



UNIVERSITY OF
PORTSMOUTH

**University of Portsmouth
School of Biological Sciences**

BSc (Hons) Biochemistry

**Transcriptional Analysis of ASCL1 and the
Effects of Neurogenesis in GABAergic Neurons**

Student Number: UP889480

2020/2021 Academic Year

Supervisor: Dr Frank Schubert

Table of Content

1. Abstract	3
2. Introduction	4
3. Aims	9
4. Methodology	10
A. Database identification and storage	10
B. The Gene Expression Omnibus database	11
C. Galaxy	12
D. RNA Seq processing and quality	12
E. Mapping and alignment	12
F. Quantification of the counts	12
G. Differential analysis of the genes	12
H. Volcano plot generation	12
I. Filtering and sorting the gene lists	12
J. Comparing of gene lists between data sets	13
K. GO Term classification and analysis	13
L. STRING analysis	14
5. Results	15
A. Visualisation of the volcano plots	15
B. Principal component analysis (PCA) plot and uniform manifold approximation and projection (UMAP) plot	16
C. Venn diagram comparative analysis	18
D. GO Term analysis	19
E. STRING PPI network analysis	20
6. Discussion	23
A. ASCL1 and the relationship with NOTCH2 and its role in expression promoting GABAergic proliferation and differentiation	23
B. The relationship between PITX2 and GATA2 in GABAergic differentiation	24
C. ASCL1 expression is important for synaptic communication of GABA	24

D. ASCL1 and SLIT2 conflicted role in cell mobility and neuronal modulation	26
E. Limitations of experiment design	26
F. Future research	27
7. Conclusion	28
8. Reference	39
9. Acknowledgements	47
10. Appendix	48
I. The list of references genomes used the experiment	48
II. The datasets used the experiment	48
III. Table used for GO enrichment analysis for molecular function	49
IV. Table used for GO enrichment analysis for biological process	50
V. Table used for GO enrichment analysis for cellular component	51
VI. Table of the gene overlaps between datasets: GSE29985, GSE31635 and GSE46791	52
VII. The processed excel gene datasets	53
11. Total number of words	54
12. Signed Declaration	55

1. Abstract

Neurogenesis refers to the process by which undifferentiated neural progenitor stem cells generate mature and functional neurons. One of the ways that neurogenesis is controlled is by ASCL1 which has roles in the central nervous system and areas of the peripheral nervous system. ASCL1 is responsible for the differentiation of cortical interneurons which are derived from the ventral telencephalon from the MGE and CGE regions and are associated with the proliferation and differentiation of immature neurons. Mutations in ASCL1 are associated with epileptic seizures and schizophrenia with strabismus. What is not known is the relationship between ASCL1 and its target genes and how they play in the development of GABAergic neurons specifically. Using the Gene Expression Omnibus database and ArrayExpress, 5 datasets were obtained which contain RNA-Seq and microarray data which are processed using Galaxy and GEO2R to obtain gene lists containing both upregulated and downregulated genes. To observe which gene overlap between the datasets and find similar biological meaning between the datasets. The datasets were placed in a Venn diagram using InteractiVenn. The gene overlap is processed in g:profiler and STRING to identify several proteins linked to ASCL1 and GABAergic neurogenesis: GAD2 and CALB1/2 are associated with synaptic communication. GATA2 is associated with GABAergic differentiation, NOTCH2 inhibition promotes neural genes related to neurogenesis and PITX2's role appears to be confined in neurogenesis in the midbrain. Slit2 and its role remains unclear in the process of neurogenesis and may play a minimal role in neural differentiation and maintaining axonal morphology.

Keywords: ASCL1, Neurogenesis, GABAergic Neurons

2. Introduction

Neurogenesis refers to the process by which undifferentiated neural progenitor stem cells generate mature and functional neurons (Begega, Alvarez-Suarez, Sampedro-Piquero & Cuesta, 2017). The process of neurogenesis occurs within the two types of neural stem cells (NSCs) that are found in the body: neuroepithelial cells and radial glial cells (Götz & Huttner, 2005). Neuroepithelial cells are NSCs and can differentiate into neuron cells and radial glial cells (Hall, 2012) and have a pseudostratified appearance as a result of the nuclei of the stem cells move up and down along the apical-basal axis during mitosis (Götz & Huttner, 2005). Radial glial cells derived from the neuroepithelium can undergo differentiation into neurons, astrocytes and oligodendrocytes and cell mitosis and are used to maintain the structural integrity of the nervous system (Hall & Miller, 2012).

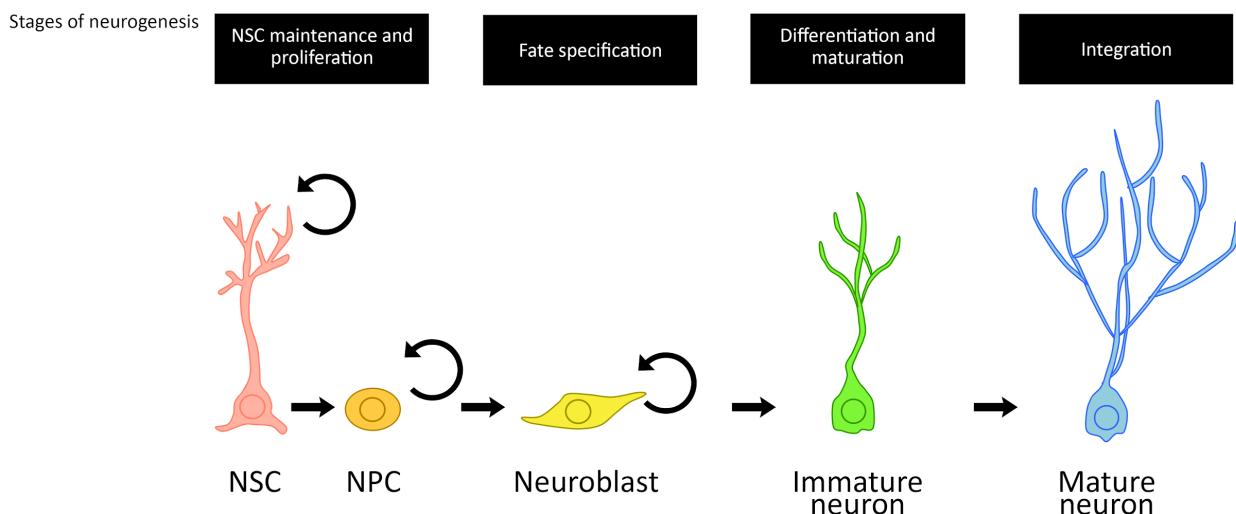


Figure 1: Adult neurogenesis is made of several different cell types: NSCs (pink), NPCs (orange), neuroblasts (yellow), immature neurons (green), and mature neurons (blue). The circular arrow represents that cell being able to undergo self-renewal. NPCs and neuroblasts although similar both can be differentiated based on protein markers associated with the different cell types. NPCs express glial fibrillary acid protein (GFAP), SOX2 brain lipid-binding protein (BLBP), and nestin and when NPCs proliferate they are characterized by the expression of SOX2 and PAX6 only. Neuroblast expresses PSA-NCAM and doublecortin (DCX). Image was drawn using information from "Boldrini et al., 2018; Hsieh, 2012; Liu & Song, 2016; "Markers of neuronal precursor cells (NPCs) / neural stem cells (NSCs)", n.d.; Niklison-Chirou, Agostini, Amelio & Melino, 2020; van Strien, van den Berge & Hol, 2011".

Neurogenesis is divided into four stages: the maintenance and proliferation of stem cells, fate determination of NSCs, differentiation and maturation and integration into the central nervous system (Liu & Song, 2016). Neuroepithelial cells and radial glia both undergo symmetric and asymmetric division. Symmetric division refers to neurogenesis describing the process where a TACs gives rise to two additional two transient amplifying cells or two differentiated cells (Götz & Huttner, 2005; Shahriyari & Komarova, 2013; Slack, 2018). Asymmetric division refers to the process where the stem cell gives rise to one stem cell and one stem cell which can commit to a TAC which further commits to a differentiated cell (Götz & Huttner, 2005; Scott, Hjelmeland, Chinnaiyan, Anderson & Basanta, 2014; Shahriyari & Komarova, 2013).

Symmetric and asymmetric division of NSCs is mainly determined by the distribution of cytoplasmic, centrosomal and polarity proteins which affects the spindle orientation within a cell; the symmetric division of cells is achieved by the equal distribution of polarity proteins into two daughter cells and asymmetric division achieved by the uneven division of polarity proteins into one daughter cell (Paridaen & Huttner, 2014). The distribution of the polarity proteins determines where the cell is divided. Vertical cleavage of a

neural stem cell results in the symmetric division producing two identical daughter cells. Horizontal cleavage of a neural stem cell results in asymmetric neurogenic division, forming a neural progenitor cell (NPC) and a neuron cell. The oblique cleavage of neural stem cells results in asymmetric neurogenic division, similarly to horizontal cleavage of a neural stem cell (Götz & Huttner, 2005). However experimental data shows that the oblique cleavage in NSCs is inconsistent in the determination of what type of cell division the mother cells take, irrespective of the angle of the cleavage that occurred in NSCs (Noctor, Martínez-Cerdeño & Kriegstein, 2008), which suggests that more research is needed to be conclusive.

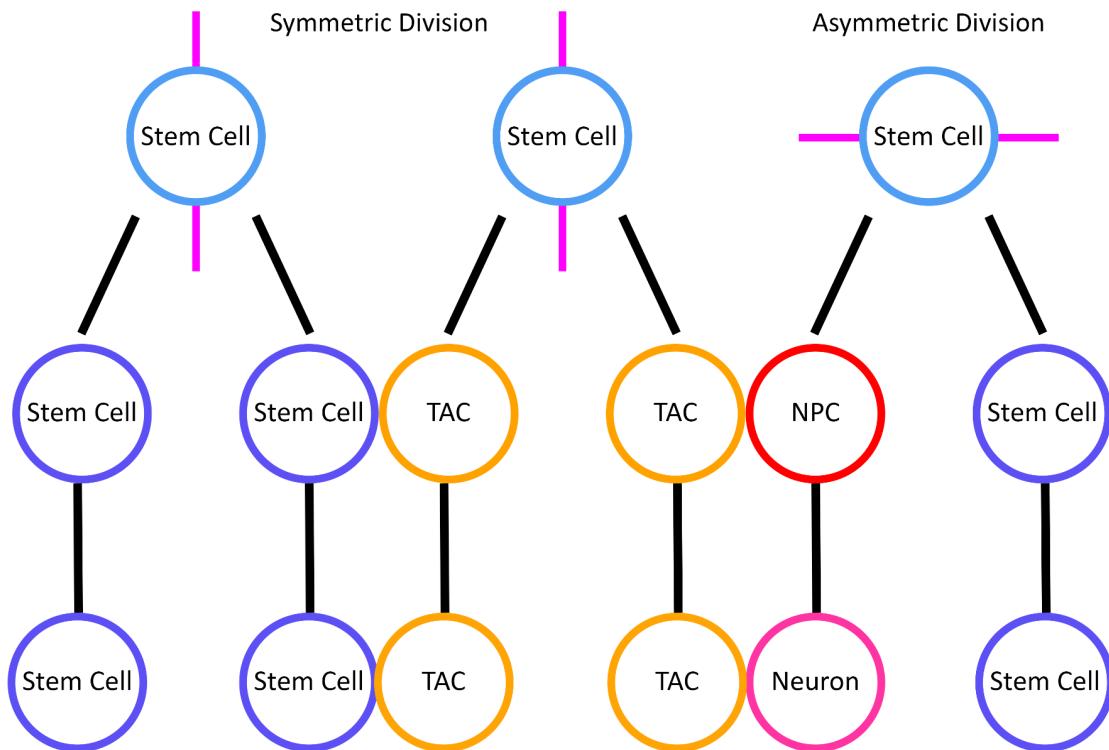


Figure 2: Stem cells (light blue) which can remain in the stem cell pool niche and divide theoretically indefinitely undergo symmetric division and undergo vertical cleavage which produce two identical daughter cells; either two stem cells (dark blue) or two transient amplifying cells which have a limited number of cell divisions before they become differentiated cells (orange) in the later stages of neuronal development. Stem cells undergo asymmetric differentiation by horizontal cleavage where two different cells are produced; one differentiated cell and one stem cell. In the case of neural stem cells. Asymmetric division produces a neural progenitor cell (NPC) (red) which is a type of transient amplifying cell which becomes further committed into a neuron cell (pink) and a daughter stem cell. Stem cells are located into the VZ and TACs are found in the SVZ in the cerebral cortex. The distribution of the polarity proteins found in apical and basal surface, the centrosome spindles and the cytoplasmic membranes determines where the cell is divided (magenta). Image was drawn using information from "Götz & Huttner, 2005; Lui, Hansen & Kriegstein, 2011; Paridaen & Huttner, 2014; Scott, Hjelmeland, Chinnaiyan, Anderson & Basanta, 2014; Shahriyari & Komarova, 2013".

One of the specific ways that neurogenesis is controlled is by bHLH proteins which are transcription factors (Jones, 2004). The bHLH motif is highly conserved among other proteins, typically made of roughly 60 amino acids (Ghosh, Yao & Chmielewski, 1999) which form two amphipathic helices connected by a linker region that forms a loop. This is referred to as a helix-loop-helix (HLH) domain (Ghosh, Yao & Chmielewski, 1999; "Helix-loop-helix DNA-binding domain superfamily", n.d.). The 15 out of the roughly 60 amino acids found at the amino-terminal that contains basic residues and are found next to the HLH domain

("Helix-loop-helix DNA-binding domain superfamily", n.d.; Jones, 2004). They bind to a core hexanucleotide sequence, "CANNTG", called the E-box ("Helix-loop-helix DNA-binding domain superfamily", n.d.; Jones, 2004; Massari & Murre, 2000) which are found in the promoter and enhancer elements of DNA (Massari & Murre, 2000; Schwab, 2009). The basic role of bHLH proteins is to influence the fate of progenitor cells; they are largely responsible for the differentiation of undifferentiated progenitor neural cells into neurons and control the development of GABAergic and glutamatergic neurons in the forebrain and the development of motor neurons from the spinal cord and its progenitors (Guillemot, 2007).

