p < 0.05$) were identified for clusters C0, C1, C2, C4, C5, C6, and C10 in upregulated and downregulated gene sets (apart from clusters 6 and 10, which showed no enriched terms for down-regulated gene sets) (Table 5.7). In particular, the GO upregulated terms observed were related to *cytosolic large/small ribosomal subunit, collagen catabolic process, endodermal cell differentiation, focal adhesion, negative regulation of macromolecule metabolic process and negative regulation of developmental process.* The GO downregulated terms observed were related to *cholesterol biosynthetic process, unfolded protein binding, RNA splicing and binding, molecular transducer activity, translational initiation, peptide biosynthetic process and negative regulation of transcription* (Table 5.7). The above findings show that DLG2 knockout does affect major biological processes during interneuron development. Also, DLG2's regulated biological processes are not limited to interneurons but are present in neural precursor cells.

Cluster	set_N	GO Term (up-regulated genes)	ann_N	N ^{overlap}	OR	P	P ^{corrected}
C0 (RG)	80	cytosolic large ribosomal subunit	58	15	87.74819	1.17×10^{-22}	5.12×10^{-19}
C0 (RG)	80	cytosolic small ribosomal subunit	40	7	47.65982	7.79×10^{-10}	3.39×10^{-06}
C1 (RG)	133	focal adhesion	352	27	12.58906	4.16×10^{-19}	1.81×10^{-15}
C1 (RG)	133	platelet aggregation	25	9	74.11743	1.96×10^{-13}	8.55×10^{-10}
C1 (RG)	133	cell leading edge	223	14	9.108860	3.26×10^{-09}	1.42×10^{-05}
C1 (RG)	133	collagen catabolic process	49	6	17.95285	2.54×10^{-06}	0.011090

C1 (RG)	133	ATP metabolic process	152	9	8.247102	$3.82x10^{-06}$	0.016629
C1 (RG)	133	structural constituent of muscle	34	5	22.03868	$7.21x10^{-06}$	0.031371
C1 (RG)	133	endodermal cell differentiation	35	5	21.29589	$8.35x10^{-06}$	0.036364
C2 (RG)	175	platelet degranulation	107	12	12.59425	$1.40x10^{-09}$	$6.11x10^{-06}$
C2 (RG)	175	endoplasmic reticulum lumen	243	14	6.122046	$3.15x10^{-07}$	0.001372
C2 (RG)	175	secretory granule membrane	238	12	5.253750	$8.90x10^{-06}$	0.038736
C4 (RG)	143	SRP-dependent cotranslational protein targeting to membrane	88	34	93.87774	$2.00x10^{-48}$	$8.73x10^{-45}$
C4 (RG)	143	structural molecule activity	361	41	20.16781	$1.35x10^{-34}$	$5.91x10^{-31}$
C4 (RG)	143	focal adhesion	325	26	10.94368	$3.43x10^{-17}$	$1.49x10^{-13}$
C4 (RG)	143	platelet aggregation	25	5	29.62055	$2.08x10^{-06}$	0.009078
C5 (Div RG)	146	platelet aggregation	25	9	67.04433	$4.61x10^{-13}$	$2.01x10^{-09}$
C5 (Div RG)	146	blood microparticle	79	10	17.37273	$1.81x10^{-09}$	$7.89x10^{-06}$
C5 (Div RG)	146	extracellular matrix	339	16	6.118615	$5.34x10^{-08}$	0.000232
C5 (Div RG)	146	polymeric cytoskeletal fiber	257	12	5.897868	$3.01x10^{-06}$	0.013073
C5 (Div RG)	146	platelet degranulation	107	8	9.533163	$4.76x10^{-06}$	0.020718
C5 (Div RG)	146	negative regulation of macromolecule metabolic process	1821	35	2.577220	$7.15x10^{-06}$	0.031131
C6 (Tran RG)	55	negative regulation of developmental process	466	10	7.805470	$2.69x10^{-06}$	0.011740
C6 (Tran RG)	55	positive regulation of receptor mediated endocytosis	38	4	37.85270	$7.44x10^{-06}$	0.032371
C6 (Tran RG)	55	transition metal ion homeostasis	82	5	21.27338	$7.60x10^{-06}$	0.033070
C10 (Div IPC)	97	cell activation	672	17	5.118063	$3.3x10^{-07}$	0.001454
C10 (Div IPC)	97	polymeric cytoskeletal fiber	257	10	7.530842	$2.66x10^{-06}$	0.011561
C10 (Div IPC)	97	platelet degranulation	107	7	12.69326	$3.10x10^{-06}$	0.013526
C10 (Div IPC)	97	extracellular matrix organization	209	9	8.298241	$3.79x10^{-06}$	0.016511

Cluster	set_N	GO Term (down-regulated genes)	ann_N	N overlap	OR	P	P corrected
C0 (RG)	274	RNA binding	1393	71	3.951150	$3.28x10^{-18}$	$1.43x10^{-14}$
C0 (RG)	274	ribonucleoprotein complex biogenesis	402	25	4.228753	$1.52x10^{-08}$	$6.59x10^{-05}$
C0 (RG)	274	cholesterol biosynthetic process	30	7	18.49009	$4.68x10^{-07}$	0.002034
C0 (RG)	274	retrograde vesicle mediated transport	76	10	9.289770	$4.84x10^{-07}$	0.002105
C0 (RG)	274	proton-transporting ATP synthase complex	20	6	25.92901	$6.26x10^{-07}$	0.002723
C0 (RG)	274	unfolded protein binding	67	9	9.480906	$1.53x10^{-06}$	0.006647
C0 (RG)	274	RNA splicing	271	17	4.166682	$2.93x10^{-06}$	0.012746
C0 (RG)	274	molecular transducer activity	774	0	0	$3.60x10^{-06}$	0.015663
C0 (RG)	274	snoRNA binding	29	6	15.79366	$6.77x10^{-06}$	0.029438
C0 (RG)	274	DNA geometric change	47	7	10.62647	$1.14x10^{-05}$	0.049403
C1 (RG)	306	translational initiation	119	25	15.26051	$1.26x10^{-19}$	$5.48x10^{-16}$
C1 (RG)	306	RNA binding	1393	70	3.340562	$6.20x10^{-15}$	$2.70x10^{-11}$
C1 (RG)	306	cytosolic small ribosomal subunit	40	9	15.82189	$3.75x10^{-08}$	$1.63x10^{-04}$
C1 (RG)	306	maturation of LSU-rRNA	23	6	19.05304	$2.97x10^{-06}$	0.012899
C2 (RG)	43	peptide biosynthetic process	318	13	22.96754	$8.43x10^{-13}$	$3.67x10^{-09}$
C2 (RG)	43	cytoplasmic translation	39	5	63.49025	$5.05x10^{-08}$	$2.20x10^{-04}$
C2 (RG)	43	cytosolic small ribosomal subunit	40	5	61.67934	$5.76x10^{-08}$	$2.51x10^{-04}$
C4 (RG)	65	negative regulation of transcription	821	18	7.468922	$1.54x10^{-09}$	$6.70x10^{-06}$
C5 (Div RG)	48	SRP-dependent cotranslational protein targeting to membrane	88	8	40.92560	$1.48x10^{-10}$	$6.43x10^{-07}$

Table 5.7. Gene ontology (GO) functional enrichment (cluster-level). Shows GO term functional enrichment (Fisher's exact test) in up-regulated (blue) and down-regulated (orange) genes sets that are significantly differentially expressed in each cluster combined KOs (*DLG2*$^{-/-}$) and WT (Bonferroni $p < 0.05$). After recursive refinement by the odds ratio (OR), the resultant enrichment from the unique sets of DEGs was accomplished. Terms arranged by P corrected value. The analysis included all expressed genes from all experimental time-points as a

comparator gene set, Noverlap = unique DEGs number captured by term. RG (radial glia), Div RG (dividing RG), Tran RG (transitioning RG), and Div IPC (dividing intermediate progenitor cells). set_N= number of genes in the DEG set (number of genes up or down regulated in a cluster); ann_N= number of genes in the GO term that are present in the all-expressed set (used as the background set for the test); and the overlap_N= number of genes in the intersection between the GO term and DEG sets.

At the group level, when eight higher clusters (RG, dividing RG, transitioning RG, IPC, dividing IPC, transitioning IPC developing neuron, developing neuron, and newborn neuron) were analysed for GO, RG, dividing RG, Tran RG, and dividing IPC showed enrichment. The functional enrichment data demonstrated that essential developmental mechanisms within KO cells had been dysregulated (Table 5.8). However, three groups did not show any enrichment (IPC, IPC- derived neuron, and newborn neuron). This could be because there were a very small number of DEGs in these groups (13 DEGs for IPC, 1 DEG for IPC-derived neurons, and 48 DEGs for newborn neurons) (Figure 5.15 A). With respect to terms from the GO source, RG group showed downregulation in *G-protein coupled receptor signalling pathway, mitochondrial gene expression, translation regulator activity, regulation of chromosome organization, system process, mitotic cell cycle process* and *translation regulator activity.* While GO enrichment showed upregulation for the *cytosolic large/small ribosomal subunit* term (Figure 5.8). Moreover, transitioning RG genes displayed downregulation enrichment for *central nervous system development, trans-synaptic signaling, synapse and cell-cell signaling*. Genes exhibiting upregulation were enriched for essential signalling mechanisms, i.e. *regulation of substrate adhesion-dependent cell spreading* and *collagen-containing extracellular matrix*. Also, dividing IPC and dividing RG showed upregulation enrichment in *collagen-containing extracellular matrix*. In neurodevelopment, cell-adhesion molecules play a significant role in controlling progenitor behaviour, axonal wiring and pathfinding, synapse formation/function, and neural migration (Geiger et al., 2001; Kerrisk et al., 2014). Cell adhesion plays important role in progenitor division/neurogenesis as cells migrate away from areas with a low concentration of extracellular matrix (ECM) where the adhesion is weak and migrate toward areas with a higher concentration of ECM, where movement will stop or slow if the adhesive strength is too strong. Also, this implies that distinct ECM molecules may play antagonistic roles in the determination of certain cell fates (Rozario and DeSimone, 2010). So, increased/prolonged expression of cell adhesion genes may alter progenitor function/development. In conclusion, the cluster and group-level analysis results reinforce the role played by *DLG2* as a major actor within the signalling pathways governing early neural development and suggest that physiological mechanisms controlled by DLG2 do not only pertain to neurons but are also evident within neural precursor cells.

Cluster	set_N	GO Term (up-regulated genes)	ann_N	N overlap	OR	P	P corrected
RG	221	cytosolic small ribosomal subunit	89	23	28.53972	$1.35x10^{-23}$	$6.29x10^{-20}$
RG	221	cytosolic large ribosomal subunit	110	23	21.62832	$2.64x10^{-21}$	$1.22x10^{-17}$
RG	221	polysomal ribosome	61	14	23.38308	$5.00x10^{-14}$	$2.31X10^{-10}$
RG	221	chemokine binding	22	5	22.15037	$8.94x10^{-06}$	0.0413962
Div RG	180	platelet aggregation	28	8	37.90480	$4.37x10^{-10}$	$2.03x10^{-06}$
Div RG	180	structural constituent of cytoskeleton	69	10	16.22471	$3.57x10^{-09}$	0.0000165
Div RG	180	platelet degranulation	107	10	9.849676	$2.58x10^{-07}$	0.0011959
Div RG	180	collagen-containing extracellular matrix	318	15	4.811153	$2.15x10^{-06}$	0.0099454
Tran RG	392	collagen-containing extracellular matrix	651	54	4.156672	$7.26 x10^{-16}$	$3.36 x10^{-12}$
Tran RG	392	tissue development	838	58	3.417396	$1.61x10^{-13}$	$7.46x10^{-10}$
Tran RG	392	platelet aggregation	41	14	22.07336	$2.79 x10^{-13}$	$1.29 x10^{-09}$
Tran RG	392	structural constituent of cytoskeleton	380	36	4.640476	$1.46 x10^{-12}$	$6.78 x10^{-09}$
Tran RG	392	platelet degranulation	1432	74	2.531443	$1.25 x10^{-10}$	$5.81 x10^{-07}$
Tran RG	392	basement membrane	387	32	3.949832	$8.88 x10^{-10}$	$4.10 x10^{-06}$
Tran RG	392	regulation of heart rate	763	47	2.932854	$1.97 x10^{-09}$	$9.16x10^{-06}$
Tran RG	392	regulation of substrate adhesion-dependent cell spreading	397	30	3.559146	$2.41 x10^{-08}$	0.0001115
Tran RG	392	recycling endosome membrane	54	11	10.80081	$4.46 x10^{-08}$	0.0002064
Tran RG	392	structural constituent of muscle	519	34	3.063606	$9.24 x10^{-08}$	0.0004275
Tran RG	392	immune effector process	664	93	2.740912	$2.67 x10^{-07}$	0.0012358
Div IPC	94	neutrophil degranulation	720	20	6.071548	$3.03x10^{-09}$	0.0000140
Div IPC	94	structural constituent of cytoskeleton	387	15	8.193561	$4.61x10^{-09}$	0.0000213
Div IPC	94	collagen-containing extracellular matrix	318	13	8.481909	$2.89x10^{-08}$	0.0001338
Div IPC	94	platelet aggregation	28	5	39.97739	$4.74x10^{-07}$	0.0021939
Dev N	49	positive regulation of supramolecular fiber organization	151	6	15.70364	$5.33x10^{-06}$	0.0246657

Cluster	set_N	GO Term (down-regulated genes)	ann_N	N overlap	OR	P	P corrected
RG	221	RNA binding	1451	472	2.364204	$1.24x10^{-42}$	$5.74x10^{-39}$
RG	221	mitochondrial gene expression	1735	508	2.018532	$2.85x10^{-32}$	$1.32x10^{-28}$
RG	221	DNA templated transcription, initiation	1256	393	2.186159	$5.51x10^{-31}$	$2.55x10^{-27}$
RG	221	integral component of plasma membrane	1197	90	0.343268	$4.16x10^{-28}$	$1.93x10^{-24}$
RG	221	molecular transducer activity	724	45	0.285121	$9.14x10^{-22}$	$4.23x10^{-18}$
RG	221	ribosome biogenesis	700	227	2.234103	$4.50x10^{-20}$	$2.08x10^{-16}$
RG	221	RNA splicing via transesterification reactions	553	180	2.226328	$2.93x10^{-16}$	$1.35x10^{-12}$
RG	221	regulation of chromosome organization	1732	446	1.642289	$3.80x10^{-16}$	$1.76x10^{-12}$
RG	221	covalent chromatin modification	1092	324	1.988437	$3.50x10^{-21}$	$2.10x10^{-10}$
RG	221	G protein-coupled receptor signaling pathway	247	7	0.130357	$1.54x10^{-13}$	$7.10x10^{-10}$
RG	221	response to unfolded protein	1708	424	1.550746	$1.29x10^{-12}$	$6.00x10^{-09}$
RG	221	mitotic cell cycle process	563	170	1.985703	$3.67x10^{-12}$	$1.69x10^{-08}$
RG	221	snRNA binding	267	94	2.470219	$4.09x10^{-11}$	$1.89x10^{-07}$
RG	221	transcription corepressor activity	1549	381	1.521270	$7.75x10^{-11}$	$3.58x10^{-07}$
RG	221	regulation of translational initiation	62	34	5.468192	$1.13x10^{-10}$	$5.23x10^{-07}$
RG	221	intracellular protein-containing complex	867	247	1.846798	$4.54x10^{-14}$	$9.19x10^{-07}$
RG	221	endoplasmic reticulum to Golgi vesicle mediated transport	175	67	2.808172	$4.00x10^{-10}$	$1.85x10^{-06}$
RG	221	translation regulator activity	103	46	3.641076	$6.21x10^{-10}$	$2.87X10^{-06}$
RG	221	transporter activity	828	91	0.537102	$4.07x10^{-09}$	0.0000188
RG	221	system process	859	98	0.560457	$1.72x10^{-08}$	0.0000796
RG	221	ficolin-1-rich granule lumen	1432	343	1.457363	$2.25x10^{-08}$	0.0001041
RG	221	Transcription elongation from RNA polymerase II promoter	66	32	4.233681	$2.19x10^{-08}$	0.0001015
RG	221	golgi vesicle budding	75	34	3.730882	$6.64x10^{-08}$	0.0003075

RG	221	anterograde trans-synaptic signaling	284	21	0.351831	$1.63x10^{-07}$	0.0007583
RG	221	protein K48-linked ubiquitination	380	110	1.848790	$2.81x10^{-07}$	0.0013031
Div RG	37	SRP-dependent cotranslational protein targeting to membrane	101	12	88.01907	$2.26x10^{-18}$	$1.04x10^{-14}$
Tran RG	72	central nervous system development	1337	23	5.406545	$5.17x10^{-09}$	$2.39x10^{-05}$
Tran RG	72	cell-cell signaling	689	16	6.696980	$3.27x10^{-08}$	0.0001511
Tran RG	72	synapse	599	13	5.963410	$1.63x10^{-06}$	0.0075307
Tran RG	72	trans-synaptic signaling	284	9	8.400418	$3.81x10^{-16}$	0.0176076

Table 5.8. Gene ontology (GO) functional enrichment (group-level). Shows GO term functional enrichment (Fisher's exact test) in up-regulated (blue) and down-regulated (orange) genes sets that are significantly differentially expressed in each group combined KOs (*DLG2*$^{-/-}$) and WT (Bonferroni $p < 0.05$). After recursive refinement by the odds ratio (OR), the resultant enrichment from unique sets of DEGs was accomplished. Terms arranged by P corrected value. The analysis included all expressed genes from all experimental time-points as a comparator gene set, Noverlap = unique DEGs number captured by term. RG (radial glia), Div RG (dividing RG), Tran RG (transitioning RG), Dev N (Developing neurons) and Div IPC (dividing intermediate progenitor cells). set_N= number of genes in the DEG set (number of genes up or down regulated in a cluster); ann_N= number of genes in the GO term that are present in the all-expressed set (used as the background set for the test); and the overlap_N= number of genes in the intersection between the GO term and DEG sets.

5.6. Increased adhesion to the ECM observed in *DLG2*$^{-/-}$ NPCs

GO enrichment analyses identified phenotypes and physiological mechanisms governed by DLG2. In particular, cellular adhesion to the ECM was shown to be dysregulated in the *DLG2*$^{-/-}$ NPCs via the GO term functional enrichment analyses in datasets from both the group and cluster levels. At day 20 (C1), GO term enrichment was seen in DEGs exhibiting upregulation for the terms *focal adhesion* and *collagen catabolic process.* At day 30 (C4, C5 and C10), up-regulated gene enrichment included the terms *extracellular matrix organization* and *focal adhesion.* Also, at group-level enrichment analysis, *collagen-containing extracellular matrix* term has been observed to be up-regulated in dividing IPC, dividing RG, and transitioning RG. According to these data, it could potentially be foreseen that in contrast to WT cells, increased expression of cell adhesion molecules in DLG2 KO should lead to increased ECM adhesion. Adhesion assay was therefore carried out to test this hypothesis.

For the purposes of this assessment, cells from the 25th day of cortical interneuron differentiation were chosen as they were deemed to be comparable to the functional enrichment analyses performed on day 20 and day 30 cells. Plating of the same numbers of WT and *DLG2*$^{-/-}$ cells was carried out onto seven dissimilar ECM substrates. These ECM substrates were collagen I, collagen II, collagen IV, fibronectin, laminin, tenascin, vitronectin and negative control. In brief, after incubating the cells for two hours, cells were washed removing unbound cells to substrates. Then, stain the cells for five minutes, wash cells (three times) and add extraction buffer for 5-10 minutes. The total number of adherent cells was quantified by using colourimetric method (measuring the colour

intensity). In contrast to WT cells, notably enhanced adhesion to the different ECM substrates was demonstrated in the *DLG2*$^{-/-}$ cell lines, particularly with respect to collagen types I and IV. Adherence to collagen type IV showed a significant increase for KO2, and KO1 showed an increased trend towards heightened adhesion (Figure 5.16). This experiment, therefore, confirms the anticipated augmented adhesion and supports the functional enrichment analysis.

In vivo, proteins of ECM bind to cell surface receptors which include integrins (Boraschi-Diaz et al., 2017). This adhesion assay includes ECM targets (Col I, Col II, etc..), from scRNAseq data, thus the integrins receptors linked to this adhesion targets were checked for expression in WT and KO (Figure 5.16 B1 and B2). The ITGB1, ITGA2, ITGA10 integrins related to collagen (I & IV) (Boraschi-Diaz et al., 2017). ITGB3, ITGA2B, ITGA5 and ITGAV integrins related to fibronectin and vitronectin (Danen et al., 2002; Felding-Habermann and Cheresh, 1993). Finally, ITGA3, ITGA6 and ITGA7 relevant to laminin (Stipp, 2010). It is clear from the expression plots that ITGB1 (Col I & IV integrins) is expressed more in KO than WT in C1. From the DEGs analysis (KO vs WT), ITGB1 was upregulated in C1 ($P= 4.237x10^{-08}$ and avg_log2FC=0.22041), C2 and C10 in KO (Supplementary Table 3).

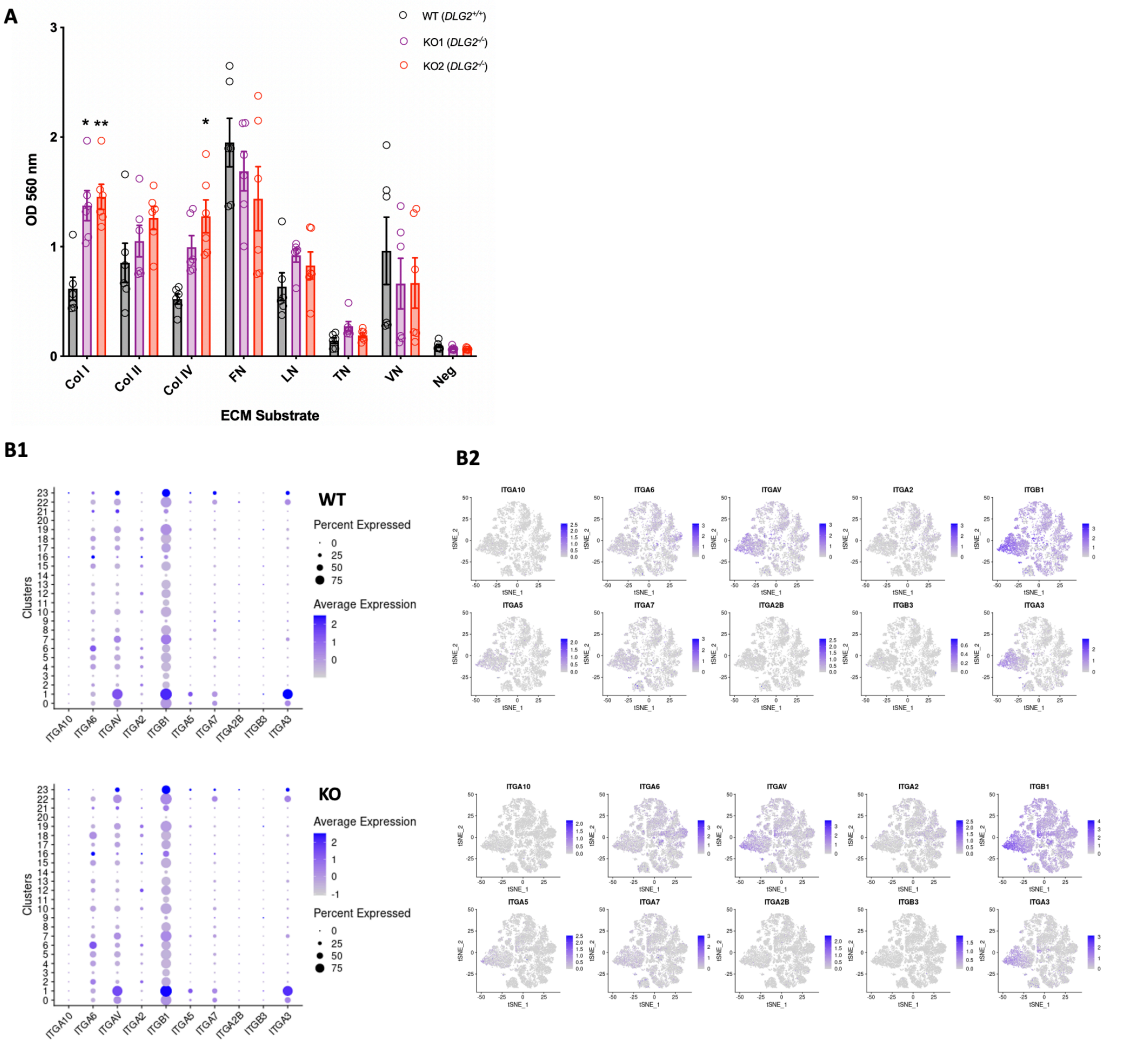


Figure 5.16. Adhesion assay and integrins expression. Day 25 of cortical interneuron differentiation WT and *DLG2*$^{-/-}$ cells adhesion to several extracellular matrix (ECM) protein substrates. On adhesion, both ECM substrate ($F_{7,120}$=35.11; P<0.0001; n=6) and genotype ($F_{2,120}$=3.292; P=0.0406; n=6) showed a significant effect; interaction ($F_{14.120}$= 2.796; P=0.0021; n=6) showed a significant effect. Two-way ANOVA analysis was used. Stars represents corrected-Bonferroni multiple comparison test: P<0.05 (*); P<0.01(**) vs. WT. Col I: collagen I, Col II: collagen II, Col IV: collagen IV, FN: fibronectin, LN: laminin, TN: tenascin, VN: vitronectin and Neg: negative control. **B1/2)** Results containing the plots with the expression of the markers in clusters (integrin genes). **B1)** Dot plot visualization: Dot size encodes the percentage of cells within a class, while the colour encodes the AverageExpression level across all cells within a class (blue indicates a high level of expression). **B2)** Expression plot visualisation: The cells are coloured on the t-SNE plot according to their gene expression (blue indicates a high level of expression).