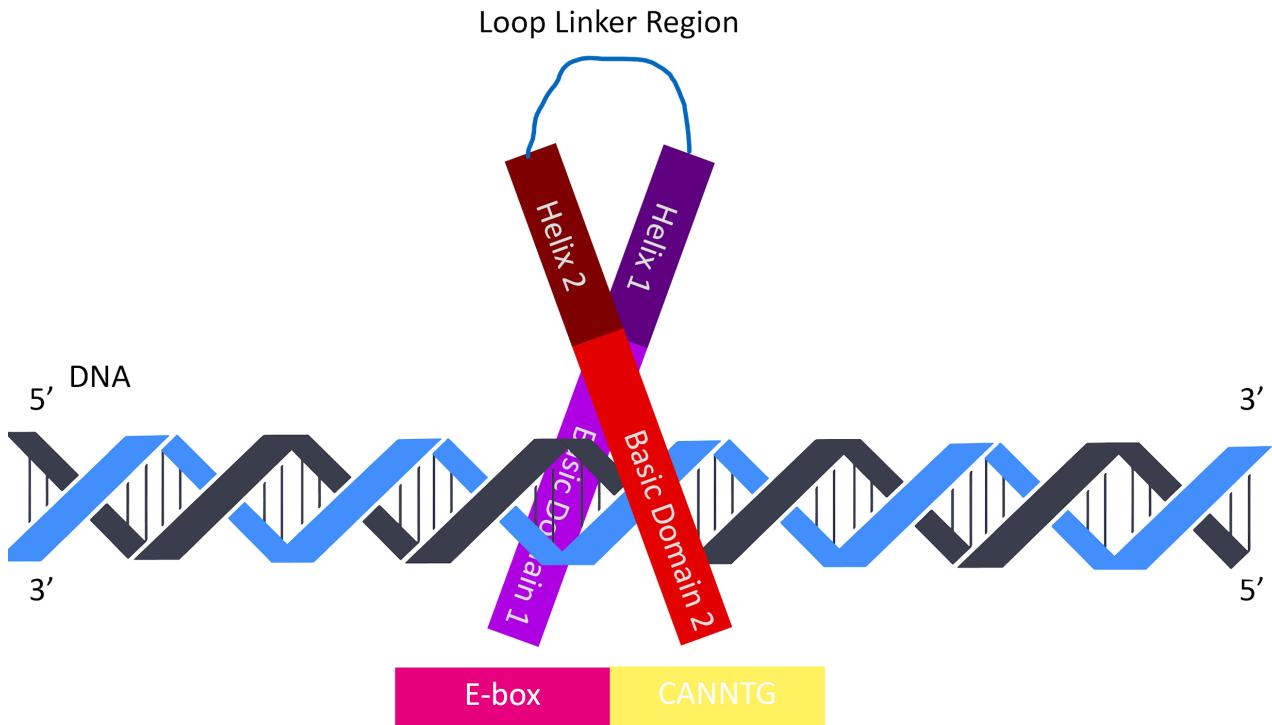


Figure 3: The bHLH structural motif seen in proneural transcription factors. The bHLH domain contains a basic region joined by the two α -helix, then a short linker loop region. The basic domain contains 15 residues and this is where it binds to the DNA via E box motif sequence "CANNTG" at the major groove. Image was drawn using information from "Dennis, Han & Schuurmans, 2019; Helix-loop-helix DNA-binding domain superfamily", n.d.; Sicoli, Vezin, Ledolter, Kress & Kurzbach, 2019; Tapscott, 2005; Jones, 2004".

Neural-specific bHLH genes are a shared homology with *Drosophila* genes (Dennis, Han & Schuurmans, 2019). The *achaete-scute* complex encodes by four proneural genes: *achaete* (ac), *scute* (sc), *lethal of scute* (lsc) and *asense* (ase) (García-Bellido & de Celis, 2009) where it was first discovered in *Drosophila* (Bertrand, Castro & Guillemot, 2002; García-Bellido, 1979). Related to the *ato gene* family. Both families are activated by the Notch pathway and allow the proliferation of neural progenitor cells from the ectoderm (Bertrand, Castro & Guillemot, 2002).

One example of a bHLH protein is ASCL1. Found on chromosome 12 ("CIViC: Gene ASCL1 Summary", n.d.). The protein sequence is made up of 236 amino acids and with a molecular weight of 25.454 kDa ("ASCL1 Gene (Protein Coding)", n.d.). It is the vertebrate homolog to the *achaete-scute* complex in *Drosophila* (Battiste et al., 2007). The expression of ASCL1 is restricted and limited to specific regions along the dorsoventral and rostrocaudal axes of the ventricular zone (VZ) and subventricular zone (SVZ) of the neural tube which is a part of the telencephalon (Battiste et al., 2007; Dennis, Han & Schuurmans, 2019). The difference between the VZ and SVZ is that SVZ ventral telencephalon is maintained throughout adulthood

compared to VZ which disappears at the end of embryo development (Dehay & Kennedy, 2007; Quiñones-Hinojosa & Chaichana, 2007) as well as containing different cell populations with VZ having a population of radial glial cells and intermediate progenitor cells compared to SVZ containing intermediate progenitor cells and daughter neurons (Noctor, Martínez-Cerdeño, Ivic & Kriegstein, 2004).

ASCL1, together with Neurog2, have similar roles involved in the proliferation and differentiation of neurons as both transcription factors are expressed complementary but not interchangeable with each other. As ASCL1 and Neurog2 bind to different genomic sites as ChIP-seq data showed that 90% of ASCL1 and Neurog2 were differentially bound; with 45% being bound to ASCL1 sites and 45% is bound to Neurog2 sites (Aydin et al., 2019). Together with both transcription factors binding of the E box sequence motif, "CANNTG" which commands the transcription of neuronal genes (Aydin et al., 2019; Dennis, Han & Schuurmans; 2019 Jones, 2004; Massari & Murre, 2000) is responsible for the different cell subtypes in development in the central nervous and areas of the peripheral nervous system. Neurog2 is necessary for the development of the glutamatergic phenotype and ASCL1 is necessary for the GABAergic phenotype. ASCL1 is responsible for the differentiation of cerebellar interneurons (Dennis, Han & Schuurmans, 2019). ASCL1 can induce a GABAergic fate on interneurons which are derived from the ventral telencephalon (Liu et al., 2017; Yang et al., 2017) from the medial ganglionic eminence (MGE) and the caudal (CGE) ganglionic eminence regions in the telencephalon (Aydin et al., 2019; Dennis, Han & Schuurmans, 2019; Faux, Rakic, Andrews & Britto, 2012; Touzot, Ruiz-Reig, Vitalis & Studer, 2016).

GABAergic neurons migrate from the CGE and MGE to the neocortex of the brain where they form three classes of GABAergic neurons: parvalbumin (PV) which accounts for 40% of the GABAergic neurons, somatostatin (SST) which accounts for 30% of the GABAergic neurons, and 5HT3a receptor (5HT3aR) which accounts for 30% of the GABAergic neurons (Kelsom & Lu, 2013; Tremblay, Lee & Rudy, 2016). All classes are involved in the release of the neurotransmitter, gamma-aminobutyric acid (GABA) to regulate the firing rate of target excitatory neurons and prevent continuous excitatory firing of neurons (Kelsom & Lu, 2013; Tremblay, Lee & Rudy, 2016). ASCL1 upregulates Dlx2 which is needed for the differentiation of GABA interneurons (Castro et al., 2011; Liu et al., 2017; Poitras, Ghanem, Hatch & Ekker, 2007; Schuurmans et al., 2004; Raina, Mahajani, Bähr & Kügler, 2019; Yang et al., 2017). It is responsible for the tangential migration of GABAergic interneurons from the ventral telencephalon to the dorsal telencephalon along the ventricular zone (VZ), subventricular zone (SVZ) and intermediate zone (IZ) (Anderson, Eisenstat, Shi & Rubenstein, 1997; Anderson, Mione, Yun & Rubenstein, 1999; Liu et al., 2017; Stuhmer, Puelles, Ekker & Rubenstein, 2002).

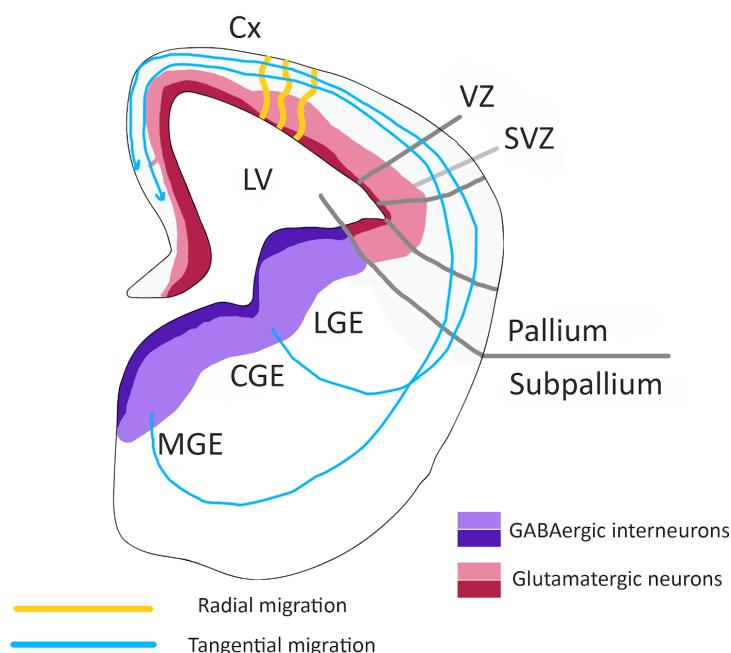


Figure 4: Schematic representation of a hemisection of a mouse embryo forebrain. Glutamatergic projection neurons are produced from the lateral ventricle (LV), the ventricular zone (VZ, dark pink) and subventricular zone (SVZ, light pink). Glutamatergic projection neurons migrate radially (orange lines) to meet the cortex (Cx). The GABAergic interneurons arise from MGE and CGE which make up the ganglionic eminence which is located in the VZ (dark violet) and SVZ (light violet). Image was drawn using information from "Aydin et al., 2019; Brandão & Romcy-Pereira, 2015; Godin & Nguyen, 2013; Hwang, Ku & Hashimoto-Torii, 2019; Moffat, Ka, Jung & Kim, 2015".

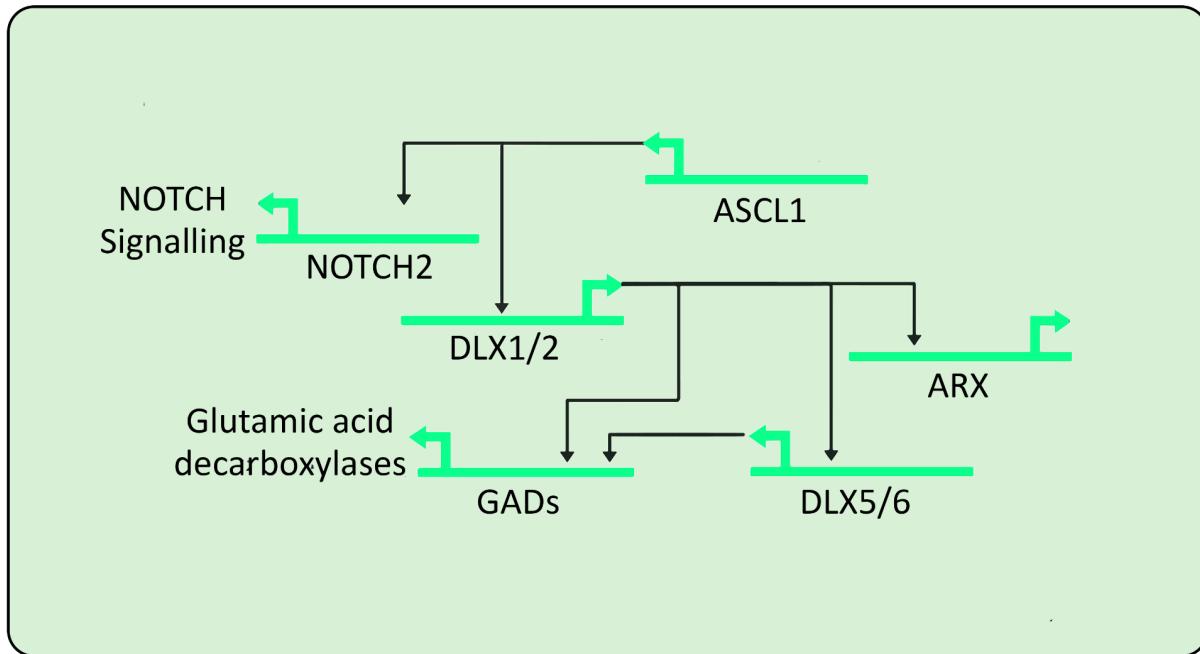


Figure 5: Model of transcription factor network interactions in the developing CGE and MCE. Green arrows indicate gene activation and the black arrows demonstrate the interactions with the network. ASCL1 promotes DLX1/2 expression and NOTCH2 which is involved in the NOTCH signalling which inhibits the expression of DLX1/2 expression. The DLX1/2 genes will regulate DLX5/6 expression and promote ARX expression aids in tangential migration of GABAergic neurons. Both DLX1/2 and DLX5/6 expression promote the expression of the GADs; glutamic acid decarboxylases (GAD2 and GAD1). Image was drawn using information from "Colasante et al., 2008; Erlander, Tillakaratne, Feldblum, Patel & Tobin, 1991; Friocourt et al., 2008; Le et al., 2017; Krueger, Stäcker & Schlosser, 2007; MacDonald et al., 2013; Stühmer, Anderson, Ekker & Rubenstein, 2002; Wang et al., 2013".

Mutations in ASCL1 affect GABAergic neurons associated with neurological diseases (Yang et al., 2017) by reducing the number of GABAergic neurons within the brain (Liu et al., 2017). The common neurological symptoms of ASCL1 mutations in GABAergic neurons in epileptic seizures are GABAergic neurons regulate the excitatory firing from the glutamatergic neurons (Barker-Haliski & White, 2015; Matsui & Hsieh, 2016; Tang, 2017). A similar relationship between the interaction of ASCL1 and PHOX2b occurs in schizophrenia with strabismus and the disruption to GABAergic neurons in the inhibitory circuits (Ide et al., 2005; Marín, 2012) however more research on general schizophrenia is needed to be conclusive to be definitive with the relationship.

Although it is known that ASCL1 is responsible for the tangential migration of GABAergic interneurons to the cortex and mutations to ASCL1 are linked to the reduced GABAergic neurons (Liu et al., 2017) and linked to neurological disorders. What is not known are the ASCL1 and its target genes and how they play the development of GABAergic neurons specifically. Therefore several aims have been proposed to guide the investigation.

3. Aims

- To understand the role of regulation of ASCL1 within GABAergic neurons
- To Identify co-expression groups that occur alongside ASCL1
- To investigate the role of ASCL1 targets genes by gene ontology

These aims will be investigated by collecting the next generation sequences and performing a comparative analysis on existing gene data sets to identify expressions of ASCL1 target genes in different cell clusters achieved by using bioinformatics software. This will result in the 2D graphical representation between the genes and their expression patterns.

4. Methodology

A. Database identification and storage

Datasets were collected, using ArrayExpress, an archive functional genomics database. A keyword search for "ASCL1 Gabaergic" and "ASCL1". E-MTAB-4840 contained various parts of the human brain which can be used as a standard reference, with a large sample size of 632, only 16 samples from the left and right cortex and the left and right telencephalon were selected (Appendix II, Table 1).

B. The Gene Expression Omnibus database

Dataset GSE31635, GSE29985, GSE78949 and GSE46791 were cross-referenced in The Gene Expression Omnibus (GEO) database is a public database that archives high-throughput gene expression datasets (Clough & Barrett, 2016). Dataset GSE31635, GSE29985, GSE78949 and GSE46791 were analysed using GEO2R, a web-based analysis program that compares datasets ("About GEO2R", n.d.; Barrett et al., 2012). Analysis was based on the grouping of the samples using default settings.

Table 1: Details of the datasets used in the present experiment, their biological function.

Database	Experiment type	Organism	Biological function and meaning	Detail of samples
GSE29985	Genome binding/occupancy profiling by genome tiling array	<i>Mus musculus</i>	Observing the ectopic expression of Arx and comparing it to GABAergic neurons.	Using the platform GPL4128, the whole brain sample was used by the control group and the cells transfected with Arx1 were used as the infected group.
GSE31635	Expression profiling by array	<i>Rattus norvegicus</i>	GE6 expresses high levels of Dlx1, 2, 5 and 6, glutamate decarboxylases (GADs), and neuropeptide Y and somatostatin. which are markers of GABAergic neurons. GE6 is compared to other precursors to find differences in the expression patterns in early stages of differentiation.	All GE6 cells (GABAergic neuronal precursor) were compared with L2.2 cells (interneuronal precursor).
E-MTAB-4 840	RNA-seq of coding RNA, observational design	<i>Homo sapiens</i>	The observation and expression of genes associated in human embryonic brain development.	16 samples were chosen from between 4 and 12 post conception weeks. During this time the major brain regions are established and there are the early stages of cortex development. 8 sample from the left and right cortex were selected from 4 sources: 8 samples from the left and right telencephalon from 4 sources: HDBR251, HDBR253, HDBR256 and HDBR259 and identifying the

				upregulated genes in the cortex and downregulated genes in the telencephalon.
GSE78949	Expression profiling by array	<i>Mus musculus</i>	The enrichment of Ascl1, Smad7 and Nr2f1 in ESC populations resulted in the upregulation of GABAergic neuron markers.	All samples containing only ASCL1 were divided based on treatment was doxycycline (dox) which when present is ASCL1 is not misexpressed either and vice versa.
GSE46791	Expression profiling by array	<i>Mus musculus</i>	Observing the difference in gene expression in murine neural stem cells derived from murine embryonic stem cells and their differentiated neuronal progeny which expressed GABAergic phenotype.	Using the platform GPL6885. expression beadchip. The samples were divided into the neuron group which contained the differentiated neuronal progeny cell and the neural stem cell which were grouped into the NSC group.

Table 2: Details of the datasets used in the present experiment and the justification of the used database being selected.

Database	Justification of the databases being selected
GSE29985	ARX and ASCL1 are associated with Dlx genes. ASCL1 expression promotes Dlx2 and ARX is a direct target of Dlx2 and aids in tangential migration of GABAergic neurons (Friocourt et al., 2008). ASCL1 specifically which are needed for GABAergic fate in ventral telencephalon (Colasante et al., 2008).
GSE31635	Dlx genes are associated with GABAergic neurons. ASCL1 binds to Dlx1/2 which promotes Dlx5/6 which causes the expression of GADs. Thus, it should be able to identify GABeric related genes linked to ASCL1.
E-MTAB-4840	The upregulated genes in the telencephalon vs cortex comparison should correspond to the ventral telencephalon which is the origin of the gabaergic neurons; this database should therefore contain Ascl1 itself.
GSE78949	The over-expression of transcription factors in Ascl1, Smad7, Nr2f1, and in ESCs which resulted in a population of GABA-like neurons cells with expression profiles similar to neural tissues and highly expressed markers of GABAergic neurons which can be used to observe in expression profile in ESCs and GABA-like cells.
GSE46791	To identify which genes expressed in murine neural stem cells derived and differentiate GABAergic like cell phenotype.

C. Galaxy

All of the raw RNA-seq data was conducted in the web-based application Galaxy. The advantage of using Galaxy is that it runs as an open-source platform. All of the computation and the storage of the data is performed on servers In addition and requires no knowledge on how to program a computer language (Goecks, Nekrutenko, Taylor & The Galaxy Team, 2010).

D. RNA Seq processing and quality control

RNA Seq data for only E-MTAB-4840 samples are checked for quality using FASTQC (Andrews, n.d.). The low-quality bases are removed to improve mappability using trimmomatic (Bolger, Lohse & Usadel, 2014; Conesa et al., 2016; Ghodsinia, 2020; "RNA-seq expression analysis hands-on tutorial: From FASTQ to differentially expressed genes", n.d.).

E. Mapping and alignment

HISAT2 (Doyle, Phipson & Dashnow, 2020; Kim, Langmead & Salzberg, 2015), is used to align reads of various lengths produced from RAW RNASeq data to the reference genome GRCh38.p13 (Alzu'bi & Clowry, 2019; Ghodsinia, 2020; Kim et al., 2013; "Read mapping or alignment | RNA sequencing", n.d.) using default settings.

F. Quantification of the counts

The aligned read counts are converted into transcripts measured in RPKM (Reads Per Kilobase of transcript, per Million mapped reads) which is a unit of transcript expression (Conesa et al., 2016; Ghodsinia, 2020; "Quantification | RNA sequencing", n.d.). The initial stringtie step using default settings and use the imported reference genome files with the no reference transcripted selected. The stringtie merge step is performed using the stringtie output files as input and use default settings using the important reference genome. The final round of stringtie uses the BAM files as input and the stringtie merge file as a reference and selects the use of reference transcript and output the file as DESeq2/edgeR/limma-voom for differential expression with the average read length obtained from the FASTQC.

G. Differential analysis of the genes

The differential analysis of the Genes used DESeq2 which analysed the counts files produced in the step above, to test for differential gene expression. The DESeq2 results are annotated using "annotate DESeq/DEXSeq".

H. Volcano plot generation

Dataset GSE31635, GSE29985, GSE78949 and GSE46791 which are analysed using GEO2R, are natively created. Datasets analysed using Galaxy, import the "annotate DESeq2/DEXSeq" output tables and follow the methodology previously outlined (Doyle, 2021).

I. Filtering and sorting the gene lists

In Microsoft Excel, all of the datasets which contain genes that are blank, have "NA" were filtered out from the table. Two new sheets are created where the gene name, log2 fold change and a P-adjusted value are imported. Upregulated genes are filtered using the "greater than 0" function on the log2 fold change and downregulated genes are filtered using the "less than 0" function on the log2 fold change. Both sheets are sorted using the criteria of less than or equal to 0.05 for the P-adjacent value. The upstream and downstream genes in each gene list are merged to identify the ASCL1 target within GABAergic neuron development or those GABAergic neuron expressed genes that are targets of ASCL1.

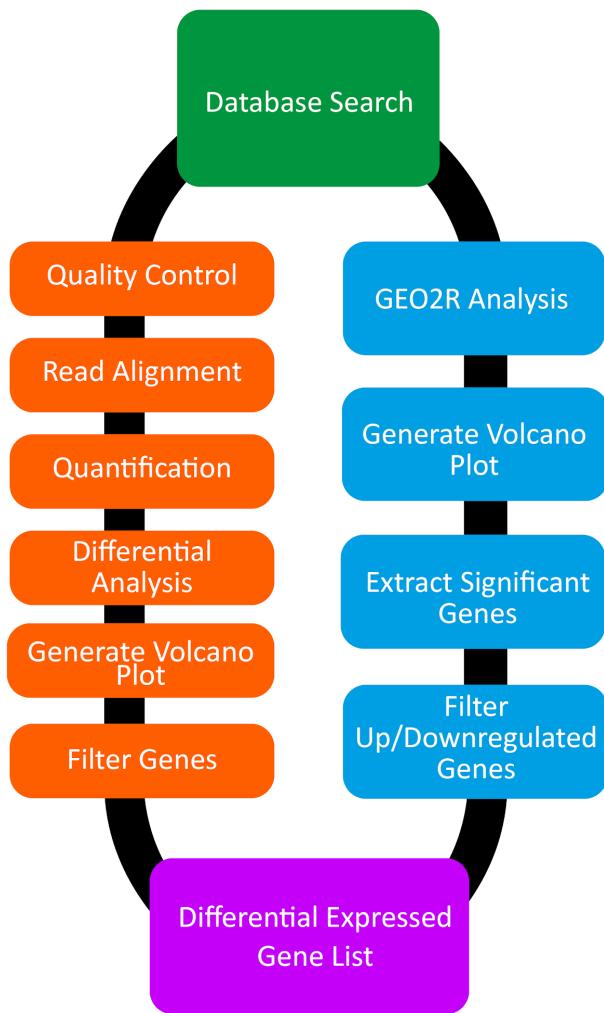


Figure 6: Flow diagram of processing the RNASeq and microarray datasets into differential expressed gene lists. The orange boxes represent the methodology for RNA Seq processing and the blue boxes represent the methodology for genome binding/occupancy profiling by genome tiling array and expression profiling by array.

J. Comparing of gene lists between data sets

All the merged genes in each dataset were compared in InteractiVenn (Heberle, Meirelles, da Silva, Telles & Minghim, 2015). A Venn diagram with the associated overlapping gene lists from the merge gene list. The datasets with the most gene overlaps are exported in Microsoft Excel (Appendix VII).

K. GO Term classification and analysis

The overlapping gene lists from InteractiVenn (Appendix VII) are used in g:Profiler, a free web platform for identifying and analysing gene lists. Using the g:GOSt function, statistical GO Term enrichment analysis can be performed. The reference genome was set as *homo sapiens* enabled ordered query and g:SCS algorithm was enabled and the GO molecular function, GO cellular component and GO biological process was enabled

to identify genes unique in that data (Ashburner et al., 2000; Raudvere et al., 2019). The CSV files are imported into Microsoft Excel and bar charts are created by highlighting -Log10 value and the GO term from each category.

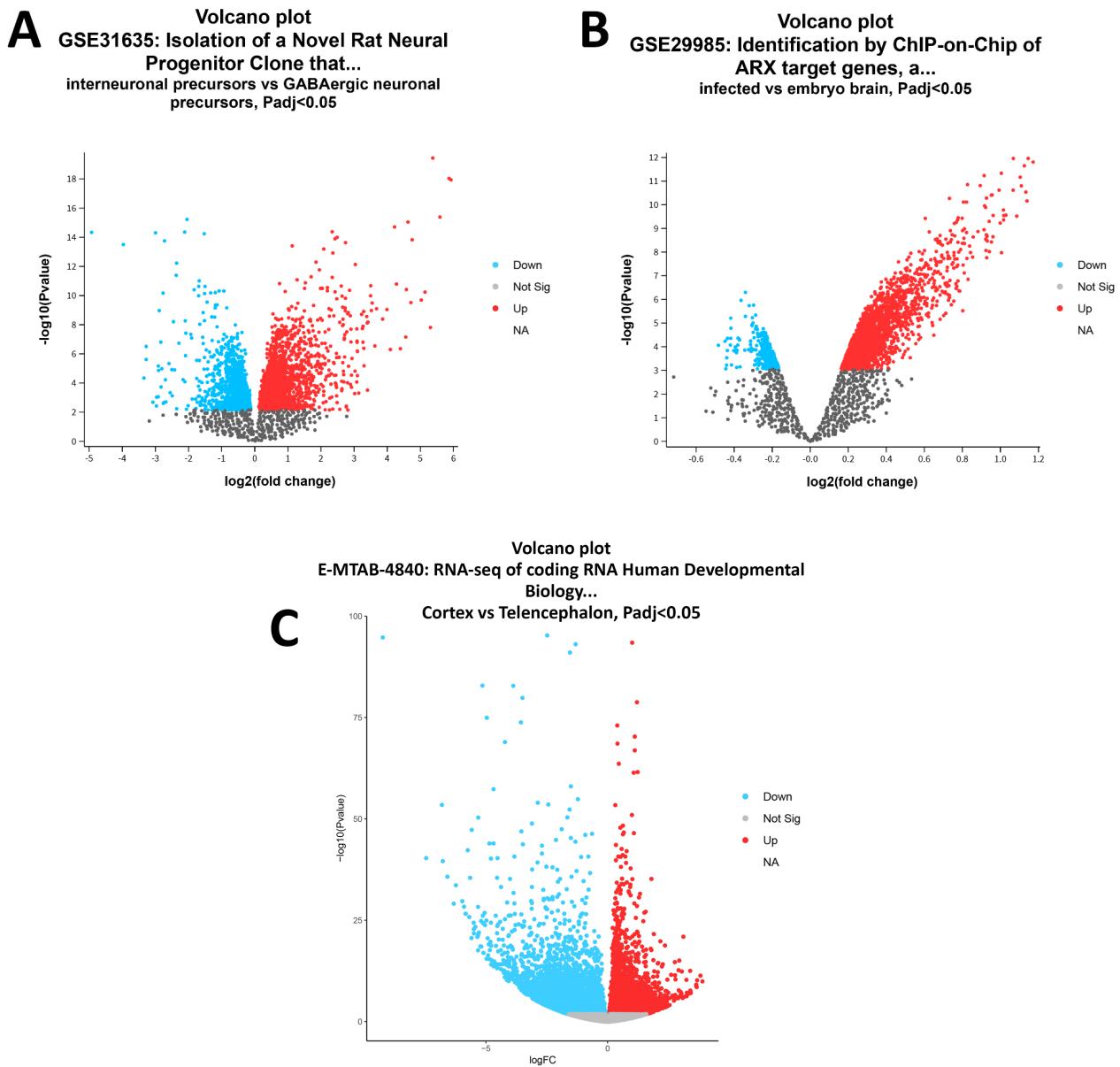
L. STRING Analysis

The same overlapping gene list (Appendix VII) generated from InteractiVenn are placed into STRING, which is a web-based database which can form a visual network between various proteins using various criteria including text mining of papers, databases and the use of computer algorithms (Szklarczyk et al., 2020). The confidence score was set at 0.65 as a compromise between the medium confidence value (0.4) and the high confidence value (0.7) The unconnected proteins were hidden in the network.