5.7. Schizophrenia enrichment analysis

Given the association between *DLG2* and SCZ (Chapter 1, Section 1.4.2, 1.5.4 and 1.5.5), from the above scRNA sequencing data, it could be possible to check whether the developmental processes regulated by *DLG2* are relevant to SCZ pathogenesis. This could be investigated owing to the accessibility of relatively large SCZ sample sizes compared to many other disorders (GWAS data) (Pardinas et al., 2018; Trubetskoy et al., 2022). As previous work has revealed DLG2 impacts disease-relevant biology in excitatory corticoneurogenesis (Sanders et al., 2022). Therefore, it was explored

whether DEGs in *DLG2*$^{-/-}$ cells in any of the clusters evaluated demonstrated SCZ common risk variant enrichment, thus establishing whether its impact on inhibitory neuron development is also disease-relevant.

As noted earlier, numerous up- and down-regulated DEGs are present in *DLG2*$^{-/-}$ cells for the identified clusters. A competitive gene set enrichment evaluation was conducted utilising MAGMA to establish whether any of the DEG sets were enriched for SCZ common variant association. The competitive gene set enrichment was carried out on each cluster (cluster-level) and each group (group-level) considering DEGs in KO versus WT cells by employing the MAGMA function conditioned on all expressed genes. The first method showed an enrichment trend in C6-DN transitioning radial glia (Table 5.9); however, there was no robust enrichment in any cluster, and the sets with nominal evidence for association enrichment contained small numbers of genes. Moreover, DEGs at group-level clusters were tested for SCZ association (Table 5.10). However, there was no enrichment for SCZ risk genes (Table 5.10). Furthermore, the DEGs from WT clusters only were tested for SCZ association to determine which aspect(s) of interneuron development was related to SCZ irrespective of whether DLG2 is involved or not (Table 5.11). The results of the analysis displayed enrichment in radial glia (C0) and (C1) down-regulated genes, developing neuron (C11), and (C13) upregulated genes.

VARIABLE	NGENES	BETA	BETA_STD	SE	P	Bonferroni-P
C0 UP	61	0.053825	0.0031069	0.13952	0.34983	1
C0 DN	240	-0.006645	-0.0007571	0.065889	0.54016	1
C1 UP	117	-0.013416	-0.0010708	0.1004	0.55315	1
C1 DN	283	0.010711	0.0013235	0.05879	0.42772	1
C2 UP	148	-0.047975	-0.0043031	0.087299	0.70868	1
C2 DN	36	0.42562	0.018886	0.16842	0.0057557	0.1266254
C3 UP	15	0.48367	0.013862	0.27023	0.03675	0.8085000
C3 DN	8	0.49298	0.01032	0.43438	0.12822	1
C4 UP	120	0.030413	0.0024582	0.092676	0.3714	1
C4 DN	54	0.095304	0.0051769	0.1503	0.26301	1
C5 UP	125	0.076032	0.0062714	0.096672	0.21579	1
C5 DN	33	0.14393	0.0061154	0.17631	0.20716	1
C6 UP	44	0.17982	0.0088197	0.1732	0.14959	1
C6 DN	15	0.88201	0.025278	0.31544	0.0025892	0.0569624
C7 UP	48	0.01431	0.000733	0.15228	0.46257	1
C7 DN	11	0.039459	0.00096853	0.3543	0.45566	1
C8 UP	16	0.00035269	1.04×10^{-05}	0.28148	0.4995	1
C8 DN	20	0.37006	0.012245	0.24763	0.067547	1
C10 UP	77	0.041713	0.002704	0.11845	0.36236	1
C10 DN	13	0.16739	0.0044664	0.25073	0.25219	1
C12 UP	14	0.27661	0.0076591	0.25552	0.13951	1
C13 UP	17	-0.09962	-0.0030393	0.25554	0.65167	1

Table 5.9. Schizophrenia association test for common risk variants (cluster-level). Enrichment of schizophrenia common risk variants for all up-regulated and down-regulated DEGs

in KO (comparative to WT) clusters. Results showed no significant outcome based on Bonferroni p < 0.05. P: p-value, NGENEs: the number of genes, UP: up-regulated genes, DN: down-regulated genes, BETA_STD: standard deviation, SE: standard error.

VARIABLE	NGENES	BETA	BETA_STD	SE	P	Bonferroni-P
Developing Neuron (KO vs WT UP)	38	0.15426	0.0070324	0.17455	0.18842	1
Developing Neuron (KO vs WT DN)	14	0.15529	0.0042998	0.34041	0.32413	1
Dividing IPC (KO vs WT UP)	76	0.011436	0.00073651	0.12427	0.46334	1
Dividing IPC (KO vs WT DN)	12	0.13059	0.0033478	0.26153	0.30878	1
Dividing RG (KO vs WT UP)	154	0.10057	0.0091998	0.088915	0.12903	1
Dividing RG (KO vs WT DN)	34	0.34519	0.014887	0.18504	0.031064	0.434896
IPC (KO vs WT UP)	5	0.64927	0.010746	0.44024	0.07014	0.98196
IPC (KO vs WT DN)	5	0.58672	0.009711	0.42072	0.081585	1
Newborn Neuron (KO vs WT UP)	9	0.1167	0.0025912	0.36746	0.3754	1
Newborn Neuron (KO vs WT DN)	27	0.18342	0.0070505	0.19687	0.17576	1
RG (KO vs WT UP)	185	0.13902	0.013927	0.078756	0.038776	0.542864
RG (KO vs WT DN)	2845	0.03154	0.011442	0.021529	0.071469	1
Transitioning RG (KO vs WT UP)	354	-0.010656	-0.0014697	0.058022	0.57285	1
Transitioning RG (KO vs WT DN)	64	0.3069	0.018144	0.14369	0.016357	0.228998

Table 5.10. Schizophrenia association test for common risk variants (group-level). Enrichment of schizophrenia common risk variants for eight higher groups in up- and down-regulated DEGs in KO (comparative to WT) clusters. Results showed no significant outcome based on Bonferroni p < 0.05. P: p-value. NGENEs: the number of genes, UP: up-regulated genes, DN: down-regulated genes, RG: radial glia, Dividing RG: dividing radial glia, Transitioning RG: transitioning radial glia, IPC: intermediate progenitor cells, and Dividing IPC: dividing intermediate progenitor cells. P: p-value. NGENEs: the number of genes, UP: up-regulated genes, DN: down-regulated genes, BETA_STD: standard deviation, SE: standard error.

VARIABLE	NGENES	BETA	BETA_STD	SE	P	Bonferroni-P
C0_UP	176	-0.037701	-0.0036848	0.078666	0.68412	1
C0_DN	**233**	**0.24616**	**0.027638**	**0.06926**	**0.0001902**	**0.0091296**
C1_UP	422	-0.073745	-0.011084	0.052195	0.92114	1
C1_DN	**348**	**0.21397**	**0.029266**	**0.05678**	**8.24×10^{-05}**	**0.003959664**
C2_UP	95	0.18428	0.013262	0.10769	0.043531	1
C2_DN	261	0.089222	0.010594	0.06656	0.090058	1
C3_UP	348	0.15814	0.02163	0.057902	0.003159	0.151632
C3_DN	379	-0.033892	-0.0048335	0.054863	0.73163	1
C4_UP	73	0.17183	0.010847	0.11188	0.062298	1
C4_DN	176	0.14747	0.014413	0.082976	0.037777	1
C5_UP	187	0.055678	0.0056076	0.072263	0.22051	1
C5_DN	155	0.23654	0.021708	0.088191	0.0036617	0.1757616
C6_UP	142	0.1563	0.013734	0.09652	0.052701	1
C6_DN	207	0.088339	0.0093555	0.074345	0.11738	1
C7_UP	149	0.063807	0.0057423	0.088748	0.23608	1
C7_DN	51	0.28415	0.015001	0.14323	0.023646	1
C8_UP	225	0.19038	0.02101	0.0747	0.0054128	0.2598144
C8_DN	188	0.052936	0.0053454	0.076301	0.24392	1
C9_UP	865	0.036628	0.0077835	0.035843	0.15342	1
C9_DN	1129	-0.00052056	-0.00012542	0.032399	0.50641	1
C10_UP	392	0.031798	0.0046104	0.052471	0.27226	1
C10_DN	132	0.047117	0.003993	0.094476	0.30899	1
C11_UP	**425**	**0.23552**	**0.035524**	**0.053203**	**4.81×10^{-06}**	**0.0002311632**

C11_DN	417	-0.038402	-0.0057386	0.051672	0.77131	1
C12_UP	221	0.13866	0.015167	0.068699	0.021785	1
C12_DN	259	0.06557	0.0077563	0.067353	0.16515	1
C13_UP	**479**	**0.18981**	**0.030347**	**0.049743**	**$6.81x10^{-05}$**	**0.00327048**
C13_DN	316	-0.049299	-0.0064312	0.059767	0.79526	1
C14_UP	203	0.12591	0.013207	0.075999	0.048792	1
C14_DN	624	0.027696	0.0050333	0.041664	0.25312	1
C15_UP	334	0.012866	0.0017247	0.0561	0.4093	1
C15_DN	84	0.13706	0.0092778	0.11857	0.12388	1
C16_UP	540	0.030441	0.0051587	0.044502	0.24698	1
C16_DN	1157	0.034521	0.0084127	0.03189	0.13952	1
C17_UP	446	0.12707	0.019622	0.051419	0.0067367	0.3233616
C17_DN	162	0.049667	0.004659	0.078647	0.26386	1
C18_UP	61	0.35414	0.020442	0.13213	0.0036835	0.176808
C18_DN	76	-0.004685	-0.00030173	0.12862	0.51453	1
C19_UP	241	0.1202	0.013723	0.070995	0.045225	1
C19_DN	31	0.20016	0.008243	0.19134	0.14777	1
C20_UP	144	0.063903	0.0056544	0.087805	0.23338	1
C20_DN	956	0.093334	0.020796	0.034194	0.003175	0.1524
C21_UP	107	0.15795	0.01206	0.10465	0.065628	1
C21_DN	1345	0.059009	0.015419	0.029697	0.023469	1
C22_UP	428	0.04118	0.0062324	0.050963	0.20954	1
C22_DN	19	0.66236	0.021362	0.28646	0.010391	0.498768
C23_UP	6	-0.48266	-0.0087508	0.40507	0.88327	1
C23_DN	818	0.074246	0.015364	0.037396	0.023559	1

Table 5.11. Schizophrenia association test on WT clusters only. Enrichment of schizophrenia common risk variants for WT clusters only in up and down-regulated DEGs. Results showed significant outcome in: C0 DN (RG-Day20), C1 DN (RG-Day20), C11 UP (developing neuron-Day50), C13 UP (developing neuron -Day30); based on Bonferroni $p < 0.05$. P: p-value. NGENEs: number of genes, UP: up-regulated genes, DN: down-regulated genes, BETA_STD: standard deviation, SE: standard error.

5.8. Discussion

5.8.1. A novel function for *DLG2* in cortical inhibitory neural development supported by scRNAseq

Single-cell RNA sequencing was carried out on WT and *DLG2*$^{-/-}$ KOs cells at several stages during the interneuron cortical differentiation process (day 20, 30 and 50). After defining the cluster types/identities (see Section 5.4), DEG analysis was conducted on KO vs WT cells. The results of this work offered additional evidence that *DLG2* is an actor in early cortical inhibitory development. When compared to WT cells, *DLG2*$^{-/-}$ cells exhibited gene expression disturbance, including in the early development phases where neural stem cells, neural precursor cells and undeveloped neurons are the principal cell forms in culture, e.g. C0 and C1 (day 20), and C4 and C5 (day 30), respectively. This finding is similar to a recent study were bulk RNA sequencing analysis showed gene disturbance in early phases of *DLG2*-deficient cortical excitatory neurons (day 15, day 20 and day 30) (Sanders et al., 2022). Also, in the current results, the peak of DEGs, was not identified on the 50th day when developing interneurons (e.g. C3 has 33 DEGs at cluster-level analysis) were evident in the culture and when DLG2 function was established within the postsynaptic density but seen at day 20 when the culture is comprised of a combination of neural stem cells and neural progenitor cells (e.g. C0 has

354 DEGs at cluster-level analysis). This is in support of a study by Sanders *et al.,* where the greater effect of DLG2 deficiency during cortical projection neuron differentiation was on progenitor cells (days 20 and 30, around 14000 DEGs) than neurons (day 60, around 3000 DEGs) (Sanders et al., 2022). Additionally, this finding indicates that some sort of recovery was evident, specifically at the RNA level.

The RG (group-level enrichment analysis) showed downregulation in *transporter activity* term (in the unrefined list, this term includes *metal ion transmembrane transporter activity, metal ion transport*, *cation channel activity, potassium ion transmembrane transporter activity,* and *potassium ion transport*) (Supplementary Table 6), which may suggest the role of DLG2 as a scaffolding channel/receptor complex early in development but acting at an earlier stage at RG before PSD formation. The notion that a DLG protein plays a role in development is not new. Invertebrate Dlg, the human DLG1-4 one orthologue, is recognised as having this activity, governing a range of cellular mechanisms during development pertaining to differentiation, migration and cell-cell adhesion, as a constituent of the Scribble signalling module. Such responsibility for Dlg in humans is attributed to DLG1, which engages in conjunction with SCRIB (loss of function of *Dlg* or *Scrib* cause excessive cell proliferation which provides a strong link to cell cycle progression regulation) (Elsum et al., 2012; Humbert et al., 2008; Stephens et al., 2018). DLG2 and DLG1 belong to the identical MAGUK subclass and therefore have a practically identical domain configuration which are three PDZ domains, one SH3 domain and one GK domain (Zheng et al., 2011). The fact that DLG1 encompasses an L27 moiety is the obvious variation (Funke et al., 2005). This resemblance in configuration together with the results offered by the current work implies that DLG2 and DLG1 inherited some of the developmental regulatory roles attributed with invertebrate Dlg.

5.8.2. Deficiency of *DLG2* affects NPCs at the early phases of interneuron differentiation

The scRNAseq was chosen in order to establish whether a phenotype of *DLG2*$^{-/-}$ was present during interneuron development (NPC stage/ early cells) because a marked variation in NKX2.1 and OLIG2 protein expression (day 20) of the differentiation process was identified via immunocytochemistry analysis. The sequencing analysis revealed numerous DEGs within DLG2-deficient progenitors from earlier stages of cortical interneuron differentiation (day 20) and before the onset of neurogenesis. Also, functional enrichment analysis for GO terms showed phenotypes of *DLG2*$^{-/-}$ present in neural progenitor cells.

Enrichment analysis in RG revealed dysregulation in gene expression and replication i.e. downregulation in *RNA splicing, RNA binding, DNA geometric change (C0), translation initiation (C1), cytoplasmic translation (C2), negative regulation of transcription* (C4) (at the cluster-level). Additionally, at the group level, enrichment analysis showed downregulation in *translation regulator activity, RNA splicing via transesterification reactions, DNA templated transcription* and *initiation, mitotic cell cycle process, response to unfolded protein, transcription corepressor activity,* and *regulation of translational initiation.* It is known that gene regulation is dependent on translational control in eukaryotic cells during development and differentiation (Sonenberg and Hinnebusch, 2009) and that transcriptional regulation is a critical biological process that allows the cell to define cell identity during development, coordinate cellular activity, and respond to a variety of intracellular and extracellular signals (Lee and Young, 2013). Whereas dividing-RG showed downregulation in *SRP-dependent cotranslational protein targeting to membrane,* this term is related to the process for the translation of a protein to its destination membrane (Elvekrog and Walter, 2015). Combining all these terms may cause a delay in transcription factors expression seen at an early point of immunocytochemistry staining in *DLG2* KO cells, encompassing the notable decrease in NKX2.1, OLIG2 (both significant), and increase trend in COUPTF2 expression. Governance of the cell cycle in neural progenitor cells is essential in establishing neuronal type identity, with their intrinsic conditions altering during numerous cell replications and the exact status directly before cell cycle exit, which dictates the produced neuron's ultimate destiny (Petros et al., 2015; Ross, 2011; Sultan et al., 2013; Wonders and Anderson, 2006). NPCs are first present in an 'introverted' format and do not react to extrinsic stimuli. NPCs then transform to the 'extrovert' type that can read the surroundings governed by the cell cycle. When the cell is in this responsive form it is the cell cycle duration that appears to control the cells capacity to appreciate and react to their environs (Ross, 2011; Telley et al., 2019). This heightens the likelihood that decreases in cell cycle duration in *DLG2*$^{-/-}$ NPCs may diminish the duration for which these cells are responsive to extrinsic triggers and thus prolong the onset of the transcription expression regimens necessary for the growth of mature cell identity. Within the nucleus, this mechanism might be augmented by any break in signalling between the intranuclear transcriptional apparatus and the cell surface receptors.

Moreover, functional enrichment analyses of DEGs from day 20 (equivalent to NSCs and NPCs phases) and day 30 (equivalent to immature neurons and NPCs phases) (Figure 5.1) suggested that several developmentally pertinent biological pathways were disrupted. This encompassed ECM adhesion. *DLG2*$^{-/-}$ cells demonstrated augmented adhesion to different extracellular matrix components in contrast to control WT cells. This

is parallel to the effect of *DLG2* deficiency observed on differentiated excitatory NPCs (ECM adhesion was upregulated in *DLG2* KO cell lines) (Sanders et al., 2022). Cell adhesion to ECM mediates cytoskeleton-ECM coupling and is vital for regular development (Frantz et al., 2010; Rozario and DeSimone, 2010). During neurodevelopment, molecules of cell-adhesion play a significant function in controlling progenitor behaviour, synapse formation/function, and neural migration (Geiger et al., 2001; Kerrisk et al., 2014). Throughout cortical differentiation, different programmes of transcriptional regulation suppress cellular ECM adhesion, stimulating cell separation and consequent migration (Geiger et al., 2001; Itoh et al., 2013; Kerrisk et al., 2014). Also, diverse ECM molecules might play antagonistic functions in the determination of certain cell fates (Rozario and DeSimone, 2010). Of note is collagen, which is the most plentiful protein component of the extracellular matrix (Rozario and DeSimone, 2010); it is known to regulate cell adhesion, direct tissue development and form scaffolding that is cross-linked with itself, cell surface receptors, and in conjunction with additional constituents of the ECM (Frantz et al., 2010; Rozario and DeSimone, 2010). Enrichment analysis at the group-level showed upregulation in terms such as *collagen-containing extracellular matrix* in dividing RG, dividing IPC, and transitioning RG; at the cluster-level, upregulation in C1 *collagen catabolic process* term was evident, which may therefore give rise to increased collagen release (*COL1A1*, upregulated within the DEGs from DLG2 KO cells in C1 with P= 0.000168 and avg_log2FC=0.26919) into the ECM creating a more solid scaffold configuration and enhanced adhesion. The collagen pro-a1 chain, a major extracellular matrix element in the human developing cortex (Long et al., 2018), is encoded by *COL1A1*.

At clusters 0 and 1 (day 20), the top DEGs within the KO cells showed downregulation in the *PRAF2* gene, which is known to defer neuroblastoma cell migration and proliferation (Yco et al., 2013). Moreover, PRAF2 is highly expressed in the brain and is known for its role in the regulation of intracellular protein transport, and is associated with the structures of cellular membranes (e.g. intracellular membranes (nucleus) and/or cell surface membranes) (Fo et al., 2006). Moreover, downregulation in the *SHH* gene was observed in C7 (day 30) transitioning RG with $P=1.14x10^{-21}$ and avg_log2FC=-0.5836; this may explain the significant decrease in NKX2.1 observed in day 20 KO cells. Besides, it has been reported that *Shh* is mandatory to maintain the expression of NKX2.1, and therefore the specification of MGE interneuron, throughout embryonic neurogenesis (Flandin et al., 2011; Gulacsi and Anderson, 2006). Also, *WNT2B* is closely linked to WNT signalling, which is downregulated in the RG group (with $P=2.12x10^{-15}$ and avg_log2FC=-0.08681). *Wnt2b* is known to define the cortical hem and is expressed at the medial edge of the telencephalon (Grove et al., 1998).

5.8.3. Several *DLG2*$^{-/-}$ phenotypes suggested cell-type identity delayed expression

The findings of the immunocytochemistry, morphology, and migration analysis documented earlier (Chapter 3) demonstrated that in contrast to the WT cells, some of these properties were disturbed in the *DLG2*$^{-/-}$ cells, e.g. NKX2.1 and OLIG2 expression on day 20, decreased trend in GABA on day 50, and neurite complexity and migration, indicating the delayed development of the identity of the mature cell form. Data from the functional enrichment analysis additionally suggest that in comparison to their WT equivalents, *DLG2*$^{-/-}$ cells were immature, indicated by DEGs pertaining to *central nervous system development* and *negative regulation of developmental process* (GO terms). When considering analysis at group-level, Tran RG GO enrichment analysis showed downregulation in *trans-synaptic signalling, central nervous system development, and synapse and cell-cell signalling.* Also, the RG cluster showed downregulation in *anterograde trans-synaptic signalling.* These signalling terms may provide an indication of the effect of DLG2 KO on NPCs at the early stages of development where cell-cell signalling is downregulated.

There is a further possible association between the heightened cell adhesion to the ECM seen in NPCs and reduced neuron cell migration at day 52 (investigated previously in Chapter 3: the work demonstrated that *DLG2*$^{-/-}$ cells were tardier during the migration process and travelled less far from baseline over a set time interval), e.g. if the ECM is compacted or if the cell-ECM interface is unable to be separated efficaciously this would hinder the capacity of the neuron cells to travel from their original locus.

Moreover, *MARCKS* has been observed to be downregulated in developing neurons (according to the group-level analysis), C18 (newborn neurons: day 50 with P=0.000406 and avg_log2FC=-0.13623) and C3 (developing neurons: day 50 with P=0.03917 and avg_log2FC=-0.11989). This gene has many roles linked to brain growth, neurite outgrowth, synaptic plasticity, and neurotransmitter release (Brudvig and Weimer, 2015). Moreover, *STMN2,* which is considered an axonal and neuronal cytoskeleton gene needed for neurite outgrowth (Benarroch, 2021; Riederer et al., 1997) was downregulated in C8 (with P=0.02659 and avg_log2FC=-0.28452). As all these genes were depleted in developing and newborn clusters, this could further support the previous findings related to migration and diminished secondary neurites.

Also, it was interesting to see downregulation in the *NR2F1* gene (also known as COUPTFI) seen in C2, 3, 6, 12, and 18 (all day 50 cells). *Nr2f1* regulates intermediate progenitor divisions to maintain the delicate balance between MGE- and CGE-derived cIN (Lodato et al., 2011). Moreover, *Nr2f1* controls neuronal cell migration; together with

COUPTF2, *Nr2f1* is needed for the caudal migration of cortical interneurons (Kanatani et al., 2008; Tripodi et al., 2004). *NR2F1* downregulation in C3 (developing neurons) and C18 (newborn neurons) may explain the defects in migration seen in day-52 neurons. Also, *LHX8* downregulation was observed in C11 and in developing neurons with P=0.0120 and avg_log2FC=-0.2601, and likely with *Lhx6*, *Lhx8* is known for its function in GABAergic fate determination (Flandin et al., 2011). This also could explain the decreased trend observed in GABA on day 50 in immunostained cells. Also, upregulation in *SOX2* in cluster 3 (developing neurons, day 50) was observed; *SOX2* is known for its role in inhibiting neural differentiation and maintaining progenitor characteristics (Zhang and Cui, 2014), which may mean that KO cells are still expressing progenitor properties at this stage and did not fully progress into neurons yet. Moreover, *SIX3* was upregulated in C16 and the upregulation of *Six3* expression *in vivo* in embryonic telencephalic progenitor cells retained cells in an undifferentiated state. Also, during mammalian neurogenesis, *Six3* controls the subtle equilibrium between differentiation and proliferation of the progenitor's populations (Appolloni et al., 2008). Therefore, these DEGs might explain a delay in development seen in KO cells.