5. Results

A. Visualisation of the volcano plots

To investigate for which genes were upregulated and downregulated visualised in the datasets. GEO2R and Galaxy were used to create the volcano plots. Only Figure 7C used Galaxy to generate the volcano plot compared to Figure 7A-B and Figure 7D-E used GEO2R. Both techniques showed the upregulated genes as red and the downregulated genes as blue. Figure 7B showed the smallest proportion of downregulated genes whilst Figure 7A,E showed the equal proportion of upregulated and downregulated genes. Figure 7C displays a greater proportion of downregulated genes and Figure 7D displays the few regulated genes in the volcano plot. Demonstrating that volcano visualisation appears to be inconsistent.



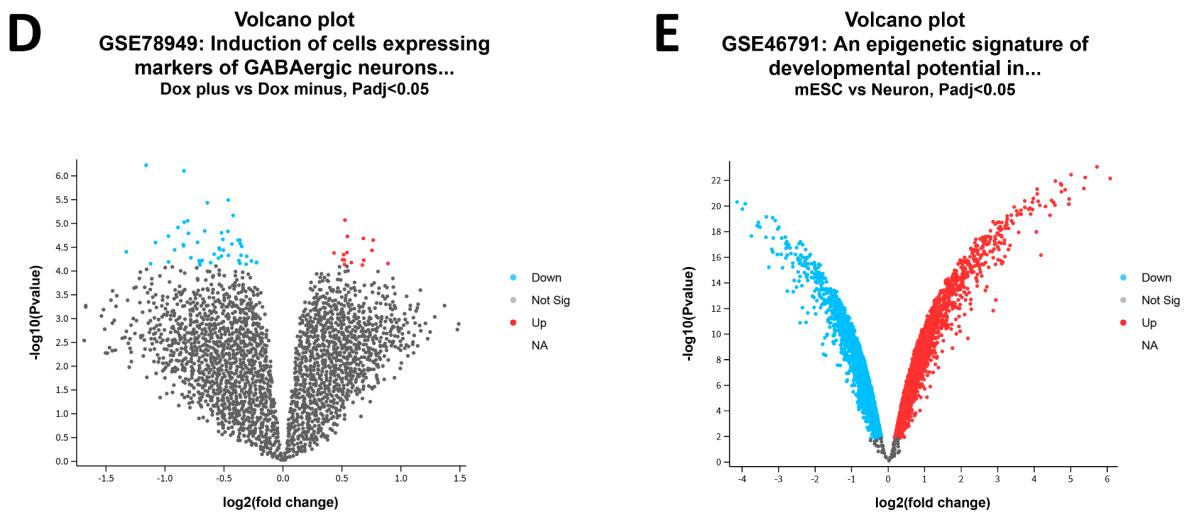


Figure 7: Volcano plots of datasets. Generated from GEO2R and Galaxy. **A**, GSE31635 compared interneuronal precursors to GABAergic neuronal precursors . **B**, GSE29985 compared Infected group to the embryo brain group. **C**, E-MTAB-4840 compared the cortex group to the telencephalon group. **D**, GSE78949 compared samples which contained Dox+ and Dox- treatment. **E**, GSE46791 compared the mouse embryonic stem cell (ESC) group to the neuron group. The volcano plot displays statistical significance ($-\log_{10} P$ value) against the magnitude of change (\log_2 fold change) visualising the differentially expressed genes using a P adjacent value of 0.05

B. Principal component analysis (PCA) plot and uniform manifold approximation and projection (UMAP) plot

To identify which samples are similar to each other in each database. The PCA plots were generated in Galaxy (Figure 8C). The UMAP plots were generated in GEO2R (Figure 8A-B, Figure 7D-E).

Figure 8E showed the highest clustering density followed by Figure 8C which displays three groups where the cortex samples cluster together, however, telencephalon samples are split into two groups. Samples telencephalon 1 and 2 left and telencephalon 5 and 6 right and telencephalon 7 and 8 right and telencephalon 3 and 4. Figure 8A is most anomalous as two groups are divided into interneuronal and GABAergic precursors but the third group includes samples from both categories. Figure 8B displays the high dispersion within the plot. Figure 8D has the weakest clustering and highest dispersion despite the samples being divided into Dox+ and Dox- samples which should provide distinctive clusters. What this means is that samples, with the exception of Figure 8E, the remaining datasets vary in their clustering and their dispersion to some degree regardless of grouping, with the most unusual in Figure 8D.

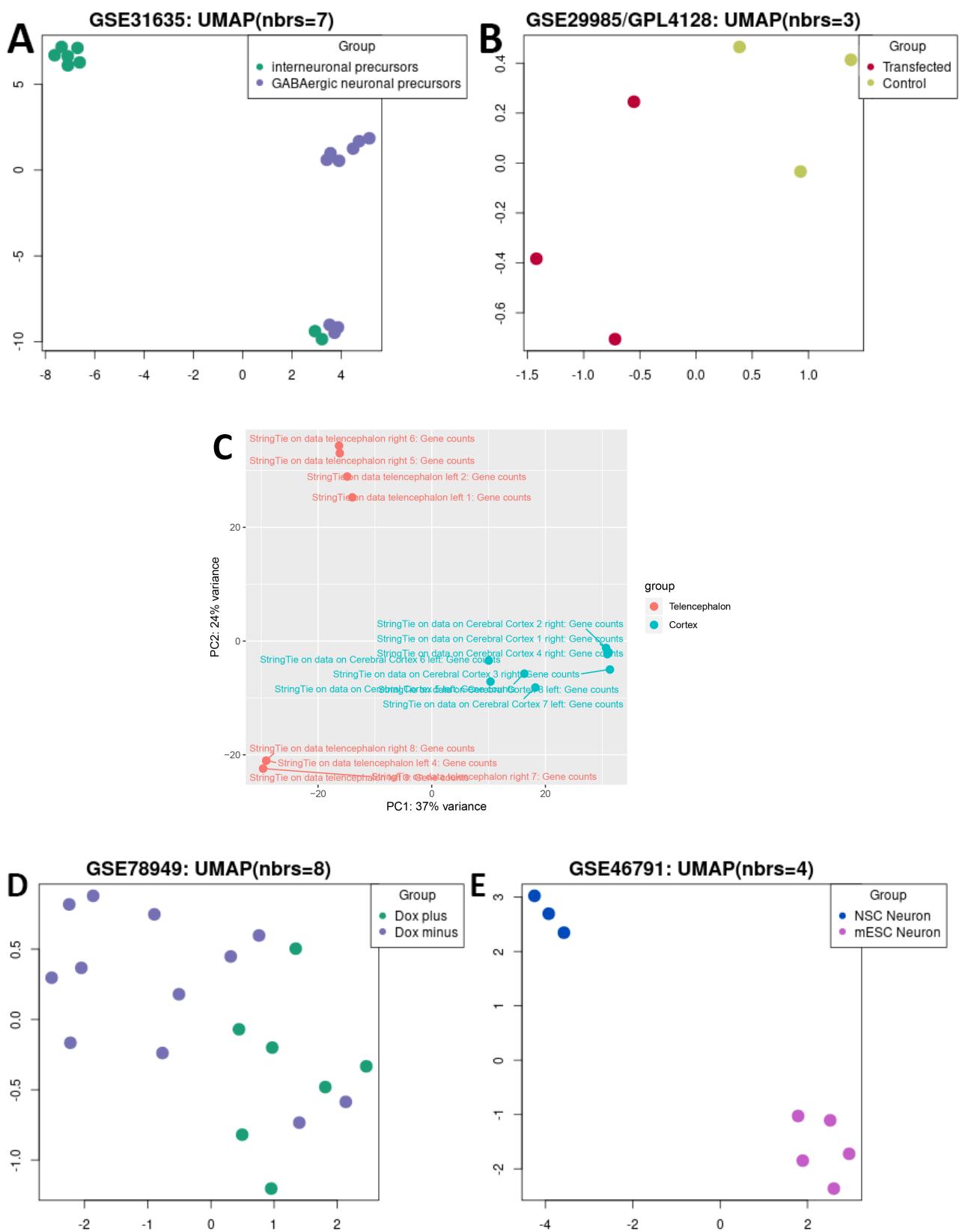


Figure 8: Principal component analysis (PCA) plots and uniform manifold approximation and projection (UMAP) plots of datasets. Generated from GEO2R and Galaxy. In each plot, the closeness of each sample to each other means that the samples are more similar there are. **A**, GSE31635 compared interneuronal precursors to GABAergic neuronal precursors . **B**, GSE29985 compared Infected group to the embryo brain group. **C**, E-MTAB-4840 compared the cortex group to the telencephalon group. **D**, GSE78949 compared

samples which contained Dox+ and Dox- treatment. E, GSE46791 compared the mouse embryonic stem cell (ESC) group to the neuron group.

C. Venn diagram comparative analysis

In order to identify which genes from the dataset overlap, the merged gene list, created from the filtering of upregulated genes using a log₂ fold change “greater than 0” and the filtering of downregulated genes which were filtered using a log₂ fold change “less than 0 function”. Both excel sheets are sorted by significance using a value of less than or equal to 0.05 for the P-adjacent value. From datasets: GSE29985 which looks at ARX and ASCL1 are associated with Dlx genes and both aid in tangential migration of GABAergic neurons (Colasante et al., 2008; Friocourt et al., 2008). GSE31635 explores the influence of ASCL1 and Dlx genes in the development of GABAergic neurons. GSE46791 which looks at the overexpression of transcription factors: Ascl1, Smad7 and Nr2f1 in embryonic stem cells (ESCs) which yield GABA-like neurons cells with expression profiles similar to neural tissues and highly expressed markers of GABAergic neurons which should contain targets of ASCL1 expression. E-MTAB-4840 looks at the upregulated genes in the telencephalon compared to the cortex and therefore the origin of the gabaergic neurons; Thus the dataset should therefore contain ASCL1 itself. GSE78949 looks at genes expressed in murine neural stem cells and compares them to differentiate GABAergic-like cells. Therefore ASCL1 target genes or GABAergic related genes should be present in the database. All the databases are placed into InteractiVenn and generated a Venn diagram (Figure 9A). Based on the purpose of looking for the gene overlap and identifying ASCL1 target genes that are involved in gabaergic neuron development or GABAergic expressed genes are associated with ASCL1.

The number of genes for GSE29985 was 1787. GSE31635 had 2315 genes. E-MTAB-4840 had 9975 genes. GSE78949 had 55 genes and GSE46791 had 1787. The datasets with the most genes were placed into another Venn diagram to emphasize the biological relevance in the overlap. As expected from the results obtained, the datasets: GSE29985, GSE31635 and GSE46791 only have 195 genes which are found between the three datasets (Appendix VI, Figure 9B). This shows that GSE29985, GSE31635 and GSE46791 share a similar biological relevance with each other. As GSE29985 ARX and ASCL1 interact with Dlx genes and promote GABAergic phenotype (Colasante et al., 2008; Figure 5; Friocourt et al., 2008; Table 2). GSE31635 looks at GABAergic precursors cells which are further committed than interneuronal precursor cells due to the influence of Dlx genes specifically which are associated with GABAergic neurons and ASCL1 binds to Dlx1/2 which promotes a GABAergic fate (Figure 5, Friocourt et al., 2008). GSE46791 compared the difference in gene expression in murine neural stem cells and differentiated neuronal progeny which expressed GABAergic phenotype (Table 1-2). Interestingly, GSE78949 which looks at the misexpression of ASCL1 in ESCs which produces GABA-like neurons and therefore, experimentally should have a greater proportion of ASCL1 target genes and a high expression profile related to the GABAergic phenotype, producing a Venn diagram with a greater degree of overlap with GSE78949.

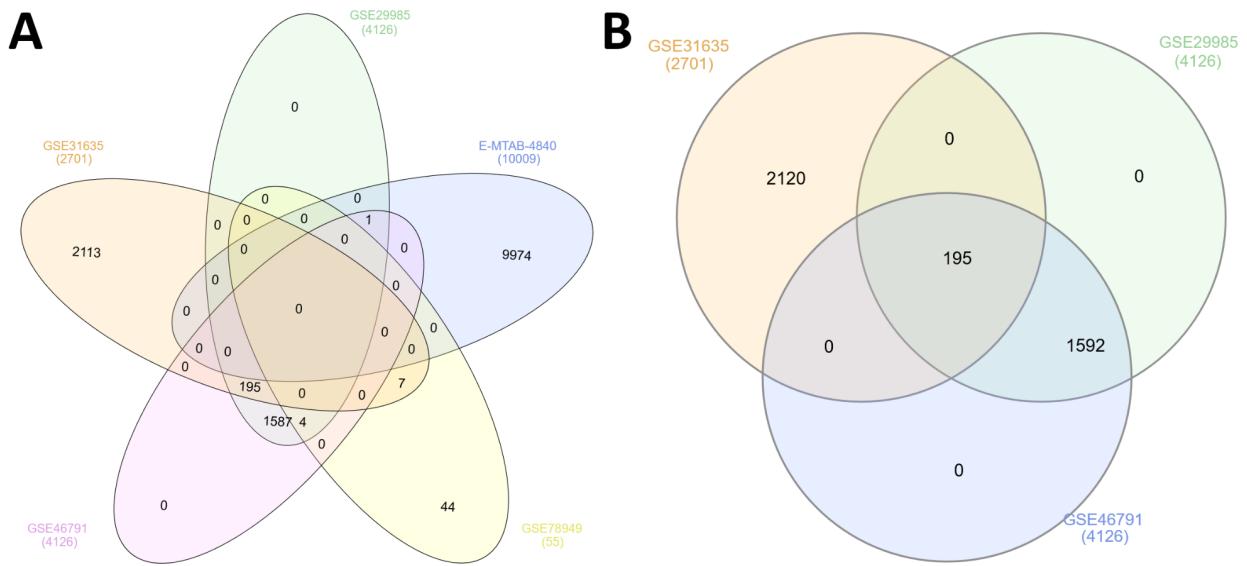


Figure 9: Venn diagrams generated by InteractiVenn. **A**, Showing that both of the upregulated and downregulated genes merged. Displaying the comparison of all 5 gene lists: GSE31635, GSE29985, E-MTAB-4840, GSE78949 and GSE46791. **B**, Showing that both of the upregulated and downregulated genes merged. Displaying the comparison of all 3 gene lists: GSE31635, GSE29985 and GSE46791.

D. GO term analysis

To investigate the relationship between datasets GSE31635, GSE29985, GSE46791 and gene ontology. A GO term analysis is performed using the same overlapping gene list generated from InteractiVenn (Appendix VI). Figure 10A displays 43 unique GO terms divided into molecular function, biological process and cellular component. 17 GO terms were attributed to cellular components, 19 GO terms were associated with biological processes and 7 GO terms were linked to molecular function. Figure 10B displays GO enrichment analysis graphs. The most enriched GO term for molecular function GO:0099510 (calcium ion binding involved in the regulation of cytosolic calcium ion concentration) and the least enriched term was GO:0016248 (channel inhibitor activity).

Likewise, in Figure 10C and Appendix IV, the most enriched GO term for the biological process was GO:0008015 (blood circulation) compared to the least enriched GO term was GO:0009653 (anatomical structure morphogenesis). Figure 10D and Appendix V showed that the GO term for cellular component, GO:0043197 (dendritic spine) was the most enriched whilst GO:0043226 (organelle) was the least enriched. This shows that the gene overlap generated from InteractiVenn (Appendix VI) contains GO terms that are associated with biological processes and cellular components can contain a greater proportion of genes rather than molecular function.

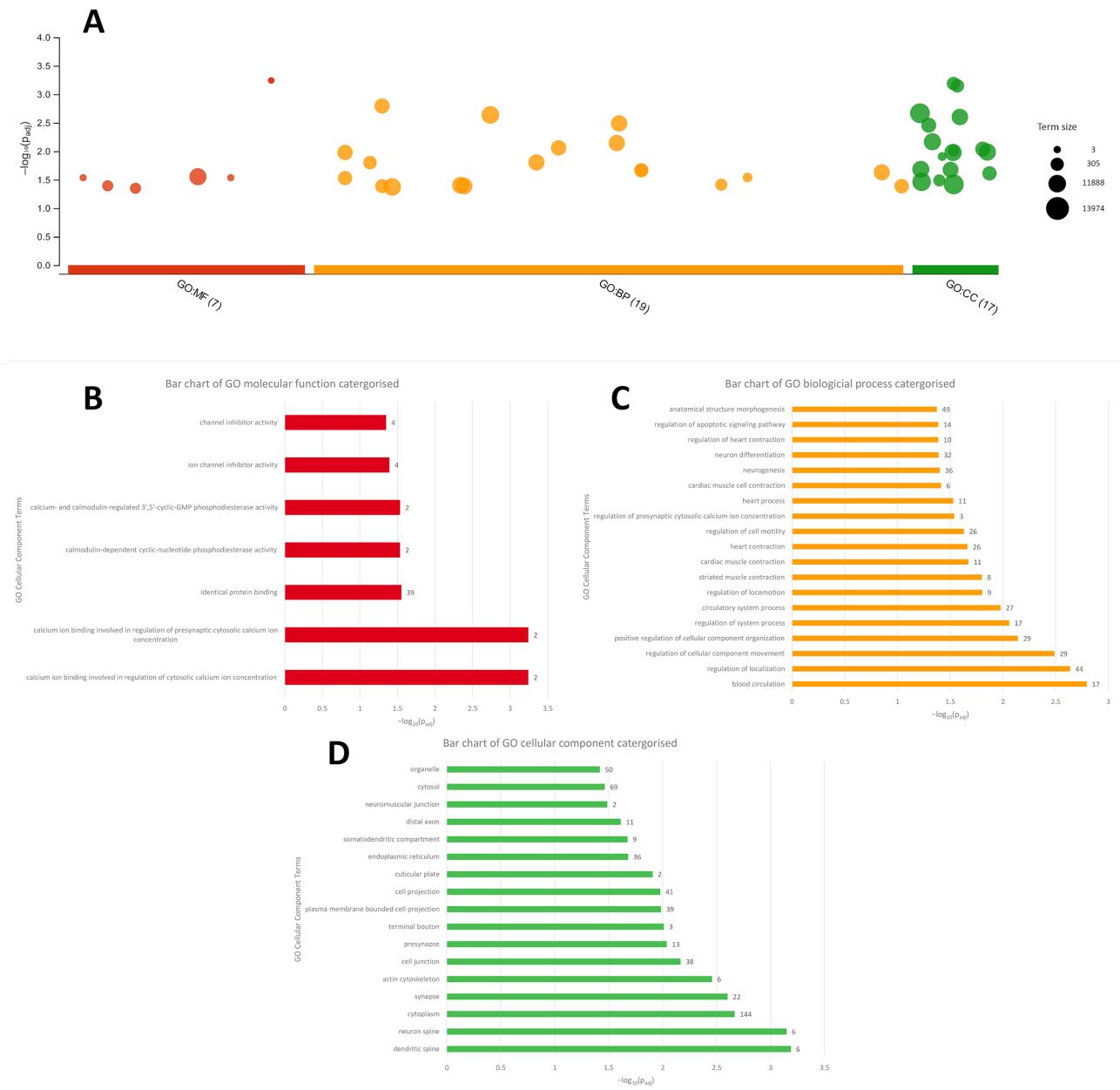


Figure 10: A, GO Term Enrichment analysis visualised as a Manhattan-like-bubble graph. The X-axis is a colour-coded and grouped GO category and the Y-axis shows the -Log10 (P adjusted value). The circle sizes represent term size. Larger terms have larger circles. Results of GO enrichment analysis as a bar chart for B, molecular function C, biological process and D, cellular component categories. The number of each bar represents the number of enriched genes annotated with the corresponding GO term. The X-axis shows the -Log10 of the adjusted P-value. The Y axis represents the GO term associated with each graph.

E. STRING PPI Network Analysis

To observe what gene-gene interactions are present in the list generated from InteractiVenn (Appendix VI), STRING was used. STRING PPI network analysis is based on the construction of identifying protein functions and interactions. Presented in Figure 10, the network identified 192 nodes, 98 edges which were higher than the expected number of edges at 69. With an average node degree of 1.02. The relevant interactions in the network include the interactions between GAD2, CALB2, CALB1 and GOT2 (Figure 11-12) which are associated with GABAergic phenotype. Other key proteins in the network which include NOTCH2 which are

associate with ASCL1 and promotion of neural genes (Engler et al., 2018; Long, Luo, Wang, Bates & Shetty, 2017), PITX2 and GATA2 which aids in the GABAergic differentiation process (Waite, Skidmore, Billi, Martin & Martin, 2011) and Slit2 which act as a molecular guidance cue in neuronal migration and axon guidance (Ba-Charvet et al., 2001). This shows that proteins highlighted in the STRING analysis have meaningful functions and are generally associated with function in the nervous system.

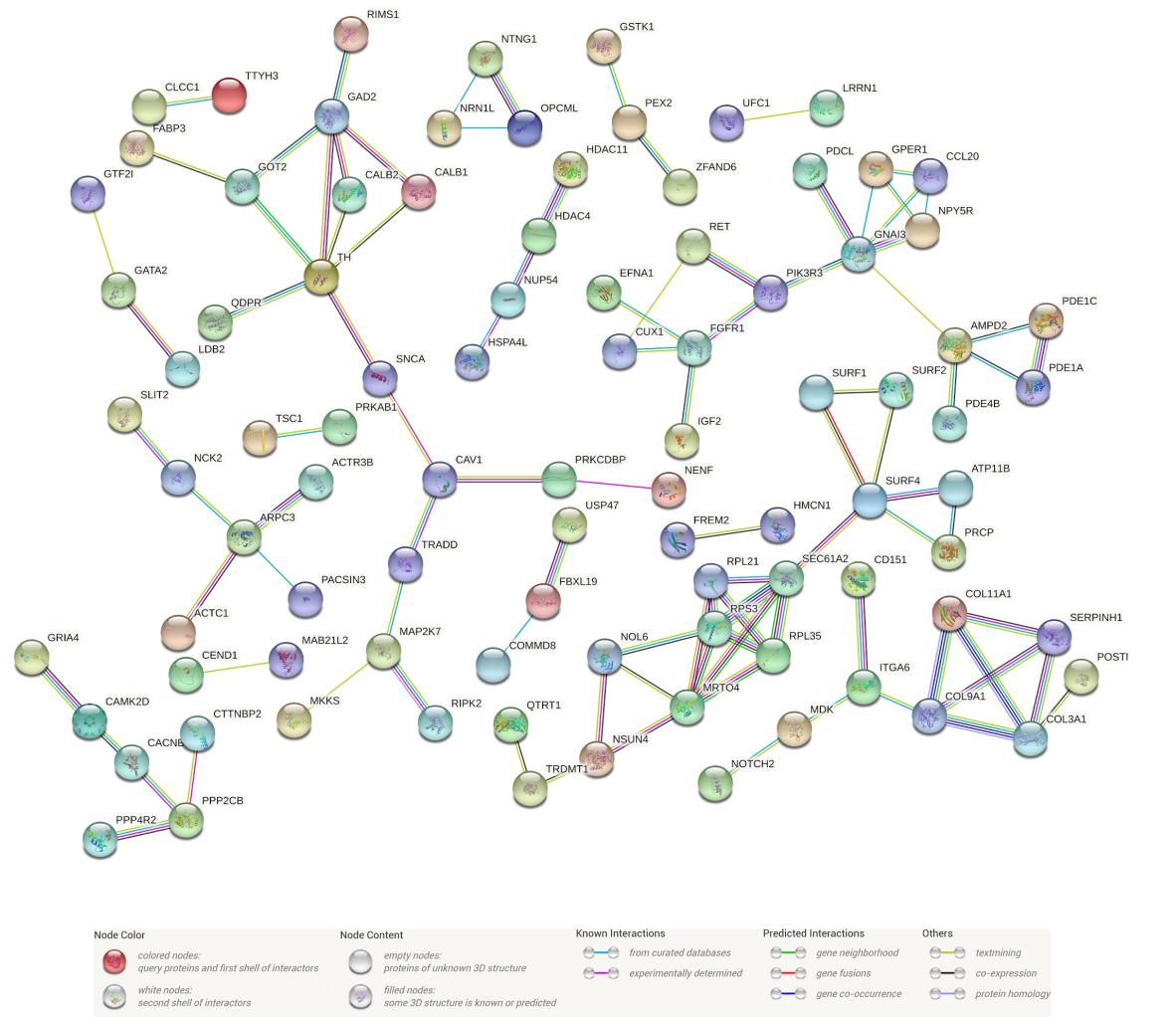


Figure 11: STRING analysis of merged genes which contain upregulated genes and downregulated genes.

Visual representation of all of the genes in the form of a network. Edges represent protein-protein associations which are specific and meaningful, proteins jointly contribute to a shared function; line colour indicates the type of interaction evidence denoted by the legend. The data was subjected to the full network type. The minimum required interaction confidence score was set at 0.65. The unconnected proteins are hidden in the network.

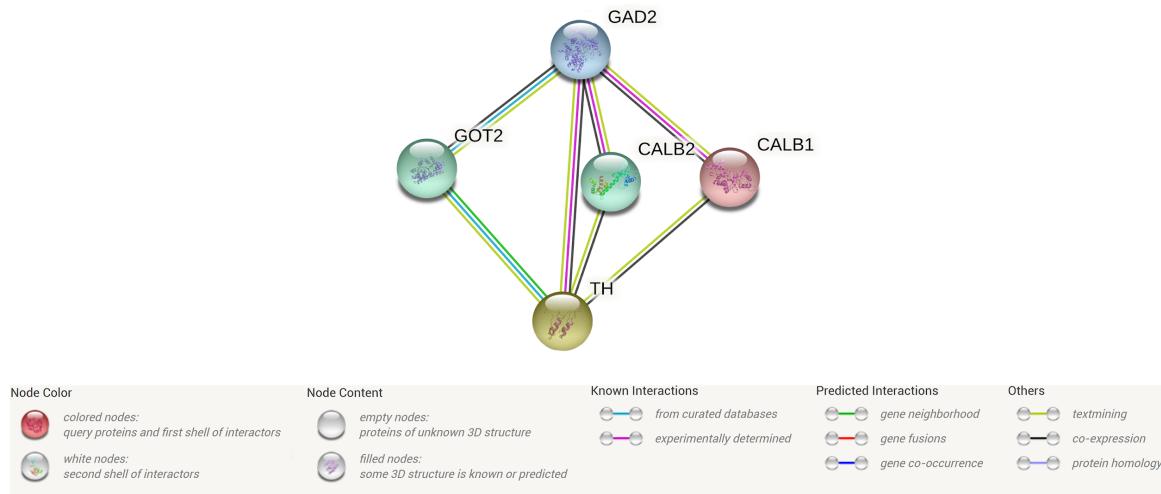


Figure 12: STRING analysis of merged genes which contain upregulated genes and downregulated genes predicted to be associated with GABAergic neurons using that analysis tab, native to STRING and selecting the reference papers which highlighted the associated genes. Visual representation of all of the genes in the form of a network. Edges represent protein-protein associations which are specific and meaningful, proteins jointly contribute to a shared function; line colour indicates the type of interaction evidence denoted by the legend. The data was subjected to the full network type. The minimum required interaction confidence score was set at 0.65. The unconnected proteins are hidden in the network.

6. Discussion

The investigation aimed to understand the role transcription factors play in the regulation of ASCL1 within GABAergic neurons. To identify co-expression groups that occur alongside ASCL1 and to investigate the relationship between the gene ontology and the ASCL1 targets genes. The findings suggest that ASCL1 expression is associated with several GO terms.

A. ASCL1 and the relationship with NOTCH2 and its role in expression promoting GABAergic proliferation and differentiation

As expected, ASCL1 expression promotes neural-specific factors as seen in the GO term analysis GO:0022008 (neurogenesis) and GO:0030182 (neuron differentiation)(Appendix IV). One of the genes mentioned is NOTCH2 which is a part of the NOTCH pathway and interacts with ASCL1 (Figure 5; Long, Luo, Wang, Bates & Shetty, 2017). NOTCH2 is expressed in VZ-SVZ and maintains quiescence of NSCs by suppressing cell cycle and neurogenic factors for GABAergic neurons. Deletions and knockouts of NOTCH2 resulted in the reactivation of the NSCs and the upregulation of neural genes (Engler et al., 2018; Long, Luo, Wang, Bates & Shetty, 2017) and promoting GABAergic-like phenotypes. This shows that NOTCH signalling inhibition contributes to the neuronal transformation in progenitor cells which need neurogenesis of the GABAergic phenotype.