Also, DEGs in developing and newborn neurons showed downregulation in the expression of mitochondrial genes (*MT-CO3*, *MT-CO1*, *MT-ND3*, *MT-ND1* in developing neurons; *MT-CO1*, *MT-CO2* in newborn neurons (Supplementary Table 5); *MT-ND6*, *MT-ND1*, *MT-ND3* in C3; *MT-CO2* in C8 (Supplementary Table 3)), which means KO cells have minimal energy needs to develop, migrate and mature, which also may be a consequence of their diminished migration and less advanced appearance than WT, with notably diminished branching of the neurites. Alternatively, this could be due to their lower energy requirements, and consequent less complex morphology and less migration. However, it is impossible to determine with certainty which explanation is correct. Also, the mitochondria-associated terms at this more advanced stage of cortical interneuron differentiation (day 30 and day 50 neurons), the neurons within the culture are electrically active and mature (Sanders et al., 2022). Potentially, the diminished energy synthesis within the mitochondria might be connected to the anomalous electrical characteristics of the *DLG2*$^{-/-}$ cells, i.e. the decrease in action potential frequency and maturity (Sanders et al., 2022). This represents an interesting area of study for further research.

5.8.4. Signalling of DLG2 may control numerous signalling pathways throughout cortical interneuron differentiation

GO term functional enrichment analyses have additionally demonstrated that the *DLG2* knockout influences gene expression associated with numerous signalling cascades. At

the group-level analysis, there was downregulation in some signalling pathways, e.g. in the RG-group, there was downregulation in the *G-protein coupled receptor signalling pathway* and *anterograde trans-synaptic signalling*. The latter term refers to neuronal signalling, so it is unclear exactly what this means in the context of RG. It may be the case that some molecules involved in synaptic signalling (such as DLG2) are also components of signalling complexes in progenitors. The G-protein-coupled receptors include the receptors for many neuropeptides, neurotransmitters, and peptide hormones. When G protein is activated, it then transports the signal to an intracellular target (either an ion channel or an enzyme) after detaching from the receptor (Rosenbaum et al., 2009). So, if this signalling pathway is depleted, this may mean that a proper radial glia sensing/response to extracellular cues will not be present at this stage.

Interestingly, there was downregulation in *anterograde trans-synaptic signalling* terms in RG, and *cell-cell signalling* and *trans-synaptic signalling were* downregulated in the transitioning RG group. The presence of these terms may represent a novel finding where signalling processes are present in this early stage. The downregulation may explain that there is a delay in cell-cell communication, which, in turn, will affect development, maturation, and migration.

The results from the mass-spectrometry, immunocytochemistry and scRNAseq analysis suggest that DLG2 expression occurs in early cortical interneuron development and implicates the protein in several signalling cascades associated with development (e.g. TFs linked to the expression of cell-type identity e.g. NKX2.1, and the decrease in *SHH* gene in Tran RG). These are founded on the alterations in different phenotypes and gene expression recognised in *DLG2* knockout neural precursor cells and developing interneurons. Since DLG2 is well-established as a scaffolding protein, these results indicate that it might perform this function in the early stages of interneuron differentiation, maintaining the components of at least one signalling cascade in the proper orientation and/or position for efficacious transduction of signals. If this is the case, DLG2 deficits may not inhibit the function of signalling pathways where it routinely forms a scaffold, as although being less well-organised, the essential components of the signalling pathway may still be present, but not linked together to form complexes (or only partially so). However, this could mean that signalling was more random and more likely to be ineffectual. Such signalling abnormalities could slow (as opposed to stopping) the stimulation of essential transcriptional cascades necessary for normal interneuron development.

5.8.5. Abnormalities in mitochondria revealed by scRNAseq in *DLG2* KO

It is established that neuron cell differentiation is linked with a significant rise in mitochondrial bulk and the copy number of mitochondrial DNA. Also, the disruption in mitochondrial signalling and functions may play roles in neuronal degeneration and impaired neuroplasticity (Cheng et al., 2010). Consequently, further evidence of the part played by *DLG2* was observed in the down-regulated genes from the *DLG2* KO cells from day 20, which demonstrates functional enrichment for terms that are connected to mitochondrial activities. In fact, particular terms pertaining to *mitochondrial transport* and *mitochondrial membrane organisation* from the unrefined results (cluster-level analysis, C0) and *mitochondrial gene expression* (group-level analysis-RG) (Supplementary Table 4 and 6) are present in the gene set that displays downregulation at day 20. Also, from cluster-level DEGs, mitochondrial related genes (*MT-ND1*, *MT-ND6*, *MT-CO1* and *MT-CO3*) from C0, C2, C4, C5 and C10 (Supplementary Table 3) showed downregulation in KO cells. This implies that at this stage, the *DLG2*$^{-/-}$ NPCs only have minimal energy needs, which may be a consequence of diminished *translational initiation;* this term is an additional item that is represented in cluster-level (C1) and group-level analysis-RG. Cortical interneuron development does not appear to be behind schedule in KO cells. However, the fact that in contrast to the WT, these downregulated processes in *DLG2*$^{-/-}$ NPCs could be viewed as implying that there is a hold-up in the manufacture of cellular constituents necessary for later maturation of the neurons, in keeping with the lack of maturity of the *DLG2*$^{-/-}$ neuronal phenotype alluded earlier. This is consistent with altered replication reported previously (given the mitosis term) as the down-regulated replication is associated with decreased mitochondrial gene expression in KO progenitors.

5.8.6. *DLG2*$^{-/-}$ regulated genes are not associated with common schizophrenia risk variants

The synaptic protein *DLG2* plays a role in development: A paper by Sanders *et al*., found that *DLG2* has a role in excitatory neuron development and this synaptic protein appears to regulate processes related to SCZ in those cells (Sanders et al., 2022). Interneurons are also important to SCZ (Akbarian and Huang, 2006; Alherz et al., 2017; Clifton et al., 2019; Dai et al., 2021; Goto et al., 2009; Hashimoto et al., 2008; Kaar et al., 2019; Ma et al., 2018; Murray and Lewis, 1987; Polioudakis et al., 2019; Skene et al., 2018; Trubetskoy et al., 2022; Walker et al., 2019; Yoon et al., 2010) and *DLG2* might have a role in interneuron development. It could potentially be anticipated that the gene sets expressed in *DLG2*$^{-/-}$ and WT cells at different time-points of interneurons differentiation demonstrated enrichment for the common risk variants linked with SCZ given the neural origin of these cells and the fact that SCZ is a neuropsychiatric pathology. Nevertheless,

differentially expressed genes in *DLG2*$^{-/-}$ cells demonstrated no robust enrichment for common risk variants for SCZ. These data are not in keeping with genetic data within the literature suggesting that *DLG2* plays a part in the disease process underlying SCZ (Kirov et al., 2012; Kushima et al., 2022; Nithianantharajah et al., 2013; Sanders et al., 2022). However, *DLG2* may contribute by disrupting excitatory neurogenesis (Sanders et al., 2022) but not inhibitory neurogenesis.

When WT DEGs were tested for SCZ association, it was intriguing to observe that there was robust enrichment for the genetic risk associated with SCZ in the UP developing neuron from day 50 (C11), which comprises the largest population of mature neurons cells, and it was interesting to observe that the robust enrichment present in the UP developing neuron from day 30 (C13), down-regulated genes in radial glia (C0) and (C1) both from day 20 (so the signal is neuronal, not progenitor) that were enriched for association. So, these cell types of interneurons (during development) are related to SCZ regardless of DLG2 involvement. This naturally aligns with the well-recognised neurodevelopment aspect of the causal process underpinning SCZ (Harrison, 1997, 1999; Murray and Lewis, 1987); contemporary research additionally reinforces the part played by a range of risk factors for SCZ in the cortical differentiation process (Walker et al., 2019). Also, the developing neuron-cell type was enriched for SCZ, which confirms the reported link between cortical inhibitory neuron cell types and SCZ (Dai et al., 2021; Skene et al., 2018; Trubetskoy et al., 2022). Therefore, the findings suggest that sets of genes switched on in developing interneurons that harbour SCZ common risk variants but DLG2 do not appear to play a major role in regulating these genes. Perhaps the timing is different for the role of DLG2 during cortical interneuron differentiation. For excitatory neurons, where bulk expression data from the culture were used, the association was reported in young neuron genes downregulated in DLG2 KO cells at day 30 (Sanders et al., 2022), the presented data do have an equivalent time-point in interneurons here (day 30) but there was no association. Maybe for interneurons, timing is more critical during later development when they start to specialise into PV, SST, or other specific types of neurons. Or in case of PV neurons then because they are born a bit later and may take longer to mature (Chu and Anderson, 2015; Xu et al., 2010b), fewer of them are present and so no signal might be evident. Day- 50 neurons are still quite immature and only a relatively small population of cells can be assigned to a specific interneuron subtype, making sub-type specific analyses difficult to perform in this dataset. In terms of culture, this may not be possible because it is unknown if neurons will develop or remain as they are. This needs further investigation.

5.8.7. Summary

The scRNAseq results reported here present the *DLG2*$^{-/-}$ phenotypes from different developmental phases/cell-types throughout cortical interneurons differentiation. Since DEGs between the WT and *DLG2*$^{-/-}$ KO cells are evident at the earliest stages of the process evaluated and before the establishment of mature interneurons, it is not possible for these observations to be the consequence of erroneous synaptic signalling, which suggests innovative function(s) for *DLG2* pertaining to the governance of early cortical interneuron development, together with its well-recognised role within the postsynaptic density. In *DLG2*$^{-/-}$ NPCs, functional enrichment analyses recognised phenotypes encompassed enhanced ECM adhesion, decreased cell-cell signalling, regulation of translational initiation, regulation of chromosome organization and decreased mitochondrial gene expression. Possible mechanisms that give rise to these phenotypes are also proposed by the scRNAseq analysis results, which featured a deferral in the triggering of numerous essential neurodevelopmental transcription regimes, potentially owing to stochastic signalling arising from an impaired DLG2 scaffold associated with the signalling complexes. During human brain development, it is well known that accurate timing is vital: Potentially, simply a transient disturbance in interneuron generation, e.g. deficits pertaining to neuronal branching, gene expression and cell migration, may have an enduring influence on the generation and ultimate functional performance of cortical interneurons networks. Additional research is needed in order to elucidate the precise role of DLG2 during interneuron development. Electrophysiological properties could be investigated in order to check if the voltage-gated channels have been affected by the DLG2 deletion.

The examination of SCZ association was performed via the use of a competitive gene set enrichment test formulated in MAGMA. This enrichment was conditioned on the entire gene population from *DLG2*$^{-/-}$ and WT cell lines and over all stages of the experiment. The *DLG2*$^{-/-}$ gene set exhibited no enrichment for SCZ common risks (their enrichment for rare variants is still to be examined). However, testing WT DEGs for SCZ association indicated that the developing neuron cell type (during development) was related to SCZ irrespective of DLG2 involvement.

6. General discussion

6.1. Research outline

The DLG2 protein is typically expressed in neuronal PSD (Won et al., 2017; Zheng et al., 2011); it regulates extracellular signal transmission to downstream signalling molecules and mediates synaptic signal transmission by interacting with other adaptor proteins (Won et al., 2017; Zhu et al., 2016). These functions of DLG2 are well-recognised in the latter stage of neurodevelopment following synaptogenesis. However, owing to the link between the *DLG2* gene and perturbed excitatory corticoneurogenesis/SCZ (Kirov et al., 2012; Kushima et al., 2022; Nithianantharajah et al., 2013; Sanders et al., 2022), it has been postulated that DLG2-dependent signalling may additionally have a role in the control of early interneuron development; this role is relevant to SCZ because SCZ has an established neurodevelopmental component with risk factors influencing cortical interneurons together with their precursors (Akbarian and Huang, 2006; Alherz et al., 2017; Clifton et al., 2019; Goto et al., 2009; Guidotti et al., 2005; Hashimoto et al., 2008; Kaar et al., 2019; Lewis et al., 2005; Murray and Lewis, 1987; Polioudakis et al., 2019; Trubetskoy et al., 2022; Walker et al., 2019; Yoon et al., 2010). The current work infers a new purpose for DLG2, i.e. the control of early interneuron generation, which was recognised by the elucidation of the properties of WT and KOs (*DLG2*$^{-/-}$) cell lines over the time course of cortical interneuron development from hESCs. This encompassed numerous interrelated methods, such as mass spectrometry, immunocytochemistry, scRNAseq, and adhesion assay.

6.2. DLG2 dependent biological processes during NPC

In the current thesis, the confirmation of the expression of the DLG2 protein during the early stages of interneurons development relied on the use of mass spectrometry as an innovative strategy (DLG2 was absent in both KO lines). This provided evidence of DLG2 expression on the 20th day of the interneuron differentiation process in WT cells. This phase is characterised by the appearance of NPCs within the culture. This is an earlier reported stage than day 30, which has been reported previously on hESCs differentiation of excitatory neurons (Sanders et al., 2022).

Using immunocytochemistry, on day 20, in contrast to WT, *DLG2*$^{-/-}$ cells demonstrated a significant decrease in the expression of NKX2.1 and OLIG2, essential developmental transcription factors within the ventral forebrain, and increased trend in COUPTF2 expression, a CGE transcription factor. This indicates that DLG2 expression (and by implication DLG2-mediated signalling) has a specific function in NPCs. Also, *DLG2*$^{-/-}$ cells on day 50 showed a slight decrease in GABA^{+} cells (not significant). To check if these results were human-cell-specific phenotypes, using three genotypes (WT, *Dlg2*$^{+/-}$

and *Dlg2*$^{-/-}$), a characterisation of *Dlg2* KO embryonic mouse brain was conducted. This was also to outline whether defects in interneuron markers identified *in vitro* correlated with those observed *in vivo*. The findings determined that *Dlg2* heterozygosity showed a decrease trend in the percentage of COUPTF2^{+} cells in the CGE area (not significant). *Dlg2* homozygosity showed a decreasing trend in NKX2.1^{+} and COUPTF2^{+} cells (not significant). Some of these differences were detected in the homozygous (*DLG2*$^{-/-}$) hESCs-derived cortical interneurons, altogether these results imply that *Dlg2* has a role in interneuron development. The support from the mouse-derived data was relatively limited due to the small sample size, which resulted in a limited ability to detect statistical differences.

Of note is that the majority of the information on interneurons derives primarily from murine research; several studies have highlighted the critical differences between human and rodent development (see Chapter 1, Sections 1.17, 1.1.9, & 1.2.4). Essentially, murine data is not entirely comparable with human development. Therefore, in one system (human cells; *in vitro*), single-cell RNA sequencing was utilised to determine genes with differential expression in *DLG2*$^{-/-}$ and WT cells and assess whether those identified genes were enriched for certain phenotypes or biological processes. At day 20, in this early phase of differentiation characterised by NPCs, the over-representation of DEGs within *DLG2*$^{-/-}$ cells pertained to the GO terms encompassing biomechanisms such as cell adhesion. When judged against controls, NPCs from *DLG2*$^{-/-}$ lineage demonstrated augmented adhesion to several ECM substrates. This is similar to the effect of *DLG2* deficiency on cortical excitatory neurons differentiated from hESCs (Sanders et al., 2022), which means that *DLG2* plays a similar role in both excitatory and inhibitory NPCs when considering augmented cell adhesion. Moreover, *DLG2* KO vs WT DEGs in the early phase of neurodevelopment demonstrated an increased expression in genes involved in cell adhesion. The adhesion assay verified the enhanced adhesion of neural precursors in *DLG2*$^{-/-}$ to a range of ECM substrates that encompassed the most frequently found constituent of the ECM, Collagen I (Rozario and DeSimone, 2010). Of note is that the expression of the Collagen I encoding gene, *COL1A1,* additionally exhibits upregulation within the DEGs from *DLG2* knockout cells (in C1/ cluster-level analysis with P=0.000168 and avg_log2FC=0.26919) at day 20. Similarly, at day 15, this gene was one of the top DEGs in *DLG2* KO cortical excitatory neurons (Sanders et al., 2022). NPC *COL1A1* expression *in vitro* has been demonstrated to occur at equivalent phases to day 20 of cortical excitatory differentiation (http://cortecon.neuralsci.org) (van de Leemput et al., 2014). Thus, increased expression of *COL1A1* might imply that elevated cell adhesion in KO cells may be a result of a higher intercellular collagen content, which creates a more solid and increasingly cross-linked collagen in the ECM.

It has been established that the MAGUK protein family are targeted by proteases for breakdown (Ivanova et al., 2011), which leads to weaker cellular binding prior to separation and implies that cell-ECM interactions that arise without the presence of routine scaffolding proteins, e.g. the loss of DLG2 simply impacts expression regulation, leading to an increased expression of ECM/adhesion-related proteins and increased adhesion. Other MAGUKs/DLGs may be involved in stabilising adhesion, therefore loss of DLG2 may not have a direct effect at cell junctions.

Moreover, it was clear that *DLG2*$^{-/-}$ NPCs at day 20 have decreased expression of mitochondrial genes (*mitochondrial gene expression* – downregulated in RG/and downregulation in mitochondrial genes from cluster-level DEGs) and it is known that a significant increase in mitochondrial bulk and copy number of mitochondrial DNA is associated with neuron cell differentiation (Cheng et al., 2010). When merged, the mass spectrometry data, scRNAseq, and cellular experiments persuasively indicated that DLG2-mediated signalling arises during the early phases of cortical interneuron differentiation. However, the entire spectrum of biomechanisms governed by such signalling and the exact pathways involved remain to be elucidated. Given the data suggesting that neuronal subtype identity is dictated by the intrinsic status of NPCs immediately before they leave the cell cycle (Brown et al., 2011; Kepecs and Fishell, 2014; Petros et al., 2015; Sultan et al., 2013; Wonders and Anderson, 2006), it is possible that the disruption of *DLG2*$^{-/-}$ NPC gene expression is responsible for the interneuron phenotypes recognised later in the process.

6.3. Schizophrenia relevant aspects of interneuron generation are not regulated by DLG2-mediated signalling in interneurons

Recent research has indicated that DLG2-regulated processes in excitatory cortical neurons are enriched with SCZ common variants (Sanders et al., 2022). From culture, expression data from bulk RNAseq were used, and the reported association was at day 30 in young neurons genes downregulated in *DLG2*$^{-/-}$ cells (Sanders et al., 2022). In contrast, with respect to interneuron generation, the signalling complexes that engage with the DLG2 scaffold do not appear to control neurogenic transcriptional schedules pertinent to SCZ. DEGs within *DLG2* knockout culture at days 20, 30, and 50 of cortical interneuron differentiation exhibited no enrichment for common risk variants linked with SCZ. This implies that the phenotypes present at these phases may not be immediately pertinent to pathology, despite being intriguing within the developmental situation and are potentially responsible for the neuron cell phenotypes seen later in the process (e.g. migration and morphology defects). However, defects in these phenotypes (migration and morphology) in excitatory corticoneurogenesis are relevant to SCZ as Sanders *et al.*

in the same paper identified GO terms such as migration and morphogenesis from day-30 *DLG2*$^{-/-}$ cells in downregulated genes. This shows that DLG2 deficiency disrupts neurogenesis-related transcriptional programmes, implicating these mechanisms in SCZ pathogenesis (Sanders et al., 2022). This means that similar pathways in interneuron development are not relevant to SCZ (or maybe they are relevant, but DLG2 loss does not significantly impact the SCZ-relevant biology so much in interneurons).

The data presented here do have comparable time-points in interneurons (day 30) but there was no association at the gene-expression level. It is possible that the timing is different for the role of *DLG2* during interneuron differentiation. However, the rare variant link is still to be examined. Of note was that DEGs within cultures comprising a higher percentage of more advanced neuron cells, at day 50, demonstrated no enrichment for the common risk variants for SCZ, implying that a lack of DLG2 fails to lead to any disruption in gene expression later in the process pertinent to pathology. That said, synaptic gene expression of the process in culture may not have fully emerged by day 50 and so it is impossible to rule out its connection to disruption in gene expression.

Comparatively, the genes displaying upregulation in developing neurons at days 30 and 50 and downregulation in radial glia at day 20 of the neurodevelopmental process within WT cultures demonstrated marked enrichment of the common risk variants for SCZ. The functional classes of gene sets enriched for these groups were *cell morphogenesis involved in neuron differentiation, generation of neurons*, *growth cone*, *signal release from synapse, dendrite, regulation of ion transmembrane transporter activity regulation of axon extension,* and *action potential* (Supplementary Table 2). This offered a connection between the recognised interneuron developmental cell types/biology and SCZ regardless of the involvement of DLG2.

6.4. Possible signalling pathways accounting for DLG2 phenotype/function

The analysis of GO term functional enrichment revealed that *DLG2*$^{-/-}$ influences gene expression related to several signalling cascades. An observation of downregulation in the *G-protein coupled receptor signalling pathway* in the RG-group was seen. It is known that G-protein coupled receptors comprise the receptors for many neurotransmitters, neuropeptides, and peptide hormones. Activation of G protein leads to the transportation of the signals to an intracellular target (either an enzyme or an ion channel) after separating from the receptor (Rosenbaum et al., 2009). Therefore, at this stage of development, the depletion of this signalling pathway may mean that a proper RG response to (or sensing of) extracellular cues will not exist. Remarkably, the downregulation in *anterograde trans-synaptic signalling* terms in RG, *cell-cell signalling*

and *trans-synaptic signalling* in the Tran RG group observed may reveal a novel finding where signalling processes exist at this early stage of cortical interneuron development. Possibly, some synaptic signalling molecules (such as DLG2) are also components of precursors' signalling complexes. The existence of these GO terms could explain a delay in cell-cell communication, affecting development, maturation, and migration.

As a scaffolding protein of the PSD, DLG2 has been well-studied and is demonstrated to bridge the extracellular signal transmission to signalling molecules downstream, preserving them in the appropriate position to facilitate signal transduction (cell-signalling) (Zhu et al., 2016). Given its documented activity, it is possible that DLG2 plays a role (or is a part of) in signal transduction complexes within the early stages of cortical interneuron development. Even though there are numerous DEGs and phenotypes present in *DLG2* KO neural progenitor cells, these cells retain the capacity to leave the cell cycle and complete differentiation into interneurons. This was investigated via SOX2 (day 20 and day 50) and NEUN (day 50) expression, indicating that exiting from the cell cycle is autonomous of *DLG2*, even though NKX2.1 and OLIG2 expression were decreased in *DLG2*$^{-/-}$ NPCs, as mentioned earlier. As interneuron generation is connected to leaving the cell cycle, an activity unaffected by a deficit in DLG2, this implies that at least two wide signalling cascades exist, i.e. (i) signalling that is independent of DLG2 that underpins departure from the cell cycle and the development of newborn/developing/mature neurons that follow, and (ii) signalling reliant on DLG2 that directs gene expression to give rise to apposite neuronal identity (Sanders et al., 2022). In view of the heterogeneous characteristics of the *DLG2*$^{-/-}$ phenotype described and the numerous signalling cascades that DLG2 may engage with, the decreased efficiency in signalling may be the likely mechanism. This mechanism was also suggested for *DLG2*-deficient excitatory neurons (Sanders et al., 2022). However, additional research is necessary to confirm this hypothesis.

On day 50 of cortical differentiation when more advanced neurons exist, the expression of most neuronal markers were similar (no significant difference) with a slight decrease trend in GABA at the protein level by the 50th day. This similar development was reinforced by data analysis of the scRNAseq from the 50th day, the most advanced stage of cortical interneuron differentiation examined, which displayed a notably smaller fraction of DEGs in the cells of *DLG2*$^{-/-}$. A total of 75 DEGs (Up + Down) were seen for newborn and developing neurons at days 20 and 30 (C8 and 13, respectively), while 40 DEGs were seen on day 50 for newborn and developing neurons (C18 and 3, respectively). However, on days 52 and 54, DLG2 KO neurons showed a reduction in migration and diminished secondary neurites, respectively, compared to WT. In support

of the above argument (that DLG2 signalling regulates the expression of cell-type identity with observed direct changes in the expression of genes related to morphology/migration reported in Sections 5.8.2 and 5.8.3), the presence of these phenotypes could also be a result of downregulation in mitochondrial genes (*MT-CO3*, *MT-CO1*, *MT-ND3*, *MT-ND1* in developing neurons; *MT-CO1*, *MT-CO2* in newborn neurons; *MT-ND6*, *MT-ND1*, *MT-ND3* in C3; *MT-CO2* in C8). As mitochondria are responsible for generating energy (ATP), mitochondria may play significant roles in regulating basic processes and comprising neurite outgrowth (Cheng et al., 2010), *DLG2*$^{-/-}$ cells appear to have fewer energy needs for development, migration, and maturation.

Merging all the data from scRNAseq, immunocytochemistry, and mass-spectrometry suggests that DLG2 expression arises in the early developmental stages of cortical interneurons and implicates the protein in several signalling cascades linked with development. Since DLG2 is recognised as a scaffolding protein, these findings suggest that it may play this role during the early stages of interneuron differentiation, retaining the components of at least one signalling cascade in the proper orientation and/or position for effective signal transduction. Supposing that this was the case, deficits of DLG2 may not inhibit the signalling pathway's function where it usually forms a scaffold, as although being less well-constructed, the key components remain present, and the signalling pathway components may still exist although not connected to form complexes. However, this could suggest that signalling was more random and more likely incompetent. Such signalling abnormalities could slow/delay (as opposed to prohibiting) the essential stimulation of transcriptional cascades that are essential for the normal development of interneurons (Figure 5.13 and 5.15), together with a range of biomechanisms recognised via the GO term functional enrichment analyses (such *translational initiation*, *RNA splicing*, *DNA templated transcription, initiation* etc.) (Table 5.7 and 5.8).