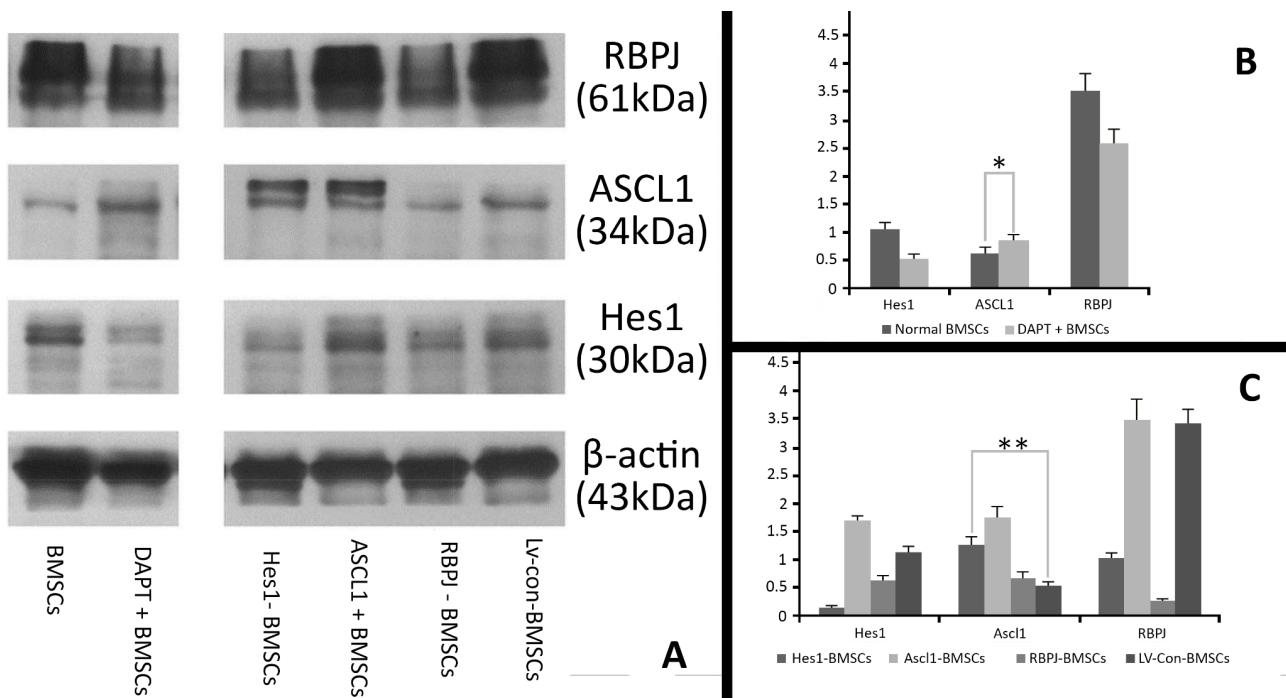


Figure 13: Protein levels RBPJ, ASCL1, Hes1 which are associated with NOTCH pathway with B-actin acting as the control housekeeping gene in bone marrow stromal cells (BMSCs) **A**, a western blotting of BMSCs that were treated by DAPT (γ -secretase inhibitor and acts as indirect inhibitor of NOTCH), or engineered by Hes1 short hairpin RNA (shRNA) (Hes1-BMSCs) alongside, lentiviral vector plasmid together with ASCL1 (ASCL1+BMSCs), RBPJ shRNA and copGFP Control which is derived from copepod *Pontellina plumata* (Lv-con-BMSCs). **B**, The protein level based on the ratio of Gauss Model Trace which displays the NOTCH signalling (RBPJ, Hes1 and ASCL1) changes in the DAPT+ BMSCs. **C**, The protein level based on the ratio of Gauss Model Trace exhibited genetically modified BMSCs. Error bars in bar graphs display standard deviation (SD). *P = 0.032, < 0.05, **P = 0.0006, < 0.01. Image modified from “Green fluorescent protein CopGFP”, n.d.; Long, Luo, Wang, Bates & Shetty, 2017”; Wang et al., 2012”.

B. The relationship between PITX2 and GATA2 in GABAergic differentiation

One of the genes found in the GO term analysis is PITX2. Present in 3/19 GO terms for biological processes (Appendix IV) and 1/7 GO terms for molecular function (Appendix III). PITX2 interacts with ASCL1 in the later transcription factor cascade (Waite, Skidmore, Billi, Martin & Martin, 2011) and promotes GATA2 which is a transcription factor expressed in post-mitotic neurons in the early stages of GABAergic differentiation (Kala et al., 2009) in the superior colliculus which activates GABAergic differentiation in the midbrain (Figure 14). GATA2 Identified in Figure 11 and 3/17 GO terms in cellular components (Appendix V) and 7/19 GO terms in biological processes (Appendix IV). However, it is not known whether a similar PITX2 and GATA2 expression mechanism are associated with the forebrain. Moreover, the role of PITX2 in neurogenesis is only applicable to the midbrain region.

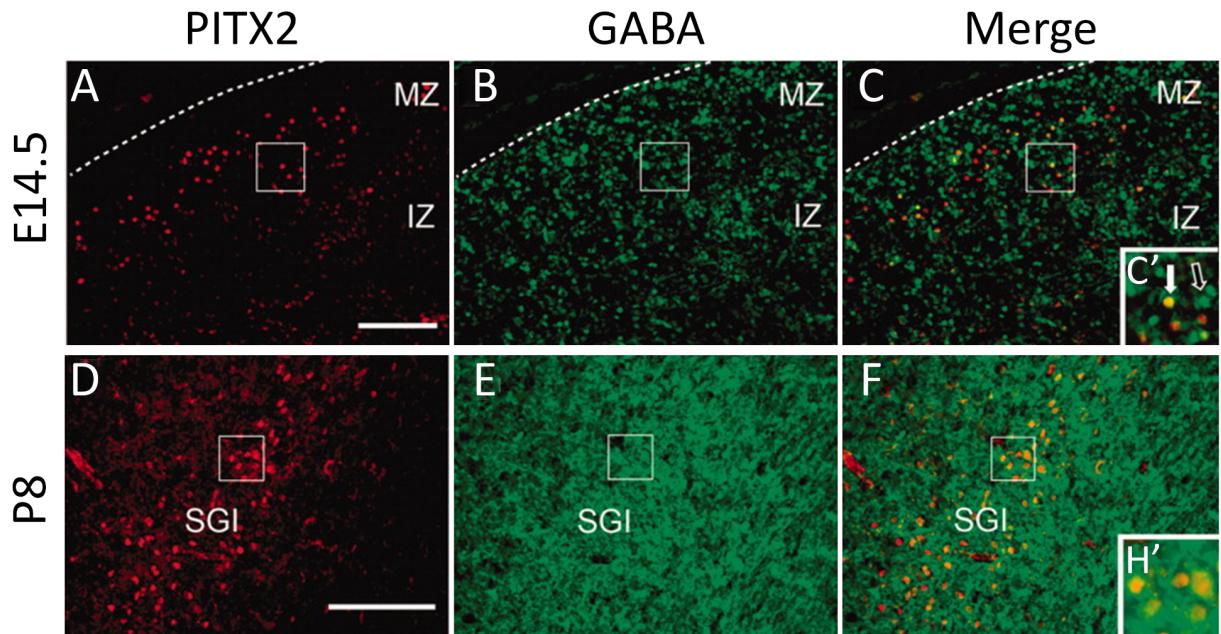


Figure 14: GABAergic interneurons located in the intermediate layer of the dorsal midbrain are associated with PITX2. **A–F**, E14.5 and P8 murine midbrains were processed for immunofluorescence for PITX2 (red) and GABA (green). **C', F'**, show enlarged boxes in **C** and **F**. At E14.5, PITX2-positive marked and GABA-positive marked cells are located in the intermediate and mantle regions of the superior colliculus, (**C'**) show PITX2-positive cells are also GABA-positive. The white filled arrow in **C'** of PITX2 and GABA are co-localised. The open white arrow demonstrates a GABAergic neuron with no PITX2 expression. **D–F'**: At P8, PITX2-positive cells are located in the intermediate layer GABAergic neurons. Dotted lines display the outline of the acetylcholinesterase positive layer. The scale bar in **C** represents 100 μ m and applies to **C–H**. The terms MZ stands for mantle zone; IZ stands for intermediate zone and SGI stands for stratum griseum intermedium. Image modified from “Waite, Skidmore, Billi, Martin & Martin, 2011”.

C. ASCL1 expression is important for synaptic communication of GABA

ASCL1 expressions can activate several genes in question. CALB1 (Calbindin, CB) is expressed in GABAergic neurons in the prefrontal cortex and cerebellum (Baimbridge, Miller & Parkes, 1982; Harris, Abel, Tejada & Rissman, 2016) and is present in 15/17 GO terms in cellular component (Appendix V); present in 2/19 GO terms for biological processes (Appendix IV) and 2/7 GO terms for molecular function (Appendix III). CALB1 acts as a high-affinity calcium sensor in GABAergic neurons. CALB1 together with CALB2 (Calretinin, CR) which is a calcium-binding protein that shares homology to CALB1 (Rogers, 1987; Taliano et al., 2013) and found in 14/17 GO terms (Appendix V) and the STRING analysis (Figure 11–12) are markers of maturation

neurogenesis. With CALB2 being a marker of early neuronal maturation and CALB1 being a marker of post-mitotic late maturation. The expression of CALB1 and CALB2 from prenatal development to adulthood remains persistent throughout (Figure 15; Kumar et al., 2020).

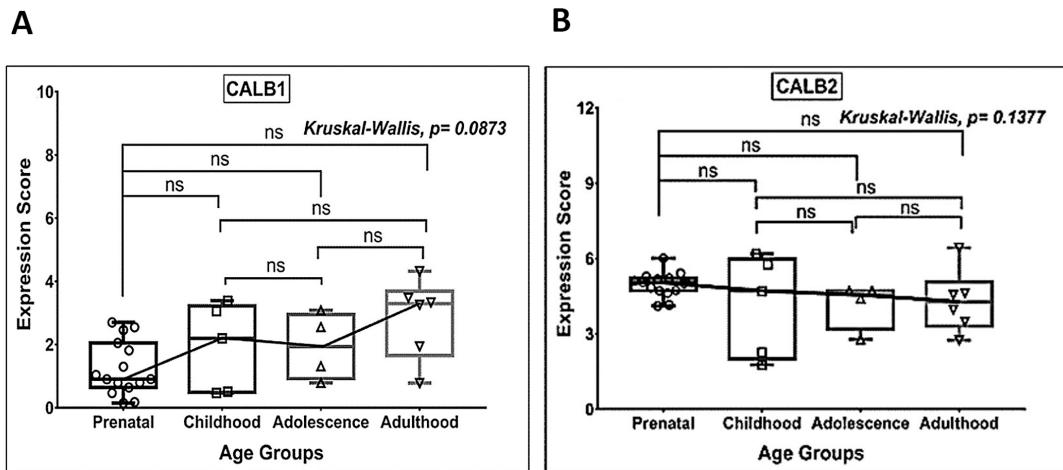


Figure 15: Box plot presentation (median with minimum and maximum bars showing all data points) neurogenic markers **A**, CALB1 and **B**, CALB2. Gene expression scores in \log_2 of RPKM (reads per kilobase of exon model per million mapped reads). Statistical comparisons were made using the Kruskal-Wallis test followed by a post hoc Mann Whitney U test and FDR correction. Adjusted p-values for MWU comparisons represented by, **** ≤ 0.0001 , *** ≤ 0.001 , ** ≤ 0.01 , and ns for non-significant. Image modified from "Kumar et al., 2020".

The functions of GABAergic neurons in the production of GABA which is the key inhibitory neurotransmitter in the CNS (Pan, 2012) which modulates the inhibitory-excitatory dichotomy in brain function (Pan, 2012; Wu & Sun, 2014). One of the genes identified in the STRING and GO term analysis is GAD2. Also known as GAD65, is found in 5/17 GO term for biological process (Appendix V, Figure 10C, Figure 11-12) is a glutamate decarboxylase which converts glutamate into GABA which is recycled in the citric acid pathway (Rashmi, Zanan, John, Khandagale & Nadaf, 2018). GAD2 shares a homology to GAD1 (Erlander, Tillakaratne, Feldblum, Patel & Tobin, 1991; Krueger, Stöcker & Schlosser, 2007) which is co-expressed with CALB1 and CALB2 (Figure 11-12). Quantitative analysis showed that 25% CALB1 expressing GABA neurons expressed GAD2 only and the total grey matter in cortical layers 1 and 6 showed 48% of the terminals expressing CALB1/GAD2. Additional quantitative analysis showed approximately 18% of CALB2 GABA terminals contained GAD2, approximately 37% contained GAD1, and approximately 45% contained both GAD proteins. Fluorescence intensity evidence, which measured the relative amount of a protein, found that CALB1/GAD2 expressing GABA terminals were approximately 34% greater compared to in CALB1/GAD2/GAD1 terminals. whereas the relative amount of GAD1 was approximately 24% greater in CALB1/GAD2/GAD1 terminals compared to that in CALB1/GAD1+ terminals. Fluorescence intensity evidence demonstrated that the total amount of GAD2 in CALB2/GAD2/GAD1+ terminals was approximately 41% greater than in CALB2 /GAD2 terminals whereas the total amount of GAD1 was approximately 63% greater in CALB2/GAD2/GAD1 terminals than in CALB2/GAD1+ terminals (Fish, Rocco & Lewis, 2018). This shows that the relationship between GADs and calcium-binding proteins are important in synaptic transmission in GABAergic neurons. Moreover, further demonstrates that neurogenesis for GABAergic neurons and the different subpopulations of GABAergic neuron terminals could highlight play different roles in the modulation of excitatory transmission from glutamatergic neurons.

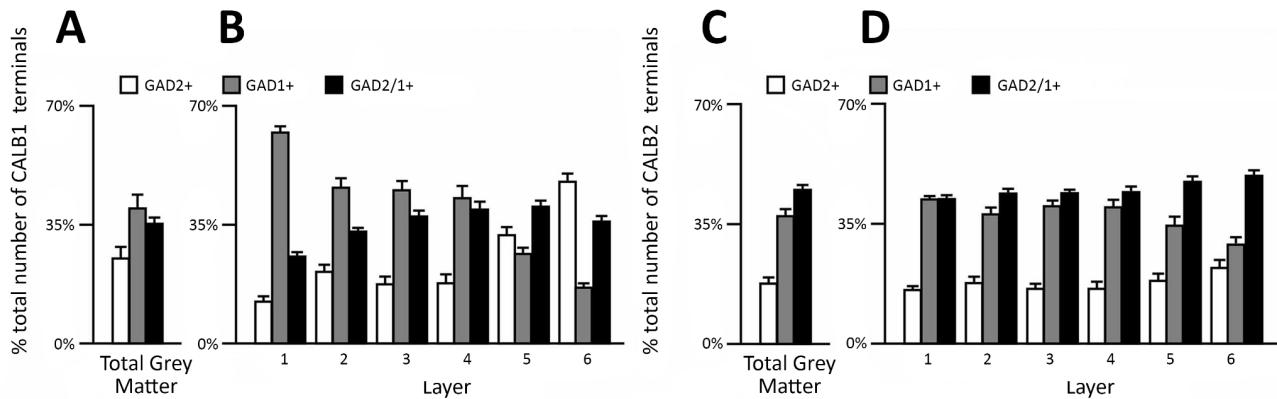


Figure 16: **A**, bar graph showing the percentage of CALB1 terminals which express GAD2, GAD1 or GAD2/1 together in the total grey matter content. **B**, Bar graphs showing the relative proportions of vGAT/CALB1+ terminals in PFC with the six cortical layers that were GAD2+, GAD1+, or GAD2/GAD1+. **C**, bar graph showing the percentage of CALB2 terminals which express GAD2, GAD1 or GAD2/1 together in the total grey matter content. **D**, Bar graphs showing the relative proportions of vGAT/CALB2+ terminals in PFC that were GAD2+, GAD1+, or GAD2/GAD1+. Error bars use the standard error of the mean. Bar charts do not contain data relating to fluorescent intensity data. Image modified from “Fish, Rocco & Lewis, 2018”.

D. ASCL1 and SLIT2 conflicted role in cell mobility and neuronal modulation

ASCL1 expression in GABAergic neurogenesis is associated with Slit2 (Appendix III-IV; Figure 11). One of three homologous proteins in the Slit family that interacts with Robo receptors (Silver, Horn, Busch & Yonkof, 2009; Tong, Jun, Nie, Hao & Fan, 2019). Results from *in situ* hybridization showed that Slit proteins in early development showed that Slit1 is expressed in MGE and LGE from the VZ and Slit2 is expressed in the MGE and expressed weakly in the VZ (Andrews et al., 2008; Yeh et al., 2014). One explanation for this difference is that VZ disappears later in neural development (Dehay & Kennedy, 2007; Quiñones-Hinojosa & Chaichana, 2007) compared to the SVZ and therefore could explain why there is a weak expression of Slit2 in the VZ. Slit proteins repel GABAergic neurons from the SVZ of LGE and inhibit tangential migration in brain explants (Zhu, Li, Zhou, Wu & Rao, 1999). However further research has shown that, when specifically Slit2 is knocked out, abnormal axonal trajectories are produced and axons undergo defasciculation but only in the hippocampal region and not the prefrontal cortex (Bagri et al., 2002). Double mice knockout for Slit1 and Slit2 shows increased activation of the neural progenitor phenotype in the LGE (Yeh et al., 2014). But this conflicts with a previous experiment that showed that tangential migration from the GE to the cortex in Slit knockout mice was normal both in Slit1/Slit2 knockout and in triple Slit1/Slit2/Ntr1 knockout mice, which suggest that Slit proteins may not be required for the migration of interneurons from the subpallium to the pallium (Andrews, Barber & Parnavelas, 2007; Marín et al., 2003). This further suggests that Slit2's role is ambiguous and may play a minuscule role in directing tangential migration and may regulate the morphology of axons and their proliferation neural progenitors.

E. Limitations of experiment design

An unusual finding in Figure 9B is that database E-MTAB-4840, only had two overlaps; between datasets E-MTAB-4840, GSE78949 and GSE29985 and E-MTAB-4840 and GSE46791. This is even though Figure 7 not only had the largest number of genes of 9975 and had the largest number of human embryonic samples at 632 which spanned from 4 post-conception weeks to 12 post-conception. Therefore theoretically, the vast majority of the genes from E-MTAB-4840 should have overlapped with 4 datasets. One reason proposed for this discrepancy is that the remaining datasets used GABAergic precursors, GABAergic adult neurons or misexpression of GABAergic factors in different organisms in each dataset. As % datasets used *Mus musculus* and ½ datasets used *Rattus norvegicus* (Table 1). This could explain why in Figure 7C there was a greater proportion of downregulated genes being expressed. As the brain matures many embryonic genes

associated with neurogenesis are downregulated. In hindsight, if the younger samples derived from the Carnegie stage 13/14 stage in forebrain development would have been more informative as it would have specifically looked at neurogenesis at an earlier stage.

Interestingly, Figure 8D has the weakest clustering and highest dispersion despite the samples being separated into Dox+ and Dox- groups which should provide distinctive clusters in dataset GSE78949. The expression of ASCL1 was repressed by doxycycline. Therefore, in theory, should result in clear sample clusters. However, the fact that Figure 8D's PCA was highly dispersed suggests that the samples are similar to each other; as the purpose of the principal component analysis (PCA) plot is to ensure and observe that the selected samples in each group are distinct from each other. This could explain why in Figure 9A, dataset GSE78949 had 55 genes present in the Venn diagram gene overlap and why Figure 7D had the fewest upregulation and downregulation of genes in the volcano plot analysis.

F. Future research

A future area of research is to reconducted aspects of the experiment using adult human brain samples and compare them to database E-MTAB 4840 and identify what genes have been upregulated and downregulated. As a way to observe embryonic neurogenesis and adult neurogenesis directly using RNA-seq analysis to directly observe ASCL1 and adult neurogenesis in research is limited. As early research detected PSA-NCAM marked cells which are associated with neurogenesis (Quartu et al., 2008) from immunohistochemistry experiments on surgically removed hippocampi and sections of the entorhinal cortex in patients with drug-resistant temporal lobe epilepsy displayed numerous PSA-NCAM cells are detectable in the subgranular zone (SGZ) of the adult human hippocampus (Mikkonen, Soininen, Tapiola, Alafuzoff & Miettinen, 1999; Seki, 2020). However, damage to the brain after death and the comorbidities found in post mortem exams as well as the limited experiment data surrounding the topic. Alternatively, a consensus can be reached by using which neural stem cells from human samples could be used to compare database E-MTAB 4840.

Another area future area of research is to explore why Slit2 is expressed weakly in VZ and when Slit2 is knockouted, abnormal axonal trajectories are produced and axons undergo defasculciation in the hippocampal region only and not the prefrontal cortex. Thus, one idea is to ectopically express Slit2 in the prefrontal cortex using reporter fusion constructs in mice models which uses the Slit1 promoter, the coding region of the Slit2 together with a reporter gene such as GFP. This will provide interesting evidence on the relationship between Slit2 and ASCL1 and role it plays in neurogenesis.

Moreover, the interactions between PITX2 with ASCL1, which promotes GATA2 in GABAergic precursors (Kala et al., 2009; Waite, Skidmore, Billi, Martin & Martin, 2011) is only observed in the midbrain. One area of research could explore whether PITX2/GATA2 expression are associated with GABAergic precursors in the forebrain as current research on PITX2 and other areas of neurogenesis is limited.

7. Conclusion

The initial aims of the experiment were: to understand the role of regulation of ASCL1 within GABAergic neurons, to identify co-expression groups that occur alongside ASCL1 and to investigate the role of ASCL1 targets genes by gene ontology using bioinformatic techniques. And throughout the experiment, the aims were achieved as ASCL1 plays an important role in the tangential migration and neural differentiation of NSCs. Several gene keys were identified and associated with GABAergic neurons and neurogenesis: GAD2, CALB1/2, PITX2, GATA2 and NOTCH2 and those same genes were identified in the GO term and STRING analysis which demonstrate their interactions and influence with each other within GABAergic neurons. However, it cannot be ignored that databases E-MTAB-4840 and GSE78949 had little to no biological relevance despite both datasets exhibiting ASCL1 expression, contained ASCL1 target genes or contained GABAergic genes. However, future research can better understand the conflicted role of Slit2 in neurogenesis, explore the interaction of PITX2 and GATA2 in the context of GABAergic neurons in and also as would investigate adult human neurogenesis and embryonic human neurogenesis directly.

8. References

- About GEO2R. Retrieved 20 February 2021, from <https://www.ncbi.nlm.nih.gov/geo/info/geo2r.html>

This website was used to give background information on GEO2R. How to use the series accession number and define Sample groups and describe the various data visualisations.

- Alzu'bi, A., & Clowry, G. (2019). Expression of ventral telencephalon transcription factors ASCL1 and DLX2 in the early fetal human cerebral cortex. *Journal Of Anatomy*, 235(3), 555-568. doi: 10.1111/joa.12971

This journal article discusses the relationship between the proneural transcription factor ASCL1 and the transcription factor DLX2 that controls differentiation. The RNAseq dataset: E-MTAB-4840 from which data were extracted for this study was acquired from this study.

- Anderson, S., Eisenstat, D., Shi, L., & Rubenstein, J. (1997). Interneuron Migration from Basal Forebrain to Neocortex: Dependence on Dlx Genes. *Science*, 278(5337), 474-476. doi: 10.1126/science.278.5337.474 10.1111/joa.12971

This journal article discusses the relationship between, γ -aminobutyric acid (GABA) expressing cells migrate in vitro from the subcortical telencephalon into the neocortex by relying on homeodomain proteins DLX-1 and DLX-2.

- Anderson, S., Mione, M., Yun, K., & Rubenstein, J. (1999). Differential Origins of Neocortical Projection and Local Circuit Neurons: Role of Dlx Genes in Neocortical Interneuronogenesis. *Cerebral Cortex*, 9(6), 646-654. doi: 10.1093/cercor/9.6.646

This journal article reviews the data supporting the argument that the relationship that neocortical projection neurons and interneurons are derived from distinct regions within the telencephalon; the ventricular zone of the neocortex and neocortical interneurons derived from the germinal zone of the basal ganglia. The Dlx homeobox genes are discussed as being important for basal ganglia differentiation and DLX2 expression in neocortical cells can induce GABAergic interneuron differentiation.

- Andrews, S. FastQC A Quality Control tool for High Throughput Sequence Data. Retrieved 25 February 2021, from <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>

This website describes the main functions and the role of FastQC in quality control in checks on raw read sequences.

- Andrews, W., Barber, M., & Parnavelas, J. (2007). Slit-Robo interactions during cortical development. *Journal Of Anatomy*, 211(2), 188-198. doi: 10.1111/j.1469-7580.2007.00750.x

This paper explores the role of Slit proteins which act through Robo receptors and play a role in axon guidance and neuronal migration of cortical interneurons.

- Andrews, W., Barber, M., Hernandez-Miranda, L., Xian, J., Rakic, S., & Sundaresan, V. et al. (2008). The role of Slit-Robo signaling in the generation, migration and morphological differentiation of cortical interneurons. *Developmental Biology*, 313(2), 648-658. doi: 10.1016/j.ydbio.2007.10.052

This paper looks at the migration of cortical interneurons tangentially into the cortex in rodents which are generated in the ventral telencephalon and migrate. The study demonstrates that Robo1 and Robo2 proteins are expressed throughout corticogenesis. The analysis of Robo and Slits mutants showed a marked

increase in the number of interneurons in the cortices of Robo1, Slit1 and Slit2 mutant knockout mice, suggesting that cell division is influenced by Slit-Robo signalling mechanisms.

- ASCL1 Gene (Protein Coding). Retrieved 24 July 2020, from <https://www.genecards.org/cgi-bin/carddisp.pl?gene=ASCL1>

This website provides the background information on ASCL1.

- Ashburner, M., Ball, C., Blake, J., Botstein, D., Butler, H., & Cherry, J. et al. (2000). Gene Ontology: tool for the unification of biology. *Nature Genetics*, 25(1), 25-29. doi: 10.1038/75556

This journal provides background information on why gene ontology is needed in the analysis of eukaryotic genomes and defines the three gene ontologies; cellular component, biological process and molecular function.

- Aydin, B., Kakumanu, A., Rossillo, M., Moreno-Estellés, M., Garippler, G., & Ringstad, N. et al. (2019). Proneural factors Ascl1 and Neurog2 contribute to neuronal subtype identities by establishing distinct chromatin landscapes. *Nature Neuroscience*, 22(6), 897-908. doi: 10.1038/s41593-019-0399-y

This paper investigates Ascl1 and Neurog2 ability to induce neuronal fates by binding to largely different sets of genomic sites which affects their chromatin accessibility and enhancer activity profiles.

- Ba-Charvet, K., Brose, K., Ma, L., Wang, K., Marillat, V., & Sotelo, C. et al. (2001). Diversity and Specificity of Actions of Slit2 Proteolytic Fragments in Axon Guidance. *The Journal Of Neuroscience*, 21(12), 4281-4289. doi: 10.1523/jneurosci.21-12-04281.2001

This paper discusses the role of Slit2 as a chemorepellent for developing axons and is involved in the control of midline crossing. The paper demonstrated *in vivo*, Slit2 is cleaved into fragments (140 kDa N-terminal (Slit2-N) and 55-60 kDa C-terminal (Slit2-C) and observed its ability to bind to Robo proteins. It was demonstrated that olfactory bulb (OB) but not dorsal root ganglia (DRG) axons are repelled by Slit2-N and Slit2-U (uncleavable protein).

- Bagri, A., Marin, O., Plump, A., Mak, J., Pleasure, S., Rubenstein, J., & Tessier-Lavigne, M. (2002). Slit Proteins Prevent Midline Crossing and Determine the Dorsoventral Position of Major Axonal Pathways in the Mammalian Forebrain. *Neuron*, 33(2), 233-248. doi: 10.1016/s0896-6273(02)00561-5

This paper reported on the role of Slit proteins and demonstrated that mice deficient in Slit2 and Slit2, show significant axon guidance errors in a variety of tract pathways. In addition, the analysis of multiple pathways resulted in several broad conclusions regarding the role of Slit proteins in the brain which include in maintaining prevention of axonal growth into ventral regions and prevention of axonal growth towards the midline.

- Baimbridge, K., Miller, J., & Parkes, C. (1982). Calcium-binding protein distribution in the rat brain. *Brain Research*, 239(2), 519-525. doi: 10.1016/0006-8993(82)90526-1

This paper look at distribution of CaBP within the rat hippocampus and The significance of this protein in relation to the CNS

- Barker-Haliski, M., & White, H. (2015). Glutamatergic Mechanisms Associated with Seizures and Epilepsy. *Cold Spring Harbor Perspectives In Medicine*, 5(8), 1-16. doi: 10.1101/cshperspect.a022863

This paper looks at the medical condition epilepsy is broadly characterized by aberrant neuronal excitability and talks about the roles of Glutamate is the predominant excitatory neurotransmitter in the adult mammalian brain and it contributes to seizures and epilepsy.

- Baronti, L., Hošek, T., Gil-Caballero, S., Raveh-Amit, H., Calçada, E., & Ayala, I. et al. (2017). Fragment-Based NMR Study of the Conformational Dynamics in the bHLH Transcription Factor Ascl1. *Biophysical Journal*, 112(7), 1366-1373. doi: 10.1016/j.bpj.2017.02.025

This paper looks at NMR analysis of the ASCL1 protein using fragment-based NMR to study and obtain resolution information on the structural properties and conformational dynamics of ASCL1.

- Barrett, T., Wilhite, S., Ledoux, P., Evangelista, C., Kim, I., & Tomashevsky, M. et al. (2012). NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Research*, 41(D1), D991-D995. doi: 10.1093/nar/gks1193

This paper refers to the public genomic database, Gene Expression Omnibus (GEO) which is an international public repository for high-throughput microarray and next-generation sequence functional genomic datasets submitted. GEO is the conduit for the R-based web application GEO2ER, which allows analysis of GEO data.