6.5. Different cell-types/stages identified from scRNAseq

As the immunocytochemistry results from Chapter 3 showed that day-20 cells were mainly progenitors, and day 50 cells were a mixture of progenitors and newborn/developing neurons, single-cell sequencing was highly informative to define the different cell types and developmental stages of cortical interneurons during the differentiation process. The cell types defined from the scRNA-seq data were in accordance with the immunocytochemistry results (Chapter 3) where cells from day 20 showed expression of SOX2 and NKX2.1 (C0 and C1) (Figures 5.9 and 5.11). At day 50, immunocytochemistry gave rise to progenitors and neurons, and scRNAseq showed expression of a mixture of radial glia, dividing radial glia, transitioning radial glia, and

newborn and developing neurons in line with the immunocytochemistry results at day 50. Besides, it is noteworthy that based on the single-cell clustering results, this differentiation protocol does give rise to cortical interneurons. Therefore, scRNAseq data could be used to assess and reveal the developmental transcriptome/processes during cortical interneuron differentiation. This could be achieved by using data of WT samples only from days 20, 30, and 50. This would help to validate the cortical inhibitory differentiation protocol used in this thesis further and uncover crucial genes needed for differentiation.

6.6. Limitations of the study

This current study has used two different knockouts (KO1 and KO2) and one single isogenic wild-type (WT) sister hESC cell lines. The CRISPR-Cas9 technology enables isogenic pairs to be created enabling evaluations and observations of phenotypic differences and downstream effects of a particular mutation. The epigenetic and transcriptional variation of hESC lines may arise from the impact of diverse factors, including the genetic background of the cell line (Bock et al., 2011). Therefore, genetic variability can be reduced significantly, allowing genotype-phenotype correlations to be made by using an isogenic line (Sigmund, 2000).

There was no large CNVs detected in KO cell lines compared to WT. Nevertheless, remains the probability that there are variants/alterations or off-target effects which has not been picked up utilising CNVs analysis. Sanders *et al.* investigated Off-target effects and no differences found (Sanders et al., 2022). Thus, for maximum certainty, whole-genome sequencing should be conducted to rule out any potential off-target effects in CRISPR-Cas9-derived hESC lines. It is possible that while phenotypes are caused by DLG2 loss, expression of these phenotypes might depend on genetic background. The genetic background describes all the genes present in an organism that affects one or more characteristics (Sigmund, 2000). For exploring this possibility, introducing KO mutations into multiple hESC lines could be conducted (e.g. creating *DLG2* KO lines from H9 or H1). Nevertheless, the majority of researchers who use mouse models in their studies, use one strain of mouse. The benefit of this is it ensures the individual mice share a high level of genetic homogeneity with the other mice in the strain. Therefore, the gold standard would be to use isogenic strains that are identical except for the target gene being absent or present. Furthermore, isogenic models ensure the experiments can be reproduced across multiple cases (Casellas, 2011; Sigmund, 2000). Standard use of one inbred strain in KO studies is because of resources and time required to produce one KO line and it applies to the same to the new human pluripotent stem cell line establishment. This is not to say though that the effect of KO will be identical for all the genetic backgrounds. It is best to study in different backgrounds in order to precisely

define the gene's function in each condition, nevertheless, it requires to be started from somewhere. In the context of using the patient iPSC lines, it requires additional lines because there is huge amount of genetic difference between the lines (Germain and Testa, 2017; McTague et al., 2021; Volpato and Webber, 2020). It is unclear exactly how many patient and healthy iPSC lines will be needed to fully determine the impact of risk alleles such as CNVs and SNPs because it would rely on many factors in each experiment (Germain and Testa, 2017; Volpato and Webber, 2020). Nevertheless, the quantity of samples used in both GWAS and exome sequencing studies guarantee that large numbers of iPSC lines will be required to identify the biological effect of risk alleles. This proposes that most studies utilising patient iPSC lines are in fact most likely to be underpowered as only handful of iPSC lines are used in most cases. In general, the lines used in the current thesis are hESCs (not iPSCs) with an isogenic background and as 1) the edited hESC line had no extra CNVs compared to the WT-sister line; 2) no modification has been found in off-target sites; and 3) every KO had different resulting DNA sequence after CRISPR, it seems highly likely that the phenotypes seen are due to DLG2 deficiency, although the further work stated above would help rule out alternatives and consolidate the DLG2 identified phenotypes.

It is important to highlight that 10x scRNAseq suffer from limitations, the capture efficiency is low, as are the sequencing and transcript coverages. Furthermore, the failure of some particular transcripts to be detected is recognised to be a common occurrence (Chen et al., 2019). The level of expression of some genes in this study was below the scRNA seq detection threshold. As mentioned, 10x observed higher noise for mRNAs with low expression levels, single cells can be clustered using Seurat analysis pipeline; however, the Seurat does have limitations. In particular, Seurat is not well suited to small datasets, but its performance improves considerably with large datasets (Adil et al., 2021). When considering the methods used to identify DEGs, t-SNE and PCA methods of dimensionality reduction have certain limitations. Assuming a normal distribution of the data, PCA does not adequately amount to the inherent structural complexities of scRNAseq data, and t-SNE has a larger time complexity (Adil et al., 2021). Yet, Seurat and above-mentioned methods are currently used in the field of scRNAseq analysis (Allaway et al., 2021; Fan et al., 2018; Polioudakis et al., 2019). However, transcriptional heterogeneity is implied by the difference in the expression genes in the same types of cells; to determine the actual level of heterogeneity, DEG analysis requires vigorous computational methods (Adil et al., 2021).

For genes in a single cell that are minimally expressed, it is difficult to capture mRNA molecules into cDNA; thus the likelihood of dropout events is heightened, and is more evident in 10x. Additionally, due to cell death or cell lysis, 10x might produce noise by

collecting some background transcripts present in the droplet. To offset the noise, the capture of single cells could be increased, which would yield more reliable clustering. Although Smart-seq2 benefits from reduced noise and enhanced sensitivity, it is disadvantaged by being expensive. Thus, a well-defined rationale and exacting design should be applied to establish the sample size in Smart-seq2 and 10x (Wang et al., 2021). Compared to Smart-seq2 and other similar tools, 10x captures fewer genes from each cell. To achieve a balance between sequencing costs and research aims, single-cell transcriptomic studies might have to use particular scRNAseq technology (Chen et al., 2019). There are unique advantages and weaknesses for each scRNAseq protocol, meaning that the features and performances of different scRNAseq approaches vary.

The identification of large numbers of DEGs was not completely surprising, as the paper by Sanders *et al.* where bulk RNA sequencing was used on DLG2 KO hESCs differentiated to cortical excitatory neurons, revealed a large number of DEGs (days 20 and 30, around 14000 DEGs) (Sanders et al., 2022). From the functional enrichment analyses, some of the identified phenotypes were similar to the current scRNAseq data such dysregulation in adhesion, DNA binding and transcription, RNA binding, mitotic cell cycle and DNA initiation (Sanders et al., 2022). Bulk RNAseq does not suffer from the limitations of scRNAseq discussed above, however, it lacks cellular specificity, and that may lead to an underestimation of the number of DEGs (e.g. by missing those specific to rare cell-types), not an over-estimation.

Using 3,000 cells for scRNAseq was reasonable as a scRNAseq study used only 466 cells and they were able to distinguished the primary brain cell types and neuron subtypes and DEGs between mature and foetal neurons (Darmanis et al., 2015). Also, a study by Zhong *et al.* identified 35 cell subtypes from 2,300 cells obtained from a human developing cortex (Zhong et al., 2018). However, the sample size used in the study by Polioudakis *et al.* is an example that this study should aim to increase the number of sequenced cells; they conducted a scRNAseq analysis of 40,000 cells derived from a developing human cortex. Interestingly, the researchers detected cell subtypes that had not been identified in humans before (Polioudakis et al., 2019). Thus, it is possible that DLG2 knockout cell lines include some exceptional cell types that is absent from the existing scRNA data due to limited number of cells; including large number of cells would give more detailed explanation.

The low number of cells analysed and limited read depth per cell may be regarded to be general limitations; additionally, the scRNAseq technologies is an emerging field and still developing. Nonetheless these data are compared against those that are published and several labs use 10x scRNAseq (Allaway et al., 2021; Mayer et al., 2018). The data

presented has a good coverage and for other published datasets it is just the technique itself that has certain limitations.

Future plans and priorities to validate the data would be to generate KOs in additional lines. Also, to include large number of cells (e.g. 10000) for scRNAseq with more replicates, as the current data show results from one replicate. Including 3 replicates (from other independent differentiations) per cell line/time point would increase the robustness of the data. Also, with more timepoints, as the current time points are from day 20, 30 and 50. Including more timepoints (day 25, 40, 100) would give a clear idea of the effect of lacking DLG2 on longer time scale during interneuron development. Additionally, co-culturing with excitatory neurons or transplanting human *DLG2*$^{-/-}$ interneurons into mouse brains to help cINs maturation/expression of identity, maintenance and developmental stimulation. Moreover, phenotypic assays could be performed to check the effect of DLG2 KO on calcium level on KOs compared to WT. Furthermore, the 10x Chromium system is one of such throughput droplet-based methods used in this thesis. This method enables 3' mRNA sequencing of up to tens of thousands of single cells per sample. Other platforms could be used (Smart-seq2) with a greater number of cells with benefit of sequencing of full-length transcript coverage. Finally, the use of other hESC cell lines such H1, H9 or iPSC lines or use 3-dimensional (3D) organoid culture system would increase the robustness of the data.

6.7. Further work

Currently only few studies reported the expression of *DLG2*/*Dlg2* during cortical excitatory neurons and neocortex area but not in cortical interneurons. Human *DLG2* can be identified during all the differentiation phases of hESCs to cortical excitatory neurones with DLG2 mRNA level in the last timepoint (high level of postmitotic neurons) is only two to three times that found during the initial phases of excitatory neuron differentiation (van de Leemput et al., 2014). Also, DLG was found to express from eight weeks after conception (Kang et al., 2011). When considering mouse *Dlg2* it has been found to express from day 18 embryonically onwards in cortex (see Section 1.5.5), however, earlier time points was not investigated (Yoo et al., 2020). Despite the transcriptome data from excitatory neurons, there are no published literature or databases that report the expression of DLG2 mRNA/protein during the early and later stages of human interneuron development. A further work could be conducted to perform a detailed analysis of *DLG2* expression in interneuron progenitors (at mRNA and protein level) in both mouse and human because currently this is unknown. This would give a more distinct interpretation of its expression during interneuron development.

A more distinct expression of *DLG2*/*Dlg2* could be gained from performing *in situ* hybridization in mouse and human from different embryonic stages to give a more detailed report about its expression. This would give clearer image of where and when *DLG2* in expressed (spatial and temporal expression). Moreover, it would be useful to test what type of interneuron progenitors/neurons express it at a high level and the level of expression during development and when they start to express it and if the expression will reduce once cell become postnatal neurons or it will continue to express.

Further studies are required to investigate how *DLG2* exerts its effects and explore the role of DLG2 in the cortical interneuron differentiation process. Concentration on spontaneous electrical currents may be examined at day 30 since the neurons at this juncture are immature and because of the part played by the electrical properties in later advancement (Luhmann et al., 2016). As there was downregulation in mitochondria term/genes (diminished energy), the appearance of *transporter activity* term *(metal ion transport*, *cation channel activity, potassium ion transmembrane transporter activity,* and *potassium ion transport* – in the unrefined list) at early stages of development (RG) and in cell-cell signalling terms (Tran RG), and electrical activity plays a role at all developmental stages with ion channel signalling involvement in neuronal induction (Spitzer, 2006), the electrophysiological properties could be deployed to explore aspects of their function in the electrophysiological characteristics of the *DLG2* KO NPCs/maturing interneurons and the impact of disturbing these traits on additionally noted phenotypes.
Since the lack of DLG2 appears to disrupt the control of gene expression, there is a need for future studies to verify which *DLG2*$^{-/-}$ phenotypes return to normal parameters during the process of cortical interneuron differentiation. Several cell-type phenotypes did recuperate solely at the level of mRNA (e.g. 793 RG DEGs on day 20 (C0 and 1) yet only 220 on day 50 (C2 and 16). However, others, e.g. cell morphology alterations and migration defects, need further time-points (e.g. day 70 or day 100) to examine if exhibited signs of a return towards normal within the time frame of cortical interneuron differentiation are achieved.

It would be beneficial to perform mass spectrometry on WT samples from earlier time-points (e.g. days 10 and 15) and later time-points (e.g. days 40 and 60) of cortical interneuron differentiation, to obtain a more precise understanding of the temporal pattern of *in vitro* DLG2 expression, which future research will undoubtedly cover. Also, a crucial area of study would be to employ mass spectrometry to identify proteins that engage with DLG2 in the early and later phases of cortical interneuron differentiation. This would verify the existence of the complexes interacting with the DLG2 scaffold at

these time-points and provide data on the signalling cascades controlled by these entities.

As some DEGs and changes in *DLG2*$^{-/-}$ phenotypes such migration and morphology seen at day 52, 54 respectively, a clear strategy would be to prolong the cortical differentiation process beyond day 50 to enable the delineation of phenotypes at subsequent junctures with the use of immunocytochemistry and scRNAseq. Alongside these, rescue studies could be conducted, e.g. co-cultures of WT and *DLG2*$^{-/-}$ cell lines throughout the interneurons developmental process in order to establish whether this would speed phenotypic recuperation in the latter. If more timely recovery occurred, it would suggest that external signalling, e.g. released molecules or intercellular links, could play a part in the *DLG2*$^{-/-}$ phenotype. Failure to speed up a return towards normal would suggest that the appropriate chemical cues may still be released by KO lines, but the lack of DLG2 impairs the reception/transduction of this signal.

Further immunohistochemistry studies are needed on embryonic mouse brains that investigate alternative markers (e.g. PAX6, FOXP2, DLX2, etc.). Such studies would provide a robust understanding of the effect of *Dlg2* homozygosity and heterozygosity on fate specification during this developmental period. Also, a clear answer to the question of the persistence of phenotypes upon *Dlg2* haplodeficiency/deficiency could be drawn from the investigation of brains of different embryonic ages (e.g. E11, E12, E13, E14.5, E16, E18, etc.) and adult mice. Mature brains could be examined in order to establish the longitudinal impact of *Dlg2* deficiency and the degree to which the early phenotypes return towards normality. Neuronal cell maturation *in vitro* is temporally limited and sustaining cells in cultures for a prolonged interval presents several difficulties. Thus, cell transplantation into murines could be a possibility for further exploration of the later phases of neurodevelopment, as the transplant study will explain how DLG2 shapes cortical interneuron-intrinsic developmental programs. Also, it is well-known that both interneurons and projection neurons are required to establish proper cerebral cortex function and a complex neural network; therefore, a transplantation study would offer such an environment to facilitate a robust investigation of the role of DLG2. Although earlier work has attempted to engineer such models (Nithianantharajah et al., 2013; Yoo et al., 2020), they were studied with respect to adult cognitive and behavioural defects as opposed to embryonic interneuron generation. Moreover, considering the similarities and dissimilarities observed from the immunostained cells (Chapter 4), scRNAseq on *Dlg2*$^{-/-}$ embryonic mouse brain cells (from the MGE, CGE, and the cortex) could be conducted. This would validate the identified phenotypes *in vivo* during early embryogenesis (see Chapter 4), reveal the effect of *Dlg2* KO on mouse transcriptome,

and confirm if *DLG2* plays the same role or a different role in rodents and humans during interneuron development.

Despite providing an enrichment analysis using Nowakowski cell-types in an effort to define the identities of the clusters, consequently, it would be advantageous if there were a bioinformatics tool or platform capable of comparing the regulated genes against all progenitor/neuronal cell datasets, voiding the need to use cell type markers. The capabilities of the cell markers used most often to establish cell identities are limited (Close et al., 2017). scRNAseq analysis data relating to cortical interneurons cell types were recently published by (Close et al., 2017; Fan et al., 2018; Nowakowski et al., 2017; Polioudakis et al., 2019). This data could be used to devise a platform capable of differentiating between various types (using gene expression profiles from well-known cell types) of progenitor and mature neurons.

The work described in this thesis could be reproduced comparably, together with additional studies on other types of human neurons (in which DLG2 is expressed) engineered *in vitro* via directed hESC differentiation, which could be performed using well-documented procedures (Arber et al., 2015; Nicholas et al., 2013; Yuan et al., 2015). Moreover, analysis of scRNAseq at the equivalent/same stages of the cortical differentiation process could be conducted to establish the function of *DLG2* in patient-derived *DLG2*$^{+/-}$ cells in order to examine a model more closely associated with pathology.

6.8. Conclusion

The results of the studies documented above support the notion that DLG2-dependent signalling controls aspects of early interneuron differentiation. One hypothesis is that DLG2 is involved in regulating signal transduction, and DLG2-reliant signalling appears to control biological processes comprising neural precursor cell adhesion and gene expression/replication, which may serve as a foundation for the phenotypes observed later in the process. Considering DLG2-mechanism during interneuron differentiation, one potential theory is that the signalling complexes held in place by the DLG2 scaffold operate at an earlier time-point in neurodevelopment than previously realised and that a lack of *DLG2* leads to chaotic signalling, which, in turn, postpones the onset of transcriptional schedule activation and the ultimate expression of particular cell characteristics necessary for routine development. However, the exact mechanism of DLG2 in the early cortical interneuron differentiation process is still obscure; additional studies are necessary in order to define this fully.

7. References

Adil, A., Kumar, V., Jan, A.T., and Asger, M. (2021). Single-Cell Transcriptomics: Current Methods and Challenges in Data Acquisition and Analysis. Front. Neurosci. *15*, 591122.

Akbarian, S., and Huang, H.S. (2006). Molecular and cellular mechanisms of altered GAD1/GAD67 expression in schizophrenia and related disorders. Brain Res. Rev. *52*, 293-304.

Al-Hallaq, R.A., Yasuda, R.P., and Wolfe, B.B. (2001). Enrichment of N-methyl-D-aspartate NR1 splice variants and synaptic proteins in rat postsynaptic densities. J. Neurochem. *77*, 110-119.

Al-Jaberi, N., Lindsay, S., Sarma, S., Bayatti, N., and Clowry, G.J. (2015). The early fetal development of human neocortical GABAergic interneurons. Cereb. Cortex *25*, 631-645.

Alherz, F., Alherz, M., and Almusawi, H. (2017). NMDAR hypofunction and somatostatin-expressing GABAergic interneurons and receptors: A newly identified correlation and its effects in schizophrenia. Schizophr. Res. Cogn. *8*, 1-6.

Allaway, K.C., Gabitto, M.I., Wapinski, O., Saldi, G., Wang, C.Y., Bandler, R.C., Wu, S.J., Bonneau, R., and Fishell, G. (2021). Genetic and epigenetic coordination of cortical interneuron development. Nature *597*, 693-697.

Alzu'bi, A., Lindsay, S., Kerwin, J., Looi, S.J., Khalil, F., and Clowry, G.J. (2017a). Distinct cortical and sub-cortical neurogenic domains for GABAergic interneuron precursor transcription factors NKX2.1, OLIG2 and COUP-TFII in early fetal human telencephalon. Brain. Struct. Funct. *222*, 2309-2328.

Alzu'bi, A., Lindsay, S.J., Harkin, L.F., McIntyre, J., Lisgo, S.N., and Clowry, G.J. (2017b). The Transcription Factors COUP-TFI and COUP-TFII have Distinct Roles in Arealisation and GABAergic Interneuron Specification in the Early Human Fetal Telencephalon. Cereb. Cortex *27*, 4971-4987.

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*, 474-476.

Anderson, S.A., Marin, O., Horn, C., Jennings, K., and Rubenstein, J.L. (2001). Distinct cortical migrations from the medial and lateral ganglionic eminences. Development *128*, 353-363.

Anselme, I., Laclef, C., Lanaud, M., Ruther, U., and Schneider-Maunoury, S. (2007). Defects in brain patterning and head morphogenesis in the mouse mutant Fused toes. Dev. Biol. *304*, 208-220.

Appolloni, I., Calzolari, F., Corte, G., Perris, R., and Malatesta, P. (2008). Six3 controls the neural progenitor status in the murine CNS. Cereb. Cortex *18*, 553-562.

Arber, C., and Li, M. (2013). Cortical interneurons from human pluripotent stem cells: prospects for neurological and psychiatric disease. Front. Cell. Neurosci. *7*, 10.

Arber, C., Precious, S.V., Cambray, S., Risner-Janiczek, J.R., Kelly, C., Noakes, Z., Fjodorova, M., Heuer, A., Ungless, M.A., Rodriguez, T.A., et al. (2015). Activin A directs striatal projection neuron differentiation of human pluripotent stem cells. Development *142*, 1375-1386.

Armbruster, D.J., Ueltzhoffer, K., Basten, U., and Fiebach, C.J. (2012). Prefrontal cortical mechanisms underlying individual differences in cognitive flexibility and stability. J. Cogn. Neurosci. *24*, 2385-2399.

Arshad, A., Vose, L.R., Vinukonda, G., Hu, F., Yoshikawa, K., Csiszar, A., Brumberg, J.C., and Ballabh, P. (2016). Extended Production of Cortical Interneurons into the Third Trimester of Human Gestation. Cereb. Cortex *26*, 2242-2256.

Asada, H., Kawamura, Y., Maruyama, K., Kume, H., Ding, R., Ji, F.Y., Kanbara, N., Kuzume, H., Sanbo, M., Yagi, T., et al. (1996). Mice lacking the 65 kDa isoform of glutamic acid decarboxylase (GAD65) maintain normal levels of GAD67 and GABA in their brains but are susceptible to seizures. Biochem. Biophys. Res. Commun. *229*, 891-895.

Asada, H., Kawamura, Y., Maruyama, K., Kume, H., Ding, R.G., Kanbara, N., Kuzume, H., Sanbo, M., Yagi, T., and Obata, K. (1997). Cleft palate and decreased brain gamma-aminobutyric acid in mice lacking the 67-kDa isoform of glutamic acid decarboxylase. Proc. Natl. Acad. Sci. U. S. A. *94*, 6496-6499.

Astarita, J.L., Acton, S.E., and Turley, S.J. (2012). Podoplanin: emerging functions in development, the immune system, and cancer. Front. Immunol. *3*, 283.

Aviles, E.C., Wilson, N.H., and Stoeckli, E.T. (2013). Sonic hedgehog and Wnt: antagonists in morphogenesis but collaborators in axon guidance. Front. Cell. Neurosci. *7*, 86.

Azim, E., Jabaudon, D., Fame, R.M., and Macklis, J.D. (2009). SOX6 controls dorsal progenitor identity and interneuron diversity during neocortical development. Nat. Neurosci. *12*, 1238-1247.

Bansod, S., Kageyama, R., and Ohtsuka, T. (2017). Hes5 regulates the transition timing of neurogenesis and gliogenesis in mammalian neocortical development. Development *144*, 3156-3167.

Bassett, A.S., Scherer, S.W., and Brzustowicz, L.M. (2010). Copy number variations in schizophrenia: critical review and new perspectives on concepts of genetics and disease. Am. J. Psychiatry *167*, 899-914.

Batista-Brito, R., Close, J., Machold, R., and Fishell, G. (2008). The distinct temporal origins of olfactory bulb interneuron subtypes. J. Neurosci. *28*, 3966-3975.

Batista-Brito, R., Rossignol, E., Hjerling-Leffler, J., Denaxa, M., Wegner, M., Lefebvre, V., Pachnis, V., and Fishell, G. (2009). The cell-intrinsic requirement of Sox6 for cortical interneuron development. Neuron *63*, 466-481.

Batista-Brito, R., and Fishell, G. (2009). The developmental integration of cortical interneurons into a functional network. Curr. Top. Dev. Biol. *87*, 81-118.

Bean, J.C., Lin, T.W., Sathyamurthy, A., Liu, F., Yin, D.M., Xiong, W.C., and Mei, L. (2014). Genetic labeling reveals novel cellular targets of schizophrenia susceptibility gene: distribution of GABA and non-GABA ErbB4-positive cells in adult mouse brain. J. Neurosci. *34*, 13549-13566.

Belgacem, Y.H., Hamilton, A.M., Shim, S., Spencer, K.A., and Borodinsky, L.N. (2016). The Many Hats of Sonic Hedgehog Signaling in Nervous System Development and Disease. J Dev Biol *4*.

Benarroch, E. (2021). What Is the Role of Stathmin-2 in Axonal Biology and Degeneration? Neurology *97*, 330-333.

Benes, F.M., McSparren, J., Bird, E.D., SanGiovanni, J.P., and Vincent, S.L. (1991). Deficits in small interneurons in prefrontal and cingulate cortices of schizophrenic and schizoaffective patients. Arch. Gen. Psychiatry *48*, 996-1001.

Berger, T.K., Perin, R., Silberberg, G., and Markram, H. (2009). Frequency-dependent disynaptic inhibition in the pyramidal network: a ubiquitous pathway in the developing rat neocortex. J. Physiol. *587*, 5411-5425.

Bernal, A., and Arranz, L. (2018). Nestin-expressing progenitor cells: function, identity and therapeutic implications. Cell. Mol. Life. Sci. *75*, 2177-2195.

Bertini, V., Milone, R., Cristofani, P., Cambi, F., Bosetti, C., Barbieri, F., Bertelloni, S., Cioni, G., Valetto, A., and Battini, R. (2022). Enhancing DLG2 Implications in Neuropsychiatric Disorders: Analysis of a Cohort of Eight Patients with 11q14.1 Imbalances. Genes (Basel) *13*.

Bock, C., Kiskinis, E., Verstappen, G., Gu, H., Boulting, G., Smith, Z.D., Ziller, M., Croft, G.F., Amoroso, M.W., Oakley, D.H., et al. (2011). Reference Maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. Cell *144*, 439-452.

Boettcher, M., and McManus, M.T. (2015). Choosing the Right Tool for the Job: RNAi, TALEN, or CRISPR. Mol. Cell *58*, 575-585.

Boldog, E., Bakken, T.E., Hodge, R.D., Novotny, M., Aevermann, B.D., Baka, J., Borde, S., Close, J.L., Diez-Fuertes, F., Ding, S.L., et al. (2018). Transcriptomic and morphophysiological evidence for a specialized human cortical GABAergic cell type. Nat. Neurosci. *21*, 1185-1195.

Boraschi-Diaz, I., Wang, J., Mort, J.S., and Komarova, S.V. (2017). Collagen type I as a ligand for receptor-mediated signaling. Frontiers in Physics *5*, 12.

Bortone, D.S., Olsen, S.R., and Scanziani, M. (2014). Translaminar inhibitory cells recruited by layer 6 corticothalamic neurons suppress visual cortex. Neuron *82*, 474-485.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., et al. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell *122*, 947-956.

Brandao, J.A., and Romcy-Pereira, R.N. (2015). Interplay of environmental signals and progenitor diversity on fate specification of cortical GABAergic neurons. Front. Cell. Neurosci. *9*, 149.

Brekken, R.A., and Sage, E.H. (2000). SPARC, a matricellular protein: at the crossroads of cell–matrix. Matrix Biology *19*, 569-580.

Brown, K.N., Chen, S., Han, Z., Lu, C.H., Tan, X., Zhang, X.J., Ding, L., Lopez-Cruz, A., Saur, D., Anderson, S.A., et al. (2011). Clonal production and organization of inhibitory interneurons in the neocortex. Science *334*, 480-486.

Brudvig, J.J., and Weimer, J.M. (2015). X MARCKS the spot: myristoylated alanine-rich C kinase substrate in neuronal function and disease. Front. Cell. Neurosci. *9*, 407.

Buchanan, K.A., Blackman, A.V., Moreau, A.W., Elgar, D., Costa, R.P., Lalanne, T., Tudor Jones, A.A., Oyrer, J., and Sjostrom, P.J. (2012). Target-specific expression of presynaptic NMDA receptors in neocortical microcircuits. Neuron *75*, 451-466.

Budnik, V., Koh, Y.H., Guan, B., Hartmann, B., Hough, C., Woods, D., and Gorczyca, M. (1996). Regulation of synapse structure and function by the Drosophila tumor suppressor gene dlg. Neuron *17*, 627-640.

Buhl, D.L., Harris, K.D., Hormuzdi, S.G., Monyer, H., and Buzsáki, G. (2003). Selective Impairment of Hippocampal Gamma Oscillations in Connexin-36 Knock-Out MouseIn Vivo. J. Neurosci. *23*, 1013-1018.

Buonanno, A. (2010). The neuregulin signaling pathway and schizophrenia: from genes to synapses and neural circuits. Brain Res Bull *83*, 122-131.

Butt, S.J., Fuccillo, M., Nery, S., Noctor, S., Kriegstein, A., Corbin, J.G., and Fishell, G. (2005). The temporal and spatial origins of cortical interneurons predict their physiological subtype. Neuron *48*, 591-604.

Butt, S.J., Sousa, V.H., Fuccillo, M.V., Hjerling-Leffler, J., Miyoshi, G., Kimura, S., and Fishell, G. (2008). The requirement of Nkx2-1 in the temporal specification of cortical interneuron subtypes. Neuron *59*, 722-732.

Cai, Y., Zhang, Q., Wang, C., Zhang, Y., Ma, T., Zhou, X., Tian, M., Rubenstein, J.L., and Yang, Z. (2013). Nuclear receptor COUP-TFII-expressing neocortical interneurons are derived from the medial and lateral/caudal ganglionic eminence and define specific subsets of mature interneurons. J. Comp. Neurol. *521*, 479-497.

Cantor-Graae, E., and Selten, J.P. (2005). Schizophrenia and migration: a meta-analysis and review. Am. J. Psychiatry *162*, 12-24.

Cardin, J.A., Carlen, M., Meletis, K., Knoblich, U., Zhang, F., Deisseroth, K., Tsai, L.H., and Moore, C.I. (2009). Driving fast-spiking cells induces gamma rhythm and controls sensory responses. Nature *459*, 663-667.

Cardno, A.G., and Gottesman, I.I. (2000). Twin studies of schizophrenia: from bow-and-arrow concordances to star wars Mx and functional genomics. Am. J. Med. Genet. *97*, 12-17.

Carelli, V., and Chan, D.C. (2014). Mitochondrial DNA: impacting central and peripheral nervous systems. Neuron *84*, 1126-1142.

Carlisle, H.J., Fink, A.E., Grant, S.G., and O'Dell, T.J. (2008). Opposing effects of PSD-93 and PSD-95 on long-term potentiation and spike timing-dependent plasticity. J. Physiol. *586*, 5885-5900.

Casarosa, S., Fode, C., and Guillemot, F. (1999). Mash1 regulates neurogenesis in the ventral telencephalon. Development *126*, 525-534.

Casellas, J. (2011). Inbred mouse strains and genetic stability: a review. Animal *5*, 1-7.

Catalina, P., Montes, R., Ligero, G., Sanchez, L., de la Cueva, T., Bueno, C., Leone, P.E., and Menendez, P. (2008). Human ESCs predisposition to karyotypic instability: Is a matter of culture adaptation or differential vulnerability among hESC lines due to inherent properties? Mol. Cancer *7*, 76.

Cauli, B., Zhou, X., Tricoire, L., Toussay, X., and Staiger, J.F. (2014). Revisiting enigmatic cortical calretinin-expressing interneurons. Front. Neuroanat. *8*, 52.

Chambers, S.M., Fasano, C.A., Papapetrou, E.P., Tomishima, M., Sadelain, M., and Studer, L. (2009). Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. Nat. Biotechnol. *27*, 275-280.

Chambers, S.M., Qi, Y., Mica, Y., Lee, G., Zhang, X.J., Niu, L., Bilsland, J., Cao, L., Stevens, E., Whiting, P., et al. (2012). Combined small-molecule inhibition accelerates developmental timing and converts human pluripotent stem cells into nociceptors. Nat. Biotechnol. *30*, 715-720.

Chandrasekaran, A., Avci, H.X., Ochalek, A., Rosingh, L.N., Molnar, K., Laszlo, L., Bellak, T., Teglasi, A., Pesti, K., Mike, A., et al. (2017). Comparison of 2D and 3D neural induction methods for the generation of neural progenitor cells from human induced pluripotent stem cells. Stem Cell Res *25*, 139-151.

Chen, G., Ning, B., and Shi, T. (2019). Single-Cell RNA-Seq Technologies and Related Computational Data Analysis. Front. Genet. *10*, 317.

Chen, L., Chetkovich, D.M., Petralia, R.S., Sweeney, N.T., Kawasaki, Y., Wenthold, R.J., Bredt, D.S., and Nicoll, R.A. (2000). Stargazin regulates synaptic targeting of AMPA receptors by two distinct mechanisms. Nature *408*, 936-943.

Cheng, A., Hou, Y., and Mattson, M.P. (2010). Mitochondria and neuroplasticity. ASN Neuro *2*, e00045.

Chu, J., and Anderson, S.A. (2015). Development of cortical interneurons. Neuropsychopharmacology *40*, 16-23.

Chubykin, A.A., Atasoy, D., Etherton, M.R., Brose, N., Kavalali, E.T., Gibson, J.R., and Sudhof, T.C. (2007). Activity-dependent validation of excitatory versus inhibitory synapses by neuroligin-1 versus neuroligin-2. Neuron *54*, 919-931.

Clancy, M.J., Clarke, M.C., Connor, D.J., Cannon, M., and Cotter, D.R. (2014). The prevalence of psychosis in epilepsy; a systematic review and meta-analysis. BMC Psychiatry *14*, 75.

Clifton, N.E., Hannon, E., Harwood, J.C., Di Florio, A., Thomas, K.L., Holmans, P.A., Walters, J.T.R., O'Donovan, M.C., Owen, M.J., Pocklington, A.J., et al. (2019). Dynamic expression of genes associated with schizophrenia and bipolar disorder across development. Transl. Psychiatry *9*, 74.

Close, J.L., Yao, Z., Levi, B.P., Miller, J.A., Bakken, T.E., Menon, V., Ting, J.T., Wall, A., Krostag, A.R., Thomsen, E.R., et al. (2017). Single-Cell Profiling of an In Vitro Model of Human Interneuron Development Reveals Temporal Dynamics of Cell Type Production and Maturation. Neuron *93*, 1035-1048 e1035.

Cobos, I., Long, J.E., Thwin, M.T., and Rubenstein, J.L. (2006). Cellular patterns of transcription factor expression in developing cortical interneurons. Cereb. Cortex *16 Suppl 1*, i82-88.

Cohen, M., Briscoe, J., and Blassberg, R. (2013). Morphogen interpretation: the transcriptional logic of neural tube patterning. Curr. Opin. Genet. Dev. *23*, 423-428.

Colgan, L.A., and Yasuda, R. (2014). Plasticity of dendritic spines: subcompartmentalization of signaling. Annu. Rev. Physiol. *76*, 365-385.

Condie, B.G., Bain, G., Gottlieb, D.I., and Capecchi, M.R. (1997). Cleft palate in mice with a targeted mutation in the gamma-aminobutyric acid-producing enzyme glutamic acid decarboxylase 67. Proc. Natl. Acad. Sci. U. S. A. *94*, 11451-11455.

Corbin, J.G., and Butt, S.J. (2011). Developmental mechanisms for the generation of telencephalic interneurons. Dev. Neurobiol. *71*, 710-732.

Curley, A.A., Arion, D., Volk, D.W., Asafu-Adjei, J.K., Sampson, A.R., Fish, K.N., and Lewis, D.A. (2011). Cortical deficits of glutamic acid decarboxylase 67 expression in schizophrenia: clinical, protein, and cell type-specific features. Am. J. Psychiatry *168*, 921-929.

D'Avino, P.P. (2017). Citron kinase - renaissance of a neglected mitotic kinase. J. Cell Sci. *130*, 1701-1708.

Dai, R., Chen, L., Liu, S., Wu, C.-T., Chen, Y., Jiang, Y., Dai, J., Wang, Q., Kopp, R., Yu, G., et al. (2021). Cell group analysis reveals changes in upper-layer neurons associated with schizophrenia. BioRxiv.

Danen, E.H., Sonneveld, P., Brakebusch, C., Fassler, R., and Sonnenberg, A. (2002). The fibronectin-binding integrins alpha5beta1 and alphavbeta3 differentially modulate RhoA-GTP loading, organization of cell matrix adhesions, and fibronectin fibrillogenesis. J. Cell Biol. *159*, 1071-1086.

Danjo, T., Eiraku, M., Muguruma, K., Watanabe, K., Kawada, M., Yanagawa, Y., Rubenstein, J.L., and Sasai, Y. (2011). Subregional specification of embryonic stem cell-derived ventral telencephalic tissues by timed and combinatory treatment with extrinsic signals. J. Neurosci. *31*, 1919-1933.

Darmanis, S., Sloan, S.A., Zhang, Y., Enge, M., Caneda, C., Shuer, L.M., Hayden Gephart, M.G., Barres, B.A., and Quake, S.R. (2015). A survey of human brain transcriptome diversity at the single cell level. Proc. Natl. Acad. Sci. *112*, 7285-7290.

Davidson, M., Reichenberg, A., Rabinowitz, J., Weiser, M., Kaplan, Z., and Mark, M. (1999). Behavioral and intellectual markers for schizophrenia in apparently healthy male adolescents. Am. J. Psychiatry *156*, 1328-1335.

Davila, J.C., Real, M.A., Olmos, L., Legaz, I., Medina, L., and Guirado, S. (2005). Embryonic and postnatal development of GABA, calbindin, calretinin, and parvalbumin in the mouse claustral complex. J. Comp. Neurol. *481*, 42-57.

de Leeuw, C.A., Mooij, J.M., Heskes, T., and Posthuma, D. (2015). MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput. Biol. *11*, e1004219.

Deacon, T., Pakzaban, P., and Isacson, O. (1994). The lateral ganglionic eminence is the origin of cells committed to striatal phenotypes: neural transplantation and developmental evidence. Brain Res. *668*, 211-219.

DeFelipe, J., Hendry, S.H., and Jones, E.G. (1989). Visualization of chandelier cell axons by parvalbumin immunoreactivity in monkey cerebral cortex. Proc. Natl. Acad. Sci. U. S. A. *86*, 2093-2097.

Del Pino, I., Garcia-Frigola, C., Dehorter, N., Brotons-Mas, J.R., Alvarez-Salvado, E., Martinez de Lagran, M., Ciceri, G., Gabaldon, M.V., Moratal, D., Dierssen, M., et al. (2013). Erbb4 deletion from fast-spiking interneurons causes schizophrenia-like phenotypes. Neuron *79*, 1152-1168.

Del Rio, J.A., Martinez, A., Auladell, C., and Soriano, E. (2000). Developmental history of the subplate and developing white matter in the murine neocortex. Neuronal organization and relationship with the main afferent systems at embryonic and perinatal stages. Cereb. Cortex *10*, 784-801.

Dessaud, E., McMahon, A.P., and Briscoe, J. (2008). Pattern formation in the vertebrate neural tube: a sonic hedgehog morphogen-regulated transcriptional network. Development *135*, 2489-2503.

Ding, Z.M., Huang, C.J., Jiao, X.F., Wu, D., and Huo, L.J. (2017). The role of TACC3 in mitotic spindle organization. Cytoskeleton (Hoboken) *74*, 369-378.

Dingledine, R., Borges, K., Bowie, D., and Traynelis, S.F. (1999). The glutamate receptor ion channels. Pharmacol. Rev. *51*, 7-61.

Dittrich, L., Heiss, J.E., Warrier, D.R., Perez, X.A., Quik, M., and Kilduff, T.S. (2012). Cortical nNOS neurons co-express the NK1 receptor and are depolarized by Substance P in multiple mammalian species. Front. Neural Circuits *6*, 31.

Draper, J.S., Smith, K., Gokhale, P., Moore, H.D., Maltby, E., Johnson, J., Meisner, L., Zwaka, T.P., Thomson, J.A., and Andrews, P.W. (2004). Recurrent gain of chromosomes 17q and 12 in cultured human embryonic stem cells. Nat. Biotechnol. *22*, 53-54.

Du, T., Xu, Q., Ocbina, P.J., and Anderson, S.A. (2008). NKX2.1 specifies cortical interneuron fate by activating Lhx6. Development *135*, 1559-1567.

Duan, W., Zhang, Y.P., Hou, Z., Huang, C., Zhu, H., Zhang, C.Q., and Yin, Q. (2016). Novel Insights into NeuN: from Neuronal Marker to Splicing Regulator. Mol. Neurobiol. *53*, 1637-1647.

Eastwood, S.L., Cairns, N.J., and Harrison, P.J. (2000). Synaptophysin gene expression in schizophrenia. Investigation of synaptic pathology in the cerebral cortex. Br. J. Psychiatry *176*, 236-242.

Eisenstat, D.D., Liu, J.K., Mione, M., Zhong, W., Yu, G., Anderson, S.A., Ghattas, I., Puelles, L., and Rubenstein, J.L. (1999). DLX-1, DLX-2, and DLX-5 expression define distinct stages of basal forebrain differentiation. J. Comp. Neurol. *414*, 217-237.

Elsum, I., Yates, L., Humbert, P.O., and Richardson, H.E. (2012). The Scribble-Dlg-Lgl polarity module in development and cancer: from flies to man. Essays Biochem *53*, 141-168.

Elvekrog, M.M., and Walter, P. (2015). Dynamics of co-translational protein targeting. Curr. Opin. Chem. Biol. *29*, 79-86.

Englund, C., Fink, A., Lau, C., Pham, D., Daza, R.A., Bulfone, A., Kowalczyk, T., and Hevner, R.F. (2005). Pax6, Tbr2, and Tbr1 are expressed sequentially by radial glia, intermediate progenitor cells, and postmitotic neurons in developing neocortex. J. Neurosci. *25*, 247-251.

Erisir, A., Lau, D., Rudy, B., and Leonard, C.S. (1999). Function of specific K(+) channels in sustained high-frequency firing of fast-spiking neocortical interneurons. J. Neurophysiol. *82*, 2476-2489.

Escudero, I., and Johnstone, M. (2014). Genetics of schizophrenia. Curr. Psychiatry. Rep. *16*, 502.

Fairen, A., Cobas, A., and Fonseca, M. (1986). Times of generation of glutamic acid decarboxylase immunoreactive neurons in mouse somatosensory cortex. J. Comp. Neurol. *251*, 67-83.

Fairen, A., and Valverde, F. (1980). A specialized type of neuron in the visual cortex of cat: a Golgi and electron microscope study of chandelier cells. J. Comp. Neurol. *194*, 761-779.

Fan, X., Dong, J., Zhong, S., Wei, Y., Wu, Q., Yan, L., Yong, J., Sun, L., Wang, X., Zhao, Y., et al. (2018). Spatial transcriptomic survey of human embryonic cerebral cortex by single-cell RNA-seq analysis. Cell Res *28*, 730-745.

Fanselow, E.E., and Connors, B.W. (2010). The roles of somatostatin-expressing (GIN) and fast-spiking inhibitory interneurons in UP-DOWN states of mouse neocortex. J. Neurophysiol. *104*, 596-606.

Fanselow, E.E., Richardson, K.A., and Connors, B.W. (2008). Selective, state-dependent activation of somatostatin-expressing inhibitory interneurons in mouse neocortex. J. Neurophysiol. *100*, 2640-2652.

Faux, C., Rakic, S., Andrews, W., and Britto, J.M. (2012). Neurons on the move: migration and lamination of cortical interneurons. Neurosignals *20*, 168-189.

Fazzari, P., Paternain, A.V., Valiente, M., Pla, R., Lujan, R., Lloyd, K., Lerma, J., Marin, O., and Rico, B. (2010). Control of cortical GABA circuitry development by Nrg1 and ErbB4 signalling. Nature *464*, 1376-1380.

Felding-Habermann, B., and Cheresh, D.A. (1993). Vitronectin and its receptors. Curr Opin Cell Biol *5*, 864-868.

Feng, W., and Zhang, M. (2009). Organization and dynamics of PDZ-domain-related supramodules in the postsynaptic density. Nat. Rev. Neurosci. *10*, 87-99.

Férézou, I., Cauli, B., Hill, E.L., Rossier, J., Hamel, E., and Lambolez, B. (2002). 5-HT3Receptors Mediate Serotonergic Fast Synaptic Excitation of Neocortical Vasoactive Intestinal Peptide/Cholecystokinin Interneurons. J. Neurosci. *22*, 7389-7397.

Ferguson, B.R., and Gao, W.J. (2018). PV Interneurons: Critical Regulators of E/I Balance for Prefrontal Cortex-Dependent Behavior and Psychiatric Disorders. Front. Neural Circuits *12*, 37.

Ferri, A., Favaro, R., Beccari, L., Bertolini, J., Mercurio, S., Nieto-Lopez, F., Verzeroli C, La Regina F, De Pietri Tonelli D, Ottolenghi S, et al. (2013). Sox2 is required for embryonic development of the ventral telencephalon through the activation of the ventral determinants Nkx2.1 and Shh. Development *140*, 1250–1261.

Fertuzinhos, S., Krsnik, Z., Kawasawa, Y.I., Rasin, M.R., Kwan, K.Y., Chen, J.G., Judas, M., Hayashi, M., and Sestan, N. (2009). Selective depletion of molecularly defined cortical interneurons in human holoprosencephaly with severe striatal hypoplasia. Cereb. Cortex *19*, 2196-2207.

Feyder, M., Karlsson, R.M., Mathur, P., Lyman, M., Bock, R., Momenan, R., Munasinghe, J., Scattoni, M.L., Ihne, J., Camp, M., et al. (2010). Association of mouse Dlg4 (PSD-95) gene deletion and human DLG4 gene variation with phenotypes relevant to autism spectrum disorders and Williams' syndrome. Am. J. Psychiatry *167*, 1508-1517.

Fishman, M.C. (1996). GAP-43: putting constraints on neuronal plasticity. Perspect. Dev. Neurobiol. *4*, 193-198.

Fitzgerald, M., Sotuyo, N., Tischfield, D.J., and Anderson, S.A. (2020). Generation of cerebral cortical GABAergic interneurons from pluripotent stem cells. Stem Cells *38*, 1375-1386.

Flames, N., Long, J.E., Garratt, A.N., Fischer, T.M., Gassmann, M., Birchmeier, C., Lai, C., Rubenstein, J.L., and Marin, O. (2004). Short- and long-range attraction of cortical GABAergic interneurons by neuregulin-1. Neuron *44*, 251-261.

Flames, N., Pla, R., Gelman, D.M., Rubenstein, J.L., Puelles, L., and Marin, O. (2007). Delineation of multiple subpallial progenitor domains by the combinatorial expression of transcriptional codes. J. Neurosci. *27*, 9682-9695.

Flandin, P., Kimura, S., and Rubenstein, J.L. (2010). The progenitor zone of the ventral medial ganglionic eminence requires Nkx2-1 to generate most of the globus pallidus but few neocortical interneurons. J. Neurosci. *30*, 2812-2823.

Flandin, P., Zhao, Y., Vogt, D., Jeong, J., Long, J., Potter, G., Westphal, H., and Rubenstein, J.L. (2011). Lhx6 and Lhx8 coordinately induce neuronal expression of Shh that controls the generation of interneuron progenitors. Neuron *70*, 939-950.

Fo, C.S., Coleman, C.S., Wallick, C.J., Vine, A.L., and Bachmann, A.S. (2006). Genomic organization, expression profile, and characterization of the new protein PRA1 domain family, member 2 (PRAF2). Gene *371*, 154-165.

Fode, C., Ma, Q., Casarosa, S., Ang, S.L., Anderson, D.J., and Guillemot, F. (2000). A role for neural determination genes in specifying the dorsoventral identity of telencephalic neurons. Genes Dev. *14*, 67-80.

Fontanet, P.A., Rios, A.S., Alsina, F.C., Paratcha, G., and Ledda, F. (2018). Pea3 Transcription Factors, Etv4 and Etv5, Are Required for Proper Hippocampal Dendrite Development and Plasticity. Cereb. Cortex *28*, 236-249.

Forsingdal, A., Jorgensen, T.N., Olsen, L., Werge, T., Didriksen, M., and Nielsen, J. (2019). Can Animal Models of Copy Number Variants That Predispose to Schizophrenia Elucidate Underlying Biology? Biol. Psychiatry *85*, 13-24.

Franco, M., and Carmena, A. (2019). Eph signaling controls mitotic spindle orientation and cell proliferation in neuroepithelial cells. J. Cell Biol. *218*, 1200-1217.

Frank, R.A., Komiyama, N.H., Ryan, T.J., Zhu, F., O'Dell, T.J., and Grant, S.G. (2016). NMDA receptors are selectively partitioned into complexes and supercomplexes during synapse maturation. Nat. Commun. *7*, 11264.

Frantseva, M.V., Fitzgerald, P.B., Chen, R., Moller, B., Daigle, M., and Daskalakis, Z.J. (2008). Evidence for impaired long-term potentiation in schizophrenia and its relationship to motor skill learning. Cereb. Cortex *18*, 990-996.

Frantz, C., Stewart, K.M., and Weaver, V.M. (2010). The extracellular matrix at a glance. J. Cell Sci. *123*, 4195-4200.

Frazer, S., Prados, J., Niquille, M., Cadilhac, C., Markopoulos, F., Gomez, L., Tomasello, U., Telley, L., Holtmaat, A., Jabaudon, D., et al. (2017). Transcriptomic and anatomic parcellation of 5-HT3AR expressing cortical interneuron subtypes revealed by single-cell RNA sequencing. Nat. Commun. *8*, 14219.

Friocourt, G., Liu, J.S., Antypa, M., Rakic, S., Walsh, C.A., and Parnavelas, J.G. (2007). Both doublecortin and doublecortin-like kinase play a role in cortical interneuron migration. J. Neurosci. *27*, 3875-3883.

Fritschy, J.M., Harvey, R.J., and Schwarz, G. (2008). Gephyrin: where do we stand, where do we go? Trends Neurosci. *31*, 257-264.

Fromer, M., Pocklington, A.J., Kavanagh, D.H., Williams, H.J., Dwyer, S., Gormley, P., Georgieva, L., Rees, E., Palta, P., Ruderfer, D.M., et al. (2014). De novo mutations in schizophrenia implicate synaptic networks. Nature *506*, 179-184.

Fung, S.J., Fillman, S.G., Webster, M.J., and Shannon Weickert, C. (2014). Schizophrenia and bipolar disorder show both common and distinct changes in cortical interneuron markers. Schizophr. Res. *155*, 26-30.

Funke, L., Dakoji, S., and Bredt, D.S. (2005). Membrane-associated guanylate kinases regulate adhesion and plasticity at cell junctions. Annu. Rev. Biochem. *74*, 219-245.

Furusho, M., Ono, K., Takebayashi, H., Masahira, N., Kagawa, T., Ikeda, K., and Ikenaka, K. (2006). Involvement of the Olig2 transcription factor in cholinergic neuron development of the basal forebrain. Dev. Biol. *293*, 348-357.

Garcia, R.A.G., Vasudevan, K., and Buonanno, A. (2000). The neuregulin receptor ErbB-4 interacts with PDZ-containing proteins at neuronal synapses. Proc. Natl. Acad. Sci. *97*, 3596-3601.

Gardoni, F., Marcello, E., and Di Luca, M. (2009). Postsynaptic density-membrane associated guanylate kinase proteins (PSD-MAGUKs) and their role in CNS disorders. Neuroscience *158*, 324-333.

Gasnier, B. (2004). The SLC32 transporter, a key protein for the synaptic release of inhibitory amino acids. Pflugers Arch. *447*, 756-759.

Geiger, B., Bershadsky, A., Pankov, R., and Yamada, K.M. (2001). Transmembrane crosstalk between the extracellular matrix--cytoskeleton crosstalk. Nat. Rev. Mol. Cell Biol. *2*, 793-805.

Gejman, P.V., Sanders, A.R., and Duan, J. (2010). The role of genetics in the etiology of schizophrenia. Psychiatr. Clin. North Am. *33*, 35-66.

Gelman, D., Griveau, A., Dehorter, N., Teissier, A., Varela, C., Pla, R., Pierani, A., and Marin, O. (2011). A wide diversity of cortical GABAergic interneurons derives from the embryonic preoptic area. J. Neurosci. *31*, 16570-16580.

Gelman, D.M., Martini, F.J., Nobrega-Pereira, S., Pierani, A., Kessaris, N., and Marin, O. (2009). The embryonic preoptic area is a novel source of cortical GABAergic interneurons. J. Neurosci. *29*, 9380-9389.

Genovese, G., Fromer, M., Stahl, E.A., Ruderfer, D.M., Chambert, K., Landen, M., Moran, J.L., Purcell, S.M., Sklar, P., Sullivan, P.F., et al. (2016). Increased burden of ultra-rare protein-altering variants among 4,877 individuals with schizophrenia. Nat. Neurosci. *19*, 1433-1441.

Germain, P.L., and Testa, G. (2017). Taming Human Genetic Variability: Transcriptomic Meta-Analysis Guides the Experimental Design and Interpretation of iPSC-Based Disease Modeling. Stem Cell Reports *8*, 1784-1796.

Gestri, G., Carl, M., Appolloni, I., Wilson, S.W., Barsacchi, G., and Andreazzoli, M. (2005). Six3 functions in anterior neural plate specification by promoting cell proliferation and inhibiting Bmp4 expression. Development *132*, 2401-2413.

Glickfeld, L.L., Roberts, J.D., Somogyi, P., and Scanziani, M. (2009). Interneurons hyperpolarize pyramidal cells along their entire somatodendritic axis. Nat. Neurosci. *12*, 21-23.

Godin, J.D., and Nguyen, L. (2014). Novel functions of core cell cycle regulators in neuronal migration. Adv. Exp. Med. Biol. *800*, 59-74.

Goldberg, E.M., Clark, B.D., Zagha, E., Nahmani, M., Erisir, A., and Rudy, B. (2008). K+ channels at the axon initial segment dampen near-threshold excitability of neocortical fast-spiking GABAergic interneurons. Neuron *58*, 387-400.

Gonzalez-Burgos, G., Hashimoto, T., and Lewis, D.A. (2010). Alterations of cortical GABA neurons and network oscillations in schizophrenia. Curr. Psychiatry. Rep. *12*, 335-344.

Gonzalez-Burgos, G., Krimer, L.S., Povysheva, N.V., Barrionuevo, G., and Lewis, D.A. (2005). Functional properties of fast spiking interneurons and their synaptic connections with pyramidal cells in primate dorsolateral prefrontal cortex. J. Neurophysiol. *93*, 942-953.

Goto, N., Yoshimura, R., Moriya, J., Kakeda, S., Ueda, N., Ikenouchi-Sugita, A., Umene-Nakano, W., Hayashi, K., Oonari, N., Korogi, Y., et al. (2009). Reduction of brain gamma-aminobutyric acid (GABA) concentrations in early-stage schizophrenia patients: 3T Proton MRS study. Schizophr. Res. *112*, 192-193.

Goulburn, A.L., Alden, D., Davis, R.P., Micallef, S.J., Ng, E.S., Yu, Q.C., Lim, S.M., Soh, C.L., Elliott, D.A., Hatzistavrou, T., et al. (2011). A targeted NKX2.1 human embryonic stem cell reporter line enables identification of human basal forebrain derivatives. Stem Cells *29*, 462-473.

Goulburn, A.L., Stanley, E.G., Elefanty, A.G., and Anderson, S.A. (2012). Generating GABAergic cerebral cortical interneurons from mouse and human embryonic stem cells. Stem Cell Res *8*, 416-426.

Gray, E.G. (1959). Axo-somatic and axo-dendritic synapses of the cerebral cortex: an electron microscope study. J. Anat. *93*, 420-433.

Grove, E.A., Tole, S., Limon, J., Yip, L., and Ragsdale, C.W. (1998). The hem of the embryonic cerebral cortex is defined by the expression of multiple Wnt genes and is compromised in Gli3-deficient mice. Development *125*, 2315-2325.

Guidotti, A., Auta, J., Davis, J.M., Di-Giorgi-Gerevini, V., Dwivedi, Y., Grayson, D.R., Impagnatiello, F., Pandey, G., Pesold, C., Sharma, R., et al. (2000). Decrease in reelin

and glutamic acid decarboxylase67 (GAD67) expression in schizophrenia and bipolar disorder: a postmortem brain study. Arch. Gen. Psychiatry *57*, 1061-1069.

Guidotti, A., Auta, J., Davis, J.M., Dong, E., Grayson, D.R., Veldic, M., Zhang, X., and Costa, E. (2005). GABAergic dysfunction in schizophrenia: new treatment strategies on the horizon. Psychopharmacology (Berl) *180*, 191-205.

Gulacsi, A., and Anderson, S.A. (2006). Shh maintains Nkx2.1 in the MGE by a Gli3-independent mechanism. Cereb. Cortex *16 Suppl 1*, i89-95.

Gunhaga, L., Marklund, M., Sjodal, M., Hsieh, J.C., Jessell, T.M., and Edlund, T. (2003). Specification of dorsal telencephalic character by sequential Wnt and FGF signaling. Nat. Neurosci. *6*, 701-707.

Gupta, S., and Kulhara, P. (2010). What is schizophrenia: A neurodevelopmental or neurodegenerative disorder or a combination of both? A critical analysis. Indian J. Psychiatry *52*, 21-27.

Hafner, H., Riecher-Rossler, A., Hambrecht, M., Maurer, K., Meissner, S., Schmidtke, A., Fatkenheuer, B., Loffler, W., and van der Heiden, W. (1992). IRAOS: an instrument for the assessment of onset and early course of schizophrenia. Schizophr. Res. *6*, 209-223.

Halabisky, B., Shen, F., Huguenard, J.R., and Prince, D.A. (2006). Electrophysiological classification of somatostatin-positive interneurons in mouse sensorimotor cortex. J. Neurophysiol. *96*, 834-845.

Hansen, D.V., Lui, J.H., Flandin, P., Yoshikawa, K., Rubenstein, J.L., Alvarez-Buylla, A., and Kriegstein, A.R. (2013). Non-epithelial stem cells and cortical interneuron production in the human ganglionic eminences. Nat. Neurosci. *16*, 1576-1587.

Hansen, K.B., Yi, F., Perszyk, R.E., Furukawa, H., Wollmuth, L.P., Gibb, A.J., and Traynelis, S.F. (2018). Structure, function, and allosteric modulation of NMDA receptors. J Gen Physiol *150*, 1081-1105.

Harrison, P.J. (1997). Schizophrenia: a disorder of neurodevelopment? Curr. Opin. Neurobiol. *7*, 285-289.

Harrison, P.J. (1999). The neuropathology of schizophrenia. A critical review of the data and their interpretation. Brain *122 (Pt 4)*, 593-624.

Harrison, P.J. (2015). Recent genetic findings in schizophrenia and their therapeutic relevance. J. Psychopharmacol. *29*, 85-96.

Hashimoto, T., Arion, D., Unger, T., Maldonado-Aviles, J.G., Morris, H.M., Volk, D.W., Mirnics, K., and Lewis, D.A. (2008). Alterations in GABA-related transcriptome in the dorsolateral prefrontal cortex of subjects with schizophrenia. Mol. Psychiatry *13*, 147-161.

Hatakeyama, J., Bessho, Y., Katoh, K., Ookawara, S., Fujioka, M., Guillemot, F., and Kageyama, R. (2004). Hes genes regulate size, shape and histogenesis of the nervous system by control of the timing of neural stem cell differentiation. Development *131*, 5539-5550.

He, M., Tucciarone, J., Lee, S., Nigro, M.J., Kim, Y., Levine, J.M., Kelly, S.M., Krugikov, I., Wu, P., Chen, Y., et al. (2016). Strategies and Tools for Combinatorial Targeting of GABAergic Neurons in Mouse Cerebral Cortex. Neuron *91*, 1228-1243.

Hebert, J.M., and Fishell, G. (2008). The genetics of early telencephalon patterning: some assembly required. Nat. Rev. Neurosci. *9*, 678-685.

Helmstaedter, M., Sakmann, B., and Feldmeyer, D. (2009). Neuronal correlates of local, lateral, and translaminar inhibition with reference to cortical columns. Cereb. Cortex *19*, 926-937.

Hernandez-Miranda, L.R., Parnavelas, J.G., and Chiara, F. (2010). Molecules and mechanisms involved in the generation and migration of cortical interneurons. ASN Neuro *2*, e00031.

Himi, T., Okazaki, T., Wang, H., McNeill, T.H., and Mori, N. (1994). Differential localization of SCG10 and p19/stathmin messenger RNAs in adult rat brain indicates distinct roles for these growth-associated proteins. Neuroscience *60*, 907-926.

Hodge, R.D., Bakken, T.E., Miller, J.A., Smith, K.A., Barkan, E.R., Graybuck, L.T., Close, J.L., Long, B., Johansen, N., Penn, O., et al. (2019). Conserved cell types with divergent features in human versus mouse cortex. Nature *573*, 61-68.

Howard, M.A., Elias, G.M., Elias, L.A., Swat, W., and Nicoll, R.A. (2010). The role of SAP97 in synaptic glutamate receptor dynamics. Proc. Natl. Acad. Sci. U. S. A. *107*, 3805-3810.

Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. Cell *157*, 1262-1278.

Hu, J.S., Vogt, D., Lindtner, S., Sandberg, M., Silberberg, S.N., and Rubenstein, J.L.R. (2017). Coup-TF1 and Coup-TF2 control subtype and laminar identity of MGE-derived neocortical interneurons. Development *144*, 2837-2851.

Huang, S.M., Mishina, Y.M., Liu, S., Cheung, A., Stegmeier, F., Michaud, G.A., Charlat, O., Wiellette, E., Zhang, Y., Wiessner, S., et al. (2009). Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling. Nature *461*, 614-620.

Humbert, P.O., Grzeschik, N.A., Brumby, A.M., Galea, R., Elsum, I., and Richardson, H.E. (2008). Control of tumourigenesis by the Scribble/Dlg/Lgl polarity module. Oncogene *27*, 6888-6907.

Husi, H., and Grant, S.G. (2001). Isolation of 2000-kDa complexes of N-methyl-D-aspartate receptor and postsynaptic density 95 from mouse brain. J. Neurochem. *77*, 281-291.

Itoh, Y., Tyssowski, K., and Gotoh, Y. (2013). Transcriptional coupling of neuronal fate commitment and the onset of migration. Curr. Opin. Neurobiol. *23*, 957-964.

Ivanova, S., Gregorc, U., Vidergar, N., Javier, R., Bredt, D.S., Vandenabeele, P., Pardo, J., Simon, M.M., Turk, V., Banks, L., et al. (2011). MAGUKs, scaffolding proteins at cell

junctions, are substrates of different proteases during apoptosis. Cell Death Dis. *2*, e116.

Jackman, M., Firth, M., and Pines, J. (1995). Human cyclins B1 and B2 are localized to strikingly different structures: B1 to microtubules, B2 primarily to the Golgi apparatus. EMBO J. *14*, 1646-1654.

Jakovcevski, I., Mayer, N., and Zecevic, N. (2011). Multiple origins of human neocortical interneurons are supported by distinct expression of transcription factors. Cereb. Cortex *21*, 1771-1782.

Jiang, X., Shen, S., Cadwell, C.R., Berens, P., Sinz, F., Ecker, A.S., Patel, S., and Tolias, A.S. (2015). Principles of connectivity among morphologically defined cell types in adult neocortex. Science *350*, aac9462.

Jimenez, D., Lopez-Mascaraque, L.M., Valverde, F., and De Carlos, J.A. (2002). Tangential migration in neocortical development. Dev. Biol. *244*, 155-169.

Jones, E.G. (1975). Varieties and distribution of non-pyramidal cells in the somatic sensory cortex of the squirrel monkey. J. Comp. Neurol. *160*, 205-267.

Joyce, E.M., and Roiser, J.P. (2007). Cognitive heterogeneity in schizophrenia. Curr. Opin. Psychiatry *20*, 268-272.

Kaar, S.J., Angelescu, I., Marques, T.R., and Howes, O.D. (2019). Pre-frontal parvalbumin interneurons in schizophrenia: a meta-analysis of post-mortem studies. J. Neural Transm (Vienna) *126*, 1637-1651.

Kageyama, R., and Ohtsuka, T. (1999). The Notch-Hes pathway in mammalian neural development. Cell Res *9*, 179-188.

Kanatani, S., Yozu, M., Tabata, H., and Nakajima, K. (2008). COUP-TFII is preferentially expressed in the caudal ganglionic eminence and is involved in the caudal migratory stream. J. Neurosci. *28*, 13582-13591.

Kang, H.J., Kawasawa, Y.I., Cheng, F., Zhu, Y., Xu, X., Li, M., Sousa, A.M., Pletikos, M., Meyer, K.A., Sedmak, G., et al. (2011). Spatio-temporal transcriptome of the human brain. Nature *478*, 483-489.

Kashyap, V., Rezende, N.C., Scotland, K.B., Shaffer, S.M., Persson, J.L., Gudas, L.J., and Mongan, N.P. (2009). Regulation of stem cell pluripotency and differentiation involves a mutual regulatory circuit of the NANOG, OCT4, and SOX2 pluripotency transcription factors with polycomb repressive complexes and stem cell microRNAs. Stem Cells Dev. *18*, 1093-1108.

Kawaguchi, Y., and Kubota, Y. (1996). Physiological and morphological identification of somatostatin- or vasoactive intestinal polypeptide-containing cells among GABAergic cell subtypes in rat frontal cortex. J. Neurosci. *16*, 2701-2715.

Kawaguchi, Y., and Kubota, Y. (1997). GABAergic cell subtypes and their synaptic connections in rat frontal cortex. Cereb. Cortex *7*, 476-486.

Kawaguchi, Y., and Kubota, Y. (1998). Neurochemical features and synaptic connections of large physiologically-identified GABAergic cells in the rat frontal cortex. Neuroscience *85*, 677-701.

Kelsom, C., and Lu, W. (2013). Development and specification of GABAergic cortical interneurons. Cell Biosci. *3*, 19.

Kepecs, A., and Fishell, G. (2014). Interneuron cell types are fit to function. Nature *505*, 318-326.

Kerrisk, M.E., Cingolani, L.A., and Koleske, A.J. (2014). ECM receptors in neuronal structure, synaptic plasticity, and behavior. Prog. Brain. Res. *214*, 101-131.

Kessaris, N., Magno, L., Rubin, A.N., and Oliveira, M.G. (2014). Genetic programs controlling cortical interneuron fate. Curr. Opin. Neurobiol. *26*, 79-87.

Khandaker, G.M., Zimbron, J., Lewis, G., and Jones, P.B. (2013). Prenatal maternal infection, neurodevelopment and adult schizophrenia: a systematic review of population-based studies. Psychol Med *43*, 239-257.

Khirug, S., Yamada, J., Afzalov, R., Voipio, J., Khiroug, L., and Kaila, K. (2008). GABAergic depolarization of the axon initial segment in cortical principal neurons is caused by the Na-K-2Cl cotransporter NKCC1. J. Neurosci. *28*, 4635-4639.

Kim, J.H., Liao, D., Lau, L.-F., and Huganir, R.L. (1998). SynGAP: a Synaptic RasGAP that Associates with the PSD-95/SAP90 Protein Family. Neuron *20*, 683-691.

Kim, T.G., Yao, R., Monnell, T., Cho, J.H., Vasudevan, A., Koh, A., Peeyush, K.T., Moon, M., Datta, D., Bolshakov, V.Y., et al. (2014). Efficient specification of interneurons from human pluripotent stem cells by dorsoventral and rostrocaudal modulation. Stem Cells *32*, 1789-1804.

Kimura, K., Mizoguchi, A., and Ide, C. (2003). Regulation of growth cone extension by SNARE proteins. J. Histochem. Cytochem. *51*, 429-433.

Kippin, T.E., Martens, D.J., and van der Kooy, D. (2005). p21 loss compromises the relative quiescence of forebrain stem cell proliferation leading to exhaustion of their proliferation capacity. Genes Dev. *19*, 756-767.

Kirov, G., Pocklington, A.J., Holmans, P., Ivanov, D., Ikeda, M., Ruderfer, D., Moran, J., Chambert, K., Toncheva, D., Georgieva, L., et al. (2012). De novo CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia. Mol. Psychiatry *17*, 142-153.

Klausberger, T., Magill, P.J., Marton, L.F., Roberts, J.D., Cobden, P.M., Buzsaki, G., and Somogyi, P. (2003). Brain-state- and cell-type-specific firing of hippocampal interneurons in vivo. Nature *421*, 844-848.

Kringlen, E. (2000). Twin studies in schizophrenia with special emphasis on concordance figures. Am. J. Med. Genet. *97*, 4-11.

Kristiansen, L.V., Beneyto, M., Haroutunian, V., and Meador-Woodruff, J.H. (2006). Changes in NMDA receptor subunits and interacting PSD proteins in dorsolateral prefrontal and anterior cingulate cortex indicate abnormal regional expression in schizophrenia. Mol. Psychiatry *11*, 737-747, 705.

Kruger, J.M., Favaro, P.D., Liu, M., Kitlinska, A., Huang, X., Raabe, M., Akad, D.S., Liu, Y., Urlaub, H., Dong, Y., et al. (2013). Differential roles of postsynaptic density-93 isoforms in regulating synaptic transmission. J. Neurosci. *33*, 15504-15517.

Kumar, P., Tan, Y., and Cahan, P. (2017). Understanding development and stem cells using single cell-based analyses of gene expression. Development *144*, 17-32.

Kushima, I., Nakatochi, M., Aleksic, B., Okada, T., Kimura, H., Kato, H., Morikawa, M., Inada, T., Ishizuka, K., Torii, Y., et al. (2022). Cross-disorder analysis of genic and regulatory copy number variations in bipolar disorder, schizophrenia, and autism spectrum disorder. Biol. Psychiatry.

Laclef, C., and Metin, C. (2018). Conserved rules in embryonic development of cortical interneurons. Semin. Cell Dev. Biol. *76*, 86-100.

Lam, M., Chen, C.Y., Li, Z., Martin, A.R., Bryois, J., Ma, X., Gaspar, H., Ikeda, M., Benyamin, B., Brown, B.C., et al. (2019). Comparative genetic architectures of schizophrenia in East Asian and European populations. Nat. Genet. *51*, 1670-1678.

Lariviere, R.C., and Julien, J.P. (2004). Functions of intermediate filaments in neuronal development and disease. J. Neurobiol. *58*, 131-148.

Lasser, M., Tiber, J., and Lowery, L.A. (2018). The Role of the Microtubule Cytoskeleton in Neurodevelopmental Disorders. Front. Cell. Neurosci. *12*, 165.

Laurent, L.C., Ulitsky, I., Slavin, I., Tran, H., Schork, A., Morey, R., Lynch, C., Harness, J.V., Lee, S., Barrero, M.J., et al. (2011). Dynamic changes in the copy number of pluripotency and cell proliferation genes in human ESCs and iPSCs during reprogramming and time in culture. Cell Stem Cell *8*, 106-118.

Lavdas, A.A., Grigoriou, M., Pachnis, V., and Parnavelas, J.G. (1999). The medial ganglionic eminence gives rise to a population of early neurons in the developing cerebral cortex. J. Neurosci. *19*, 7881-7888.

Le Bon-Jego, M., and Yuste, R. (2007). Persistently active, pacemaker-like neurons in neocortex. Front. Neurosci. *1*, 123-129.

Lee, S., Hjerling-Leffler, J., Zagha, E., Fishell, G., and Rudy, B. (2010). The largest group of superficial neocortical GABAergic interneurons expresses ionotropic serotonin receptors. J. Neurosci. *30*, 16796-16808.

Lee, T.I., and Young, R.A. (2013). Transcriptional regulation and its misregulation in disease. Cell *152*, 1237-1251.

Leeson, V.C., Robbins, T.W., Matheson, E., Hutton, S.B., Ron, M.A., Barnes, T.R., and Joyce, E.M. (2009). Discrimination learning, reversal, and set-shifting in first-episode

schizophrenia: stability over six years and specific associations with medication type and disorganization syndrome. Biol. Psychiatry *66*, 586-593.

Letinic, K., and Kostovic, I. (1998). Postnatal development of calcium-binding proteins calbindin and parvalbumin in human visual cortex. Cereb. Cortex *8*, 660-669.

Letinic, K., Zoncu, R., and Rakic, P. (2002). Origin of GABAergic neurons in the human neocortex. Nature *417*, 645-649.

Levitt, P., Eagleson, K.L., and Powell, E.M. (2004). Regulation of neocortical interneuron development and the implications for neurodevelopmental disorders. Trends Neurosci. *27*, 400-406.

Lewis, D.A., Hashimoto, T., and Volk, D.W. (2005). Cortical inhibitory neurons and schizophrenia. Nat. Rev. Neurosci. *6*, 312-324.

Li, G., Adesnik, H., Li, J., Long, J., Nicoll, R.A., Rubenstein, J.L., and Pleasure, S.J. (2008). Regional distribution of cortical interneurons and development of inhibitory tone are regulated by Cxcl12/Cxcr4 signaling. J. Neurosci. *28*, 1085-1098.

Liebau, S., Vaida, B., Storch, A., and Boeckers, T.M. (2007). Maturation of synaptic contacts in differentiating neural stem cells. Stem Cells *25*, 1720-1729.

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

Lin, M., Pedrosa, E., Hrabovsky, A., Chen, J., Puliafito, B.R., Gilbert, S.R., Zheng, D., and Lachman, H.M. (2016). Integrative transcriptome network analysis of iPSC-derived neurons from schizophrenia and schizoaffective disorder patients with 22q11.2 deletion. BMC Syst. Biol. *10*, 105.

Liu, P., Kaplan, A., Yuan, B., Hanna, J.H., Lupski, J.R., and Reiner, O. (2014). Passage number is a major contributor to genomic structural variations in mouse iPSCs. Stem Cells *32*, 2657-2667.

Liu, Y., Liu, H., Sauvey, C., Yao, L., Zarnowska, E.D., and Zhang, S.C. (2013). Directed differentiation of forebrain GABA interneurons from human pluripotent stem cells. Nat. Protoc. *8*, 1670-1679.

Lodato, S., Tomassy, G.S., De Leonibus, E., Uzcategui, Y.G., Andolfi, G., Armentano, M., Touzot, A., Gaztelu, J.M., Arlotta, P., Menendez de la Prida, L., et al. (2011). Loss of COUP-TFI alters the balance between caudal ganglionic eminence- and medial ganglionic eminence-derived cortical interneurons and results in resistance to epilepsy. J. Neurosci. *31*, 4650-4662.

Long, J.E., Cobos, I., Potter, G.B., and Rubenstein, J.L. (2009). Dlx1&2 and Mash1 transcription factors control MGE and CGE patterning and differentiation through parallel and overlapping pathways. Cereb. Cortex *19 Suppl 1*, i96-106.

Long, K.R., Newland, B., Florio, M., Kalebic, N., Langen, B., Kolterer, A., Wimberger, P., and Huttner, W.B. (2018). Extracellular Matrix Components HAPLN1, Lumican, and

Collagen I Cause Hyaluronic Acid-Dependent Folding of the Developing Human Neocortex. Neuron *99*, 702-719 e706.

Lopez-Bendito, G., Sanchez-Alcaniz, J.A., Pla, R., Borrell, V., Pico, E., Valdeolmillos, M., and Marin, O. (2008). Chemokine signaling controls intracortical migration and final distribution of GABAergic interneurons. J. Neurosci. *28*, 1613-1624.

Luhmann, H.J., Sinning, A., Yang, J.W., Reyes-Puerta, V., Stuttgen, M.C., Kirischuk, S., and Kilb, W. (2016). Spontaneous Neuronal Activity in Developing Neocortical Networks: From Single Cells to Large-Scale Interactions. Front. Neural Circuits *10*, 40.

Ma, C., Gu, C., Huo, Y., Li, X., and Luo, X.J. (2018). The integrated landscape of causal genes and pathways in schizophrenia. Transl. Psychiatry *8*, 67.

Ma, T., Wang, C., Wang, L., Zhou, X., Tian, M., Zhang, Q., Zhang, Y., Li, J., Liu, Z., Cai, Y., et al. (2013). Subcortical origins of human and monkey neocortical interneurons. Nat. Neurosci. *16*, 1588-1597.

Ma, T., Zhang, Q., Cai, Y., You, Y., Rubenstein, J.L., and Yang, Z. (2012). A subpopulation of dorsal lateral/caudal ganglionic eminence-derived neocortical interneurons expresses the transcription factor Sp8. Cereb. Cortex *22*, 2120-2130.

Ma, Y., Hu, H., Berrebi, A.S., Mathers, P.H., and Agmon, A. (2006). Distinct subtypes of somatostatin-containing neocortical interneurons revealed in transgenic mice. J. Neurosci. *26*, 5069-5082.

MacLaren, E.J., Charlesworth, P., Coba, M.P., and Grant, S.G. (2011). Knockdown of mental disorder susceptibility genes disrupts neuronal network physiology in vitro. Mol. Cell. Neurosci. *47*, 93-99.

Maden, M. (2007). Retinoic acid in the development, regeneration and maintenance of the nervous system. Nat. Rev. Neurosci. *8*, 755-765.

Maitra, A., Arking, D.E., Shivapurkar, N., Ikeda, M., Stastny, V., Kassauei, K., Sui, G., Cutler, D.J., Liu, Y., Brimble, S.N., et al. (2005). Genomic alterations in cultured human embryonic stem cells. Nat. Genet. *37*, 1099-1103.

Mani, S., Shen, Y., Schaefer, J., and Meiri, K.F. (2001). Failure to express GAP-43 during neurogenesis affects cell cycle regulation and differentiation of neural precursors and stimulates apoptosis of neurons. Mol. Cell. Neurosci. *17*, 54-66.

Marin, O., Anderson, S.A., and Rubenstein, J.L. (2000). Origin and molecular specification of striatal interneurons. J. Neurosci. *20*, 6063-6076.

Marin, O., and Rubenstein, J.L. (2001). A long, remarkable journey: tangential migration in the telencephalon. Nat. Rev. Neurosci. *2*, 780-790.

Marin, O., Valdeolmillos, M., and Moya, F. (2006). Neurons in motion: same principles for different shapes? Trends Neurosci. *29*, 655-661.

Markram, H., Toledo-Rodriguez, M., Wang, Y., Gupta, A., Silberberg, G., and Wu, C. (2004). Interneurons of the neocortical inhibitory system. Nat. Rev. Neurosci. *5*, 793-807.

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., et al. (2013). Directed differentiation and functional maturation of cortical interneurons from human embryonic stem cells. Cell Stem Cell *12*, 559-572.

Mayer, C., Hafemeister, C., Bandler, R.C., Machold, R., Batista Brito, R., Jaglin, X., Allaway, K., Butler, A., Fishell, G., and Satija, R. (2018). Developmental diversification of cortical inhibitory interneurons. Nature *555*, 457-462.

McGarry, L.M., Packer, A.M., Fino, E., Nikolenko, V., Sippy, T., and Yuste, R. (2010). Quantitative classification of somatostatin-positive neocortical interneurons identifies three interneuron subtypes. Front. Neural Circuits *4*, 12.

McGee, A.W., Topinka, J.R., Hashimoto, K., Petralia, R.S., Kakizawa, S., Kauer, F., Aguilera-Moreno, A., Wenthold, R.J., Kano, M., and Bredt, D.S. (2001). PSD-93 Knock-Out Mice Reveal That Neuronal MAGUKs Are Not Required for Development or Function of Parallel Fiber Synapses in Cerebellum. J. Neurosci. *21*, 3085-3091.

McTague, A., Rossignoli, G., Ferrini, A., Barral, S., and Kurian, M.A. (2021). Genome Editing in iPSC-Based Neural Systems: From Disease Models to Future Therapeutic Strategies. Front Genome Ed *3*, 630600.

Mei, L., and Xiong, W.C. (2008). Neuregulin 1 in neural development, synaptic plasticity and schizophrenia. Nat. Rev. Neurosci. *9*, 437-452.

Mendoza, C., Olguin, P., Lafferte, G., Thomas, U., Ebitsch, S., Gundelfinger, E.D., Kukuljan, M., and Sierralta, J. (2003). Novel isoforms of Dlg are fundamental for neuronal development in Drosophila. J. Neurosci. *23*, 2093-2101.

Menezes, J.R., and Luskin, M.B. (1994). Expression of neuron-specific tubulin defines a novel population in the proliferative layers of the developing telencephalon. J. Neurosci. *14*, 5399-5416.

Metin, C., Baudoin, J.P., Rakic, S., and Parnavelas, J.G. (2006). Cell and molecular mechanisms involved in the migration of cortical interneurons. Eur. J. Neurosci. *23*, 894-900.

Meyer, G., Schaaps, J.P., Moreau, L., and Goffinet, A.M. (2000). Embryonic and early fetal development of the human neocortex. J. Neurosci. *20*, 1858-1868.

Mi, D., Li, Z., Lim, L., Li, M., Moissidis, M., Yang, Y., Gao, T., Hu, T.X., Pratt, T., Price, D.J., et al. (2018). Early emergence of cortical interneuron diversity in the mouse embryo. Science *360*, 81-85.

Miller, M.W. (1985). Cogeneration of retrogradely labeled corticocortical projection and GABA-immunoreactive local circuit neurons in cerebral cortex. Brain Res. *355*, 187-192.

Miyoshi, G., Butt, S.J., Takebayashi, H., and Fishell, G. (2007). Physiologically distinct temporal cohorts of cortical interneurons arise from telencephalic Olig2-expressing precursors. J. Neurosci. *27*, 7786-7798.

Miyoshi, G., Hjerling-Leffler, J., Karayannis, T., Sousa, V.H., Butt, S.J., Battiste, J., Johnson, J.E., Machold, R.P., and Fishell, G. (2010). Genetic fate mapping reveals that the caudal ganglionic eminence produces a large and diverse population of superficial cortical interneurons. J. Neurosci. *30*, 1582-1594.

Miyoshi, G., Young, A., Petros, T., Karayannis, T., McKenzie Chang, M., Lavado, A., Iwano, T., Nakajima, M., Taniguchi, H., Huang, Z.J., et al. (2015). Prox1 Regulates the Subtype-Specific Development of Caudal Ganglionic Eminence-Derived GABAergic Cortical Interneurons. J. Neurosci. *35*, 12869-12889.

Morris, G.F., and Mathews, M.B. (1989). Regulation of proliferating cell nuclear antigen during the cell cycle. J. Biol. Chem. *264*, 13856-13864.

Mowry, B.J., and Gratten, J. (2013). The emerging spectrum of allelic variation in schizophrenia: current evidence and strategies for the identification and functional characterization of common and rare variants. Mol. Psychiatry *18*, 38-52.

Mueser, K.T., and McGurk, S.R. (2004). Schizophrenia. Lancet *363*, 2063-2072.

Murray, R.M., and Lewis, S.W. (1987). Is schizophrenia a neurodevelopmental disorder? Br. Med. J. (Clin Res Ed). *295*, 681-682.

Muzio, L., Di Benedetto, B., Stoykova, A., Boncinelli, E., Gruss, P., and Mallamaci, A. (2002). Conversion of cerebral cortex into basal ganglia in Emx2(-/-) Pax6(Sey/Sey) double-mutant mice. Nat. Neurosci. *5*, 737-745.

Nadarajah, B., Alifragis, P., Wong, R.O., and Parnavelas, J.G. (2002). Ventricle-directed migration in the developing cerebral cortex. Nat. Neurosci. *5*, 218-224.

Naka, H., Nakamura, S., Shimazaki, T., and Okano, H. (2008). Requirement for COUP-TFI and II in the temporal specification of neural stem cells in CNS development. Nat. Neurosci. *11*, 1014-1023.

Nakamura, Y., Sakakibara, S.-i., Miyata, T., Ogawa, M., Shimazaki, T., Weiss, S., Kageyama, R., and Okano, H. (2000). The bHLH GeneHes1as a Repressor of the Neuronal Commitment of CNS Stem Cells. J. Neurosci. *20*, 283-293.

Nakazawa, K., Zsiros, V., Jiang, Z., Nakao, K., Kolata, S., Zhang, S., and Belforte, J.E. (2012). GABAergic interneuron origin of schizophrenia pathophysiology. Neuropharmacology *62*, 1574-1583.

Nazir, F.H., Becker, B., Brinkmalm, A., Hoglund, K., Sandelius, A., Bergstrom, P., Satir, T.M., Ohrfelt, A., Blennow, K., Agholme, L., et al. (2018). Expression and secretion of synaptic proteins during stem cell differentiation to cortical neurons. Neurochem. Int. *121*, 38-49.

Neddens, J., Fish, K.N., Tricoire, L., Vullhorst, D., Shamir, A., Chung, W., Lewis, D.A., McBain, C.J., and Buonanno, A. (2011). Conserved interneuron-specific ErbB4

expression in frontal cortex of rodents, monkeys, and humans: implications for schizophrenia. Biol. Psychiatry *70*, 636-645.

Nery, S., Fishell, G., and Corbin, J.G. (2002). The caudal ganglionic eminence is a source of distinct cortical and subcortical cell populations. Nat. Neurosci. *5*, 1279-1287.

Nery, S., Wichterle, H., and Fishell, G. (2001). Sonic hedgehog contributes to oligodendrocyte specification in the mammalian forebrain. Development *128*, 527-540.

Network, and Pathway Analysis Subgroup of Psychiatric Genomics, C. (2015). Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. Nat. Neurosci. *18*, 199-209.

Nicholas, C.R., Chen, J., Tang, Y., Southwell, D.G., Chalmers, N., Vogt, D., Arnold, C.M., Chen, Y.J., Stanley, E.G., Elefanty, A.G., et al. (2013). Functional maturation of hPSC-derived forebrain interneurons requires an extended timeline and mimics human neural development. Cell Stem Cell *12*, 573-586.

Nithianantharajah, J., Komiyama, N.H., McKechanie, A., Johnstone, M., Blackwood, D.H., St Clair, D., Emes, R.D., van de Lagemaat, L.N., Saksida, L.M., Bussey, T.J., et al. (2013). Synaptic scaffold evolution generated components of vertebrate cognitive complexity. Nat. Neurosci. *16*, 16-24.

Nobrega-Pereira, S., Kessaris, N., Du, T., Kimura, S., Anderson, S.A., and Marin, O. (2008). Postmitotic Nkx2-1 controls the migration of telencephalic interneurons by direct repression of guidance receptors. Neuron *59*, 733-745.

Nowakowski, T.J., Bhaduri, A., Pollen, A.A., Alvarado, B., Mostajo-Radji, M.A., Di Lullo, E., Haeussler, M., Sandoval-Espinosa, C., Liu, S.J., Velmeshev, D., et al. (2017). Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. Science *358*, 1318-1323.

Nozumi, M., Togano, T., Takahashi-Niki, K., Lu, J., Honda, A., Taoka, M., Shinkawa, T., Koga, H., Takeuchi, K., Isobe, T., et al. (2009). Identification of functional marker proteins in the mammalian growth cone. Proc. Natl. Acad. Sci. U. S. A. *106*, 17211-17216.

Olah, S., Komlosi, G., Szabadics, J., Varga, C., Toth, E., Barzo, P., and Tamas, G. (2007). Output of neurogliaform cells to various neuron types in the human and rat cerebral cortex. Front. Neural Circuits *1*, 4.

Oliva, C., Escobedo, P., Astorga, C., Molina, C., and Sierralta, J. (2012). Role of the MAGUK protein family in synapse formation and function. Dev. Neurobiol. *72*, 57-72.

Olsson, M., Bjorklund, A., and Campbell, K. (1998). Early specification of striatal projection neurons and interneuronal subtypes in the lateral and medial ganglionic eminence. Neuroscience *84*, 867-876.

Ongur, D., Prescot, A.P., McCarthy, J., Cohen, B.M., and Renshaw, P.F. (2010). Elevated gamma-aminobutyric acid levels in chronic schizophrenia. Biol. Psychiatry *68*, 667-670.

Orquera, D.P., Nasif, S., Low, M.J., Rubinstein, M., and de Souza, F.S.J. (2016). Essential function of the transcription factor Rax in the early patterning of the mammalian hypothalamus. Dev. Biol. *416*, 212-224.

Osen-Sand, A., Catsicas, M., Staple, J.K., Jones, K.A., Ayala, G., Knowles, J., Grenningloh, G., and Catsicas, S. (1993). Inhibition of axonal growth by SNAP-25 antisense oligonucleotides in vitro and in vivo. Nature *364*, 445-448.

Overstreet-Wadiche, L., and McBain, C.J. (2015). Neurogliaform cells in cortical circuits. Nat. Rev. Neurosci. *16*, 458-468.

Palay, S.L. (1956). Synapses in the central nervous system. J. Biophys. Biochem. Cytol. *2*, 193-202.

Pardinas, A.F., Holmans, P., Pocklington, A.J., Escott-Price, V., Ripke, S., Carrera, N., Legge, S.E., Bishop, S., Cameron, D., Hamshere, M.L., et al. (2018). Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. Nat. Genet. *50*, 381-389.

Paredes, M.F., James, D., Gil-Perotin, S., Kim, H., Cotter, J.A., Ng, C., Sandoval, K., Rowitch, D.H., Xu, D., McQuillen, P.S., et al. (2016). Extensive migration of young neurons into the infant human frontal lobe. Science *354*.

Pass, R., Haan, N., Humby, T., Wilkinson, L.S., Hall, J., and Thomas, K.L. (2022). Selective behavioural impairments in mice heterozygous for the cross disorder psychiatric risk gene DLG2. Genes Brain Behav. *21*, e12799.

Patel, K.R., Cherian, J., Gohil, K., and Atkinson, D. (2014). Schizophrenia: overview and treatment options. P T *39*, 638-645.

Patz, S., Grabert, J., Gorba, T., Wirth, M.J., and Wahle, P. (2004). Parvalbumin expression in visual cortical interneurons depends on neuronal activity and TrkB ligands during an Early period of postnatal development. Cereb. Cortex *14*, 342-351.

Petanjek, Z., Berger, B., and Esclapez, M. (2009a). Origins of cortical GABAergic neurons in the cynomolgus monkey. Cereb. Cortex *19*, 249-262.

Petanjek, Z., Kostovic, I., and Esclapez, M. (2009b). Primate-specific origins and migration of cortical GABAergic neurons. Front. Neuroanat. *3*, 26.

Petilla Interneuron Nomenclature Group, Ascoli, G.A., Alonso-Nanclares, L., Anderson, S.A., Barrionuevo, G., Benavides-Piccione, R., Burkhalter, A., Buzsaki, G., Cauli, B., Defelipe, J., et al. (2008). Petilla terminology: nomenclature of features of GABAergic interneurons of the cerebral cortex. Nat. Rev. Neurosci. *9*, 557-568.

Petros, T.J., Bultje, R.S., Ross, M.E., Fishell, G., and Anderson, S.A. (2015). Apical versus Basal Neurogenesis Directs Cortical Interneuron Subclass Fate. Cell Rep. *13*, 1090-1095.

Peyre, E., Silva, C.G., and Nguyen, L. (2015). Crosstalk between intracellular and extracellular signals regulating interneuron production, migration and integration into the cortex. Front. Cell. Neurosci. *9*, 129.

Philips, A.K., Siren, A., Avela, K., Somer, M., Peippo, M., Ahvenainen, M., Doagu, F., Arvio, M., Kaariainen, H., Van Esch, H., et al. (2014). X-exome sequencing in Finnish families with intellectual disability--four novel mutations and two novel syndromic phenotypes. Orphanet J Rare Dis *9*, 49.

Pinto, D.J., Brumberg, J.C., and Simons, D.J. (2000). Circuit dynamics and coding strategies in rodent somatosensory cortex. J. Neurophysiol. *83*, 1158-1166.

Pocklington, A.J., Rees, E., Walters, J.T., Han, J., Kavanagh, D.H., Chambert, K.D., Holmans, P., Moran, J.L., McCarroll, S.A., Kirov, G., et al. (2015). Novel Findings from CNVs Implicate Inhibitory and Excitatory Signaling Complexes in Schizophrenia. Neuron *86*, 1203-1214.

Poglia, L., Muller, D., and Nikonenko, I. (2011). Ultrastructural modifications of spine and synapse morphology by SAP97. Hippocampus *21*, 990-998.

Polioudakis, D., de la Torre-Ubieta, L., Langerman, J., Elkins, A.G., Shi, X., Stein, J.L., Vuong, C.K., Nichterwitz, S., Gevorgian, M., Opland, C.K., et al. (2019). A Single-Cell Transcriptomic Atlas of Human Neocortical Development during Mid-gestation. Neuron *103*, 785-801 e788.

Pouille, F., Marin-Burgin, A., Adesnik, H., Atallah, B.V., and Scanziani, M. (2009). Input normalization by global feedforward inhibition expands cortical dynamic range. Nat. Neurosci. *12*, 1577-1585.

Pronneke, A., Scheuer, B., Wagener, R.J., Mock, M., Witte, M., and Staiger, J.F. (2015). Characterizing VIP Neurons in the Barrel Cortex of VIPcre/tdTomato Mice Reveals Layer-Specific Differences. Cereb. Cortex *25*, 4854-4868.

Purcell, S.M., Moran, J.L., Fromer, M., Ruderfer, D., Solovieff, N., Roussos, P., O'Dushlaine, C., Chambert, K., Bergen, S.E., Kahler, A., et al. (2014). A polygenic burden of rare disruptive mutations in schizophrenia. Nature *506*, 185-190.

Qi, J., Su, Y., Sun, R., Zhang, F., Luo, X., Yang, Z., and Luo, X. (2005). CASK inhibits ECV304 cell growth and interacts with Id1. Biochem. Biophys. Res. Commun. *328*, 517-521.

Raghanti, M.A., Conley, T., Sudduth, J., Erwin, J.M., Stimpson, C.D., Hof, P.R., and Sherwood, C.C. (2013). Neuropeptide Y-immunoreactive neurons in the cerebral cortex of humans and other haplorrhine primates. Am. J. Primatol. *75*, 415-424.

Rakic, S., and Zecevic, N. (2003a). Early oligodendrocyte progenitor cells in the human fetal telencephalon. Glia *41*, 117-127.

Rakic, S., and Zecevic, N. (2003b). Emerging complexity of layer I in human cerebral cortex. Cereb. Cortex *13*, 1072-1083.

Rallu, M., Machold, R., Gaiano, N., Corbin, J.G., McMahon, A.P., and Fishell, G. (2002). Dorsoventral patterning is established in the telencephalon of mutants lacking both Gli3 and Hedgehog signaling. Development *129*, 4963-4974.

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

Reggiani, C., Coppens, S., Sekhara, T., Dimov, I., Pichon, B., Lufin, N., Addor, M.C., Belligni, E.F., Digilio, M.C., Faletra, F., et al. (2017). Novel promoters and coding first exons in DLG2 linked to developmental disorders and intellectual disability. Genome Med. *9*, 67.

Rhinn, M., and Brand, M. (2001). The midbrain--hindbrain boundary organizer. Curr. Opin. Neurobiol. *11*, 34-42.

Rico, B., and Marin, O. (2011). Neuregulin signaling, cortical circuitry development and schizophrenia. Curr. Opin. Genet. Dev. *21*, 262-270.

Riederer, B.M., Pellier, V., Antonsson, B., Di Paolo, G., Stimpson, S.A., Lutjens, R., Catsicas, S., and Grenningloh, G. (1997). Regulation of microtubule dynamics by the neuronal growth-associated protein SCG10. Proc. Natl. Acad. Sci. U. S. A. *94*, 741-745.

Rigby, M.J., Gomez, T.M., and Puglielli, L. (2020). Glial Cell-Axonal Growth Cone Interactions in Neurodevelopment and Regeneration. Front. Neurosci. *14*, 203.

Ripke, S., O'Dushlaine, C., Chambert, K., Moran, J.L., Kahler, A.K., Akterin, S., Bergen, S.E., Collins, A.L., Crowley, J.J., Fromer, M., et al. (2013). Genome-wide association analysis identifies 13 new risk loci for schizophrenia. Nat. Genet. *45*, 1150-1159.

Rosenbaum, D.M., Rasmussen, S.G., and Kobilka, B.K. (2009). The structure and function of G-protein-coupled receptors. Nature *459*, 356-363.

Rosenthal, N., and Brown, S. (2007). The mouse ascending: perspectives for human-disease models. Nat. Cell Biol. *9*, 993-999.

Ross, M.E. (2011). Cell cycle regulation and interneuron production. Dev. Neurobiol. *71*, 2-9.

Rossignol, E. (2011). Genetics and function of neocortical GABAergic interneurons in neurodevelopmental disorders. Neural Plast *2011*, 649325.

Roussos, P., Guennewig, B., Kaczorowski, D.C., Barry, G., and Brennand, K.J. (2016). Activity-Dependent Changes in Gene Expression in Schizophrenia Human-Induced Pluripotent Stem Cell Neurons. JAMA Psychiatry *73*, 1180-1188.

Rozario, T., and DeSimone, D.W. (2010). The extracellular matrix in development and morphogenesis: a dynamic view. Dev. Biol. *341*, 126-140.

Rudy, B., Fishell, G., Lee, S., and Hjerling-Leffler, J. (2011). Three groups of interneurons account for nearly 100% of neocortical GABAergic neurons. Dev. Neurobiol. *71*, 45-61.

Rujescu, D., Ingason, A., Cichon, S., Pietilainen, O.P., Barnes, M.R., Toulopoulou, T., Picchioni, M., Vassos, E., Ettinger, U., Bramon, E., et al. (2009). Disruption of the neurexin 1 gene is associated with schizophrenia. Hum. Mol. Genet. *18*, 988-996.

Rund, B.R. (1998). A review of longitudinal studies of cognitive functions in schizophrenia patients. Schizophr. Bull. *24*, 425-435.

Sahara, S., Yanagawa, Y., O'Leary, D.D., and Stevens, C.F. (2012). The fraction of cortical GABAergic neurons is constant from near the start of cortical neurogenesis to adulthood. J. Neurosci. *32*, 4755-4761.

Sanders, B., D'Andrea, D., Collins, M.O., Rees, E., Steward, T.G.J., Zhu, Y., Chapman, G., Legge, S.E., Pardinas, A.F., Harwood, A.J., et al. (2022). Transcriptional programs regulating neuronal differentiation are disrupted in DLG2 knockout human embryonic stem cells and enriched for schizophrenia and related disorders risk variants. Nat. Commun. *13*, 27.

Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., and Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. Nat. Biotechnol. *33*, 495-502.

Sato, Y., Tao, Y.X., Su, Q., and Johns, R.A. (2008). Post-synaptic density-93 mediates tyrosine-phosphorylation of the N-methyl-D-aspartate receptors. Neuroscience *153*, 700-708.

Schizophrenia Working Group of the Psychiatric Genomics Consortium (2014). Biological insights from 108 schizophrenia-associated genetic loci. Nature *511*, 421-427.

Schluter, O.M., Xu, W., and Malenka, R.C. (2006). Alternative N-terminal domains of PSD-95 and SAP97 govern activity-dependent regulation of synaptic AMPA receptor function. Neuron *51*, 99-111.

Schneider, F., Duong, T.A., Metz, I., Winkelmeier, J., Hubner, C.A., Endesfelder, U., and Rust, M.B. (2021). Mutual functional dependence of cyclase-associated protein 1 (CAP1) and cofilin1 in neuronal actin dynamics and growth cone function. Prog. Neurobiol. *202*, 102050.

Schnell, E., Sizemore, M., Karimzadegan, S., Chen, L., Bredt, D.S., and Nicoll, R.A. (2002). Direct interactions between PSD-95 and stargazin control synaptic AMPA receptor number. Proc. Natl. Acad. Sci. U. S. A. *99*, 13902-13907.

Semple, B.D., Blomgren, K., Gimlin, K., Ferriero, D.M., and Noble-Haeusslein, L.J. (2013). Brain development in rodents and humans: Identifying benchmarks of maturation and vulnerability to injury across species. Prog. Neurobiol. *106-107*, 1-16.

Sigmund, C.D. (2000). Viewpoint: are studies in genetically altered mice out of control? Arterioscler Thromb Vasc Biol *20*, 1425-1429.

Silberberg, G., and Markram, H. (2007). Disynaptic inhibition between neocortical pyramidal cells mediated by Martinotti cells. Neuron *53*, 735-746.

Silbereis, J.C., Pochareddy, S., Zhu, Y., Li, M., and Sestan, N. (2016). The Cellular and Molecular Landscapes of the Developing Human Central Nervous System. Neuron *89*, 248-268.

Singh, T., Poterba, T., Curtis, D., Akil, H., Al Eissa, M., Barchas, J.D., Bass, N., Bigdeli, T.B., Breen, G., Bromet, E.J., et al. (2022). Rare coding variants in ten genes confer substantial risk for schizophrenia. Nature *604*, 509-516.

Sinha, S., and Chen, J.K. (2006). Purmorphamine activates the Hedgehog pathway by targeting Smoothened. Nat Chem Biol *2*, 29-30.

Skarnes, W.C., Rosen, B., West, A.P., Koutsourakis, M., Bushell, W., Iyer, V., Mujica, A.O., Thomas, M., Harrow, J., Cox, T., et al. (2011). A conditional knockout resource for the genome-wide study of mouse gene function. Nature *474*, 337-342.

Skene, N.G., Bryois, J., Bakken, T.E., Breen, G., Crowley, J.J., Gaspar, H.A., Giusti-Rodriguez, P., Hodge, R.D., Miller, J.A., Munoz-Manchado, A.B., et al. (2018). Genetic identification of brain cell types underlying schizophrenia. Nat. Genet. *50*, 825-833.

Sohal, V.S., Zhang, F., Yizhar, O., and Deisseroth, K. (2009). Parvalbumin neurons and gamma rhythms enhance cortical circuit performance. Nature *459*, 698-702.

Solanki, R.K., Singh, P., Midha, A., and Chugh, K. (2008). Schizophrenia: Impact on quality of life. Indian J. Psychiatry *50*, 181-186.

Soler, J., Fañanás, L., Parellada, M., Krebs, M.-O., Rouleau, G.A., and Fatjó-Vilas, M. (2018). Genetic variability in scaffolding proteins and risk for schizophrenia and autism-spectrum disorders: a systematic review. J. Psychiatry Neurosci. *43*, 223-244.

Soltani, M.H., Pichardo, R., Song, Z., Sangha, N., Camacho, F., Satyamoorthy, K., Sangueza, O.P., and Setaluri, V. (2005). Microtubule-associated protein 2, a marker of neuronal differentiation, induces mitotic defects, inhibits growth of melanoma cells, and predicts metastatic potential of cutaneous melanoma. Am. J. Pathol. *166*, 1841-1850.

Somogyi, P. (1977). A specific 'axo-axonal' interneuron in the visual cortex of the rat. Brain Res. *136*, 345-350.

Sonenberg, N., and Hinnebusch, A.G. (2009). Regulation of translation initiation in eukaryotes: mechanisms and biological targets. Cell *136*, 731-745.

Sousa, V.H., and Fishell, G. (2010). Sonic hedgehog functions through dynamic changes in temporal competence in the developing forebrain. Curr. Opin. Genet. Dev. *20*, 391-399.

Spitzer, N.C. (2006). Electrical activity in early neuronal development. Nature *444*, 707-712.

Stanton, B.Z., and Peng, L.F. (2010). Small-molecule modulators of the Sonic Hedgehog signaling pathway. Mol Biosyst *6*, 44-54.

Stephens, R., Lim, K., Portela, M., Kvansakul, M., Humbert, P.O., and Richardson, H.E. (2018). The Scribble Cell Polarity Module in the Regulation of Cell Signaling in Tissue Development and Tumorigenesis. J. Mol. Biol. *430*, 3585-3612.

Stipp, C.S. (2010). Laminin-binding integrins and their tetraspanin partners as potential antimetastatic targets. Expert reviews in molecular medicine *12*.

Straub, R.E., Lipska, B.K., Egan, M.F., Goldberg, T.E., Callicott, J.H., Mayhew, M.B., Vakkalanka, R.K., Kolachana, B.S., Kleinman, J.E., and Weinberger, D.R. (2007). Allelic variation in GAD1 (GAD67) is associated with schizophrenia and influences cortical function and gene expression. Mol. Psychiatry *12*, 854-869.

Stumm, R.K., Zhou, C., Ara, T., Lazarini, F., Dubois-Dalcq, M., Nagasawa, T., Hollt, V., and Schulz, S. (2003). CXCR4 regulates interneuron migration in the developing neocortex. J. Neurosci. *23*, 5123-5130.

Sultan, K.T., Brown, K.N., and Shi, S.H. (2013). Production and organization of neocortical interneurons. Front. Cell. Neurosci. *7*, 221.

Sultan, K.T., and Shi, S.H. (2018). Generation of diverse cortical inhibitory interneurons. Wiley Interdiscip Rev Dev Biol *7*, e306.

Sun, X., and Kaufman, P.D. (2018). Ki-67: more than a proliferation marker. Chromosoma *127*, 175-186.

Sussel, L., Marin, O., Kimura, S., and Rubenstein, J.L. (1999). Loss of Nkx2.1 homeobox gene function results in a ventral to dorsal molecular respecification within the basal telencephalon: evidence for a transformation of the pallidum into the striatum. Development *126*, 3359-3370.

Szabadics, J., Varga, C., Molnar, G., Olah, S., Barzo, P., and Tamas, G. (2006). Excitatory effect of GABAergic axo-axonic cells in cortical microcircuits. Science *311*, 233-235.

Szabo, N.E., Zhao, T., Cankaya, M., Theil, T., Zhou, X., and Alvarez-Bolado, G. (2009a). Role of neuroepithelial Sonic hedgehog in hypothalamic patterning. J. Neurosci. *29*, 6989-7002.

Szabo, N.E., Zhao, T., Zhou, X., and Alvarez-Bolado, G. (2009b). The role of Sonic hedgehog of neural origin in thalamic differentiation in the mouse. J. Neurosci. *29*, 2453-2466.

Szatkiewicz, J.P., O'Dushlaine, C., Chen, G., Chambert, K., Moran, J.L., Neale, B.M., Fromer, M., Ruderfer, D., Akterin, S., Bergen, S.E., et al. (2014). Copy number variation in schizophrenia in Sweden. Mol. Psychiatry *19*, 762-773.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2008). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Obstetrical & Gynecological Survey *63*, 153.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663-676.

Tallon-Baudry, C., Bertrand, O., Peronnet, F., and Pernier, J. (1998). Induced gamma-band activity during the delay of a visual short-term memory task in humans. J. Neurosci. *18*, 4244-4254.

Tamas, G., Buhl, E.H., Lorincz, A., and Somogyi, P. (2000). Proximally targeted GABAergic synapses and gap junctions synchronize cortical interneurons. Nat. Neurosci. *3*, 366-371.

Tanaka, D., Nakaya, Y., Yanagawa, Y., Obata, K., and Murakami, F. (2003). Multimodal tangential migration of neocortical GABAergic neurons independent of GPI-anchored proteins. Development *130*, 5803-5813.

Tanaka, D.H., Mikami, S., Nagasawa, T., Miyazaki, J., Nakajima, K., and Murakami, F. (2010). CXCR4 is required for proper regional and laminar distribution of cortical somatostatin-, calretinin-, and neuropeptide Y-expressing GABAergic interneurons. Cereb. Cortex *20*, 2810-2817.

Taniguchi, H., Lu, J., and Huang, Z.J. (2013). The spatial and temporal origin of chandelier cells in mouse neocortex. Science *339*, 70-74.

Tarpey, P., Parnau, J., Blow, M., Woffendin, H., Bignell, G., Cox, C., Cox, J., Davies, H., Edkins, S., Holden, S., et al. (2004). Mutations in the DLG3 gene cause nonsyndromic X-linked mental retardation. Am. J. Hum. Genet. *75*, 318-324.

Tasic, B., Menon, V., Nguyen, T.N., Kim, T.K., Jarsky, T., Yao, Z., Levi, B., Gray, L.T., Sorensen, S.A., Dolbeare, T., et al. (2016). Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. Nat. Neurosci. *19*, 335-346.

Tasic, B., Yao, Z., Graybuck, L.T., Smith, K.A., Nguyen, T.N., Bertagnolli, D., Goldy, J., Garren, E., Economo, M.N., Viswanathan, S., et al. (2018). Shared and distinct transcriptomic cell types across neocortical areas. Nature *563*, 72-78.

Tekki-Kessaris, N., Woodruff, R., Hall, A.C., Gaffield, W., Kimura, S., Stiles, C.D., Rowitch, D.H., and Richardson, W.D. (2001). Hedgehog-dependent oligodendrocyte lineage specification in the telencephalon. Development *128*, 2545-2554.

Telley, L., Agirman, G., Prados, J., Amberg, N., Fievre, S., Oberst, P., Bartolini, G., Vitali, I., Cadilhac, C., Hippenmeyer, S., et al. (2019). Temporal patterning of apical progenitors and their daughter neurons in the developing neocortex. Science *364*.

Thévenaz, P. (2011). StackReg - An ImageJ plugin for the recursive alignment of a stack of images. *27-41*.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145-1147.

Tiveron, M.C., Rossel, M., Moepps, B., Zhang, Y.L., Seidenfaden, R., Favor, J., Konig, N., and Cremer, H. (2006). Molecular interaction between projection neuron precursors and invading interneurons via stromal-derived factor 1 (CXCL12)/CXCR4 signaling in the cortical subventricular zone/intermediate zone. J. Neurosci. *26*, 13273-13278.

Tremblay, R., Lee, S., and Rudy, B. (2016). GABAergic Interneurons in the Neocortex: From Cellular Properties to Circuits. Neuron *91*, 260-292.

Tretter, V., Mukherjee, J., Maric, H.M., Schindelin, H., Sieghart, W., and Moss, S.J. (2012). Gephyrin, the enigmatic organizer at GABAergic synapses. Front. Cell. Neurosci. *6*, 23.

Tripodi, M., Filosa, A., Armentano, M., and Studer, M. (2004). The COUP-TF nuclear receptors regulate cell migration in the mammalian basal forebrain. Development *131*, 6119-6129.

Trubetskoy, V., Pardinas, A.F., Qi, T., Panagiotaropoulou, G., Awasthi, S., Bigdeli, T.B., Bryois, J., Chen, C.Y., Dennison, C.A., Hall, L.S., et al. (2022). Mapping genomic loci implicates genes and synaptic biology in schizophrenia. Nature *604*, 502-508.

Uematsu, M., Hirai, Y., Karube, F., Ebihara, S., Kato, M., Abe, K., Obata, K., Yoshida, S., Hirabayashi, M., Yanagawa, Y., et al. (2008). Quantitative chemical composition of cortical GABAergic neurons revealed in transgenic venus-expressing rats. Cereb. Cortex *18*, 315-330.

Ulloa, F., and Marti, E. (2010). Wnt won the war: antagonistic role of Wnt over Shh controls dorso-ventral patterning of the vertebrate neural tube. Dev. Dyn. *239*, 69-76.

Ursini, G., Punzi, G., Chen, Q., Marenco, S., Robinson, J.F., Porcelli, A., Hamilton, E.G., Mitjans, M., Maddalena, G., Begemann, M., et al. (2018). Convergence of placenta biology and genetic risk for schizophrenia. Nat. Med. *24*, 792-801.

van de Leemput, J., Boles, N.C., Kiehl, T.R., Corneo, B., Lederman, P., Menon, V., Lee, C., Martinez, R.A., Levi, B.P., Thompson, C.L., et al. (2014). CORTECON: a temporal transcriptome analysis of in vitro human cerebral cortex development from human embryonic stem cells. Neuron *83*, 51-68.

Van der Maaten, L., and Hinton, G. (2008). Visualizing data using t-SNE. *9*, 2579–2605.

Vassos, E., Pedersen, C.B., Murray, R.M., Collier, D.A., and Lewis, C.M. (2012). Meta-analysis of the association of urbanicity with schizophrenia. Schizophr. Bull. *38*, 1118-1123.

Volpato, V., and Webber, C. (2020). Addressing variability in iPSC-derived models of human disease: guidelines to promote reproducibility. Dis Model Mech *13*.

von Engelhardt, J., Khrulev, S., Eliava, M., Wahlster, S., and Monyer, H. (2011). 5-HT(3A) receptor-bearing white matter interstitial GABAergic interneurons are functionally integrated into cortical and subcortical networks. J. Neurosci. *31*, 16844-16854.

Vullhorst, D., Neddens, J., Karavanova, I., Tricoire, L., Petralia, R.S., McBain, C.J., and Buonanno, A. (2009). Selective expression of ErbB4 in interneurons, but not pyramidal cells, of the rodent hippocampus. J. Neurosci. *29*, 12255-12264.

Waclaw, R.R., Allen, Z.J., 2nd, Bell, S.M., Erdelyi, F., Szabo, G., Potter, S.S., and Campbell, K. (2006). The zinc finger transcription factor Sp8 regulates the generation and diversity of olfactory bulb interneurons. Neuron *49*, 503-516.

Waclaw, R.R., Wang, B., Pei, Z., Ehrman, L.A., and Campbell, K. (2009). Distinct temporal requirements for the homeobox gene Gsx2 in specifying striatal and olfactory bulb neuronal fates. Neuron *63*, 451-465.

Walker, R.L., Ramaswami, G., Hartl, C., Mancuso, N., Gandal, M.J., de la Torre-Ubieta, L., Pasaniuc, B., Stein, J.L., and Geschwind, D.H. (2019). Genetic Control of Expression and Splicing in Developing Human Brain Informs Disease Mechanisms. Cell *179*, 750-771 e722.

Wang, X., He, Y., Zhang, Q., Ren, X., and Zhang, Z. (2021). Direct Comparative Analyses of 10X Genomics Chromium and Smart-seq2. Genomics Proteomics Bioinformatics *19*, 253-266.

Wang, Y., Li, G., Stanco, A., Long, J.E., Crawford, D., Potter, G.B., Pleasure, S.J., Behrens, T., and Rubenstein, J.L. (2011). CXCR4 and CXCR7 have distinct functions in regulating interneuron migration. Neuron *69*, 61-76.

Wang, Y., Toledo-Rodriguez, M., Gupta, A., Wu, C., Silberberg, G., Luo, J., and Markram, H. (2004). Anatomical, physiological and molecular properties of Martinotti cells in the somatosensory cortex of the juvenile rat. J. Physiol. *561*, 65-90.

Warland, A., Kendall, K.M., Rees, E., Kirov, G., and Caseras, X. (2020). Schizophrenia-associated genomic copy number variants and subcortical brain volumes in the UK Biobank. Mol. Psychiatry *25*, 854-862.

Whitfield, M.L., Sherlock, G., Saldanha, A.J., Murray, J.I., Ball, C.A., Alexander, K.E., Matese, J.C., Perou, C.M., Hurt, M.M., Brown, P.O., et al. (2002). Identification of genes periodically expressed in the human cell cycle and their expression in tumors. Mol. Biol. Cell *13*, 1977-2000.

Wianny, F., Real, F., Mummery, C., Van Rooijen, M., Lahti, J., Samarut, J., and Savatier, P. (1998). G1-phase regulators, cyclin D1, cyclin D2, and cyclin D3: Up-regulation at gastrulation and dynamic expression during neurulation. Dev. Dyn: an official publication of the American Association of Anatomists *212*, 49-62.

Wichterle, H., Garcia-Verdugo, J.M., Herrera, D.G., and Alvarez-Buylla, A. (1999). Young neurons from medial ganglionic eminence disperse in adult and embryonic brain. Nat. Neurosci. *2*, 461-466.

Wichterle, H., Turnbull, D.H., Nery, S., Fishell, G., and Alvarez-Buylla, A. (2001). In utero fate mapping reveals distinct migratory pathways and fates of neurons born in the mammalian basal forebrain. Development *128*, 3759-3771.

Wilkins, O.G., Capitanchik, C., Luscombe, N.M., and Ule, J. (2021). Ultraplex: A rapid, flexible, all-in-one fastq demultiplexer. Wellcome Open Res. *6*, 141.

Winkler, D., Daher, F., Wustefeld, L., Hammerschmidt, K., Poggi, G., Seelbach, A., Krueger-Burg, D., Vafadari, B., Ronnenberg, A., Liu, Y., et al. (2018). Hypersocial behavior and biological redundancy in mice with reduced expression of PSD95 or PSD93. Behav. Brain Res. *352*, 35-45.

Won, S., Levy, J.M., Nicoll, R.A., and Roche, K.W. (2017). MAGUKs: multifaceted synaptic organizers. Curr. Opin. Neurobiol. *43*, 94-101.

Wonders, C.P., and Anderson, S.A. (2006). The origin and specification of cortical interneurons. Nat. Rev. Neurosci. *7*, 687-696.

Woodruff, A.R., McGarry, L.M., Vogels, T.P., Inan, M., Anderson, S.A., and Yuste, R. (2011). State-dependent function of neocortical chandelier cells. J. Neurosci. *31*, 17872-17886.

Woods, D.F., Hough, C., Peel, D., Callaini, G., and Bryant, P.J. (1996). Dlg protein is required for junction structure, cell polarity, and proliferation control in Drosophila epithelia. J. Cell Biol. *134*, 1469-1482.

Xu, Q., Guo, L., Moore, H., Waclaw, R.R., Campbell, K., and Anderson, S.A. (2010a). Sonic hedgehog signaling confers ventral telencephalic progenitors with distinct cortical interneuron fates. Neuron *65*, 328-340.

Xu, Q., Tam, M., and Anderson, S.A. (2008). Fate mapping Nkx2.1-lineage cells in the mouse telencephalon. J. Comp. Neurol. *506*, 16-29.

Xu, X., Roby, K.D., and Callaway, E.M. (2010b). Immunochemical characterization of inhibitory mouse cortical neurons: three chemically distinct classes of inhibitory cells. J. Comp. Neurol. *518*, 389-404.

Yan, X.X., Zheng, D.S., and Garey, L.J. (1992). Prenatal development of GABA-immunoreactive neurons in the human striate cortex. Brain Res. Dev. Brain Res. *65*, 191-204.

Yco, L.P., Geerts, D., Koster, J., and Bachmann, A.S. (2013). PRAF2 stimulates cell proliferation and migration and predicts poor prognosis in neuroblastoma. Int. J. Oncol. *42*, 1408-1416.

Yoo, T., Kim, S.G., Yang, S.H., Kim, H., Kim, E., and Kim, S.Y. (2020). A DLG2 deficiency in mice leads to reduced sociability and increased repetitive behavior accompanied by aberrant synaptic transmission in the dorsal striatum. Mol. Autism *11*, 19.

Yoon, J.H., Maddock, R.J., Rokem, A., Silver, M.A., Minzenberg, M.J., Ragland, J.D., and Carter, C.S. (2010). GABA concentration is reduced in visual cortex in schizophrenia and correlates with orientation-specific surround suppression. J. Neurosci. *30*, 3777-3781.

You, Z., Komamura, Y., and Ishimi, Y. (1999). Biochemical analysis of the intrinsic Mcm4-Mcm6-mcm7 DNA helicase activity. Mol. Cell. Biol. *19*, 8003-8015.

Yuan, F., Fang, K.H., Cao, S.Y., Qu, Z.Y., Li, Q., Krencik, R., Xu, M., Bhattacharyya, A., Su, Y.W., Zhu, D.Y., et al. (2015). Efficient generation of region-specific forebrain neurons from human pluripotent stem cells under highly defined condition. Sci. Rep. *5*, 18550.

Yun, K., Garel, S., Fischman, S., and Rubenstein, J.L. (2003). Patterning of the lateral ganglionic eminence by the Gsh1 and Gsh2 homeobox genes regulates striatal and olfactory bulb histogenesis and the growth of axons through the basal ganglia. J. Comp. Neurol. *461*, 151-165.

Zanni, G., van Esch, H., Bensalem, A., Saillour, Y., Poirier, K., Castelnau, L., Ropers, H.H., de Brouwer, A.P., Laumonnier, F., Fryns, J.P., et al. (2010). A novel mutation in the DLG3 gene encoding the synapse-associated protein 102 (SAP102) causes non-syndromic mental retardation. Neurogenetics *11*, 251-255.

Zecevic, N., Hu, F., and Jakovcevski, I. (2011). Interneurons in the developing human neocortex. Dev. Neurobiol. *71*, 18-33.

Zecevic, N., and Milosevic, A. (1997). Initial development of gamma-aminobutyric acid immunoreactivity in the human cerebral cortex. J. Comp. Neurol. *380*, 495-506.

Zhang, M., Xu, J.T., Zhu, X., Wang, Z., Zhao, X., Hua, Z., Tao, Y.X., and Xu, Y. (2010). Postsynaptic density-93 deficiency protects cultured cortical neurons from N-methyl-D-aspartate receptor-triggered neurotoxicity. Neuroscience *166*, 1083-1090.

Zhang, S., and Cui, W. (2014). Sox2, a key factor in the regulation of pluripotency and neural differentiation. World J. Stem Cells *6*, 305-311.

Zhao, Y., Flandin, P., Long, J.E., Cuesta, M.D., Westphal, H., and Rubenstein, J.L. (2008). Distinct molecular pathways for development of telencephalic interneuron subtypes revealed through analysis of Lhx6 mutants. J. Comp. Neurol. *510*, 79-99.

Zhao, Y., Marin, O., Hermesz, E., Powell, A., Flames, N., Palkovits, M., Rubenstein, J.L., and Westphal, H. (2003). The LIM-homeobox gene Lhx8 is required for the development of many cholinergic neurons in the mouse forebrain. Proc. Natl. Acad. Sci. U. S. A. *100*, 9005-9010.

Zheng, C.Y., Seabold, G.K., Horak, M., and Petralia, R.S. (2011). MAGUKs, synaptic development, and synaptic plasticity. Neuroscientist *17*, 493-512.

Zhong, S., Zhang, S., Fan, X., Wu, Q., Yan, L., Dong, J., Zhang, H., Li, L., Sun, L., Pan, N., et al. (2018). A single-cell RNA-seq survey of the developmental landscape of the human prefrontal cortex. Nature *555*, 524-528.

Zhu, J., Shang, Y., and Zhang, M. (2016). Mechanistic basis of MAGUK-organized complexes in synaptic development and signalling. Nat. Rev. Neurosci. *17*, 209-223.

8. Supplementary Information

Supplementary Table 1. The list of DEGs from Nowakowski dataset (up and down-regulated) between each specific cluster and the others (related to section 5.4).

Supplementary Table 2. The list of DEGs (up and down-regulated) from WT cells comparing the cells in each specific cluster versus all other cells. Also, full list of gene ontology classes enriched for up- and down-regulated protein coding genes significantly differentially expressed using Fisher's exact test for GO term functional enrichment (Bonferroni $p < 0.05$) in WT clusters (unrefined and refined) (related to section 5.4).

Supplementary Table 3. The list of DEGs (up and down-regulated) obtained comparing KO cells versus WT cells in each specific cluster (related to section 5.5).

Supplementary Table 4. The list of gene ontology classes enriched for up- and down-regulated protein coding genes significantly differentially expressed between KO and WT using Fisher's exact test for GO term functional enrichment (Bonferroni $p < 0.05$) in each specific cluster (unrefined and refined) (related to section 5.5).

Supplementary Table 5. The list of DEGs (up and down-regulated) obtained comparing KO vs WT on the identified cell type groups (eight higher groups), comparing the cells in each group versus all other cells (related to section 5.5).

Supplementary Table 6. The list of gene ontology classes enriched for up- and down-regulated protein coding genes significantly differentially expressed in eight higher groups, using Fisher's exact test for GO term functional enrichment (Bonferroni $p < 0.05$) (unrefined and refined) (related to section 5.5).

Supplementary Table 7. The list of up and down regulated proteins that are recognized in mass spectrometry. WT (JSD4), KO1 (JSD2) and KO2 (JSD21). Student t-test q values (FDR <0.05) was used to select significant differences. And then among those below 0.05, student T-test differences (mean difference) was used to show what is increased (red is increased in either JSD2 or JSD21) or decreased (green is decreased in either JSD2 or JSD21) in KOs (related to section 3.3).

Supplementary file 1. The functional gene-set enrichment tests R script: Fisher’s exact test and refinement by odds ratio. Generously offered by Dr Andrew Pocklington (related to sections 5.4 & 5.5).

Supplementary file 2. The script used to perform the pipeline to analyse the single cell data and identify Differentially Expressed Genes among cell types in WT samples.

Supplementary file 3. The script used to perform the pipeline to analyse the single cell data and identify Differentially Expressed Genes between WT and KO cell types.

Supplementary File 4. The list of genes in CNVs. The highlighted genes (yellow) are involved in synaptogenesis or involved in neurodevelopment (related to section 3.4).