- Battiste, J., Helms, A., Kim, E., Savage, T., Lagace, D., & Mandyam, C. et al. (2007). Ascl1 defines sequentially generated lineage-restricted neuronal and oligodendrocyte precursor cells in the spinal cord. *Development*, 134(2), 285-293. doi: 10.1242/dev.02727

This paper looks at ASCL1 expression in progenitors to both neurons and oligodendrocytes, but not astrocytes. They found that ASCL1 expression results in rapid cell-cycle exit and differentiation; ASCL1 expression differs based on the age of the embryo and what cells are expressed.

- Begega, A., Alvarez-Suarez, P., Sampedro-Piquero, P., & Cuesta, M. (2017). Effects of Physical Activity on the Cerebral Networks. *Physical Activity And The Aging Brain*, 3-11. doi: 10.1016/b978-0-12-805094-1.00001-0

This journal reviews the process of ageing and the association of a decline in cognitive functions. This paper also investigates methods of preservation and maintaining cognitive function primarily through physical exercise and mental training.

- Bertrand, N., Castro, D., & Guillemot, F. (2002). Proneural genes and the specification of neural cell types. *Nature Reviews Neuroscience*, 3(7), 517-530. doi: 10.1038/nrn874

This paper discusses the morphological, physiological and molecular characteristics of neurons and identifies differences and similarities between invertebrate and vertebrate neural lineages.

- Boldrini, M., Fulmore, C., Tartt, A., Simeon, L., Pavlova, I., & Poposka, V. et al. (2018). Human Hippocampal Neurogenesis Persists throughout Aging. *Cell Stem Cell*, 22(4), 589-599.e5. doi: 10.1016/j.stem.2018.03.015

This paper looks at the neural cell populations and found that in young and old adults found that similar numbers of intermediate neural progenitors and immature neurons to the differentiated cell population and showed that although older individuals have less angiogenesis and neuroplasticity and a smaller quiescent progenitor pool. Healthy older subjects that do not suffer from cognitive impairment and neuropsychiatric disease can have hippocampal neurogenesis occur in their lives.

- Bolger, A., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114-2120. doi: 10.1093/bioinformatics/btu170

This original paper describes the purpose of trimmomatic in the processing of RNA seq data.

- Brandão, J., & Romcy-Pereira, R. (2015). Interplay of environmental signals and progenitor diversity on fate specification of cortical GABAergic neurons. *Frontiers In Cellular Neuroscience*, 9. doi: 10.3389/fncel.2015.00149

This paper reviews the diversity of cortical GABAergic interneurons and discusses the production of cortical GABAergic neuron diversity. It also focuses on the various types of calcium-binding proteins which are expressed alongside various biochemical markers found in GABAergic neurons subtypes.

- Castro, D., Martynoga, B., Parras, C., Ramesh, V., Pacary, E., & Johnston, C. et al. (2011). A novel function of the proneural factor Ascl1 in progenitor proliferation identified by genome-wide characterization of its targets. *Genes & Development*, 25(9), 930-945. doi: 10.1101/gad.627811

This paper discusses ASCL1 and the associated role of promoting cell cycle exit and neuronal differentiation when expressed in neural progenitor cells. The mutational analysis in the embryonic brain due to ASCL1 manipulation demonstrates it is required for normal proliferation of neural progenitors.

- CIViC: Gene ASCL1 Summary. Retrieved 24 July 2020, from <https://civicdb.org/events/genes/436/summary#gene>

This website provides a background gene summary on ASCL1

- Clough, E., & Barrett, T. (2016). The Gene Expression Omnibus Database. *Methods In Molecular Biology*, 1418, 93-110. doi: 10.1007/978-1-4939-3578-9_5

This paper discusses the Gene Expression Omnibus (GEO) database is a public genomic international database which archives high-throughput gene expression and other functional genomics data sets. As well as provide the ability to process and visual the data using various web-based tools.

- Colasante, G., Collombat, P., Raimondi, V., Bonanomi, D., Ferrai, C., & Maira, M. et al. (2008). Arx Is a Direct Target of Dlx2 and Thereby Contributes to the Tangential Migration of GABAergic Interneurons. *Journal Of Neuroscience*, 28(42), 10674-10686. doi: 10.1523/jneurosci.1283-08.2008

This paper discusses the relationship between ARX and Dlx2 and its relationship to ASCL1 and GABAergic neurons.

- Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., & McPherson, A. et al. (2016). A survey of best practices for RNA-seq data analysis. *Genome Biology*, 17. doi: 10.1186/s13059-016-0881-8

This paper discusses the pipeline in the analysis of RNA seq data.

- Dehay, C., & Kennedy, H. (2007). Cell-cycle control and cortical development. *Nature Reviews Neuroscience*, 8(6), 438-450. doi: 10.1038/nrn2097

This paper discusses the process of corticogenesis in mammals and differences found in several species including mice and primates as well as the features and mechanism in the cell cycle for neural cells.

- Dennis, D., Han, S., & Schuurmans, C. (2019). bHLH transcription factors in neural development, disease, and reprogramming. *Brain Research*, 1705, 48-65. doi: 10.1016/j.brainres.2018.03.013

This paper reviews the role of basic-helix-loop-helix (bHLH) proteins in neural specification and differentiation and investigates mutation in bHLH which are associated with various neurological diseases and cancers.

- Doyle, M., Phipson, B., & Dashnow, H. (2020). RNA-Seq reads to counts (Galaxy Training Materials). Retrieved 5 December 2020, from <https://training.galaxyproject.org/training-material/topics/transcriptomics/tutorials/rna-seq-reads-to-counts/tutorial.html#citing-this-tutorial>

The website describes how to import, process, count and map RNA seq data in the program galaxy.

- Doyle, M. (2021). Visualization of RNA-Seq results with Volcano Plot. Retrieved 15 March 2021, from <https://training.galaxyproject.org/training-material/topics/transcriptomics/tutorials/rna-seq-viz-with-volcanoplot/tutorial.html>

The website describes how to visualise RNA-Seq results with volcano plot in the program galaxy.

- Engler, A., Rolando, C., Giachino, C., Saotome, I., Erni, A., & Brien, C. et al. (2018). Notch2 Signaling Maintains NSC Quiescence in the Murine Ventricular-Subventricular Zone. *Cell Reports*, 22(4), 992-1002. doi: 10.1016/j.celrep.2017.12.094

This paper describes the process of neurogenesis that occurs in the ventricular-subventricular zone(V-SVZ)of the adult forebrain from quiescent neural stem cells (NSCs). The paper also discusses the relationship with Notch2 is used to repress cell-cycle related genes and inhibits neurogenesis in V-SVZ NSCs. Loss of Notch2 activates quiescent NSCs, which proliferate and generate new neurons.

- Erlander, M., Tillakaratne, N., Feldblum, S., Patel, N., & Tobin, A. (1991). Two genes encode distinct glutamate decarboxylases. *Neuron*, 7(1), 91-100. doi: 10.1016/0896-6273(91)90077-d

This paper describes the process demonstrated that GAD2 and GAD 1 are synthesised by different mRNAs and not from a single mRNA strand. This paper demonstrated this by using *Clustal Multiple Alignment Program* and aligned the rat GAD genes to the *Drosophila melanogaster* version of GAD, together using

immunochemistry techniques and the several visualization techniques to show that GAD2 and GAD1 are distinct.

- Faux, C., Rakic, S., Andrews, W., & Britto, J. (2012). Neurons on the Move: Migration and Lamination of Cortical Interneurons. *Neurosignals*, 20(3), 168-189. doi: 10.1159/000334489

This paper reviews and summarises findings associated with the migration and molecular cues in the migratory cues interneuron migration. The paper also illustrates how the disruption of early developmental events can result in a loss of GABAergic neurons.

- Fish, K., Rocco, B., & Lewis, D. (2018). Laminar Distribution of Subsets of GABAergic Axon Terminals in Human Prefrontal Cortex. *Frontiers In Neuroanatomy*, 12. doi: 10.3389/fnana.2018.00009

This journal article discusses the different types of γ-aminobutyric acid (GABA) neurons based on the presence of calbindin (CB), calretinin (CR) or parvalbumin (PV) which are all calcium-ion channels and the different types of glutamic acid decarboxylase 65 (GAD65) and glutamic acid decarboxylase 67 (GAD67) which are found in GABA neurons. The paper observed that PV terminals comprised two subsets based on the presence of only GAD67 or both GADs (GAD65 and GAD67) whereas CB and CR terminals comprised three subsets GAD65, GAD67, or GAD65 and GAD67. The densities of the different CB, CR and PV GAD terminal subpopulations also differed across layers.

- Friocourt, G., Kanatani, S., Tabata, H., Yozu, M., Takahashi, T., & Antypa, M. et al. (2008). Cell-Autonomous Roles of ARX in Cell Proliferation and Neuronal Migration during Corticogenesis. *Journal Of Neuroscience*, 28(22), 5794-5805. doi: 10.1523/jneurosci.1067-08.2008

This paper discusses the role of ARX. The paper highlights that ARX expression is not necessarily in the expression of the GABAergic phenotype.

- García-Bellido, A. (1979). GENETIC ANALYSIS OF THE ACHAETE-SCUTE SYSTEM OF DROSOPHILA MELANOGASTER. *Genetics*, 91(3), 491-520. doi: 10.1093/genetics/91.3.491

The original paper which describes the *achaete-scute* gene complex in *Drosophila Melanogaster*.

- García-Bellido, A., & de Celis, J. (2009). The Complex Tale of the achaete–scute Complex: A Paradigmatic Case in the Analysis of Gene Organization and Function During Development. *Genetics*, 182(3), 631-639. doi: 10.1534/genetics.109.104083

This paper explores the *achaete-scute* gene complex and how these genes play a key role in the differentiation of neural fate during development.

- Ghodsinia, A. (2020). *Tutorial: RNA-Seq Workflow with Galaxy | No Coding Involved (Step-by-Step)* [Video]. Retrieved from <https://www.youtube.com/watch?v=Kvh98S89yUU>

This video provided by Arman Ghodsinia shows the main workflow of analysing RNA-seq data in Galaxy

- Ghosh, I., Yao, S., & Chmielewski, J. (1999). DNA-binding Peptides. *Comprehensive Natural Products Chemistry*, 7, 477-490. doi: 10.1016/b978-0-08-091283-7.00069-2

This paper provides a overview on the various protein transcription factors which bind to DNA as well as providing examples of the various DNA binding proteins

- Godin, J., & Nguyen, L. (2013). Novel Functions of Core Cell Cycle Regulators in Neuronal Migration. *Advances In Experimental Medicine And Biology*, 59-74. doi: 10.1007/978-94-007-7687-6_4

This paper explores the cerebral cortex and the role of p27 Kip1 acts as a multifunctional protein that can control critical steps of neuronal migration through activities that go far beyond cell cycle regulation.

- Goecks, J., Nekrutenko, A., Taylor, J., & The Galaxy Team. (2010). Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biology*, 11(8). doi: 10.1186/gb-2010-11-8-r86

This paper talks about the rationale and the need of using computer-based approaches in the processing of next-generation sequencing data. This paper offers Galaxy, an open web-based platform for genomic research, which addresses these problems and provides tutorials and an open community to users to analyse data without the need for programming.

- Götz, M., & Huttner, W. (2005). The cell biology of neurogenesis. *Nature Reviews Molecular Cell Biology*, 6(10), 777-788. doi: 10.1038/nrm1739

This paper reports about neural stem cells that generate neurons by asymmetric and symmetric divisions and the development from neuroepithelial to radial glial cells.

- Green fluorescent protein CopGFP. Retrieved 1 May 2021, from https://evrogen.com/products/CopGFP/CopGFP_Detailed_description.shtml

This website reports on the Green fluorescent protein CopGFP extracted from copepod *Pontellina plumata*.

- Guillemot, F. (2007). Spatial and temporal specification of neural fates by transcription factor codes. *Development*, 134(21), 3771-3780. doi: 10.1242/dev.006379

This paper looks at the importance of homeodomain and basic helix-loop-helix families and their influence in the generation of neurons and glial cells.

- Hall, A., & Miller, R. (Eds.). (2012). Stem Cells in the Nervous System. *Basic Neurochemistry (Eighth Edition)*, 558-568. doi: 10.1016/b978-0-12-374947-5.00030-4

The chapter from the paper " Basic Neurochemistry" provides an overview of stem cells from their function and purpose to their therapeutic applications.

- Hall, A. (2012). Development of the Nervous System. *Basic Neurochemistry (Eighth Edition)*, 533-545. doi: 10.1016/b978-0-12-374947-5.00028-6

The chapter from the paper " Basic Neurochemistry" provides an overview of the development of the nervous system from development to neurogenesis to synapse formation.

- Harris, E., Abel, J., Tejada, L., & Rissman, E. (2016). Calbindin Knockout Alters Sex-Specific Regulation of Behavior and Gene Expression in Amygdala and Prefrontal Cortex. *Endocrinology*, 157(5), 1967-1979. doi: 10.1210/en.2016-1055

This paper looks at Calbindin (Calb1) and investigated novel aspects of calbindin function on social and anxiety behaviour in adult mice by comparing wild-type to Calb1 KO mice. The paper also explored the effects on γ-aminobutyric acid signalling through the study.

- Heberle, H., Meirelles, G., da Silva, F., Telles, G., & Minghim, R. (2015). InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinformatics*, 16(1). doi: 10.1186/s12859-015-0611-3

This paper reviews the importance of InteractiVenn as the method of bioinformatic analysis and describes the unique feature not found in other analysis tools.

- Helix-loop-helix DNA-binding domain superfamily. Retrieved 12 July 2020, from <https://www.ebi.ac.uk/interpro/entry/InterPro/IPR036638/>

This website provides a description on the helix-loop-helix DNA-binding domain.

- Hsieh, J. (2012). Orchestrating transcriptional control of adult neurogenesis. *Genes & Development*, 26(10), 1010-1021. doi: 10.1101/gad.187336.112

This journal looks at stem cells and the process called adult neurogenesis and explores the various transcription factors associated with mature neurons.

- Hwang, H., Ku, R., & Hashimoto-Torii, K. (2019). Prenatal Environment That Affects Neuronal Migration. *Frontiers In Cell And Developmental Biology*, 7. doi: 10.3389/fcell.2019.00138

This paper looks at the migration of neurons from development into infancy and investigates the role of external influences.

- Ide, M., Yamada, K., Toyota, T., Iwayama, Y., Ishitsuka, Y., & Minabe, Y. et al. (2005). Genetic association analyses of PHOX2B and ASCL1 in neuropsychiatric disorders: evidence for association of ASCL1 with Parkinson's disease. *Human Genetics*, 117(6), 520-527. doi: 10.1007/s00439-005-1342-8

This paper discusses the relationship between PHOX2B and ASCL1 and Parkinson's disease.

- Jones, S. (2004). An overview of the basic helix-loop-helix proteins. *Genome Biology*, 5(6), 226. doi: 10.1186/gb-2004-5-6-226

This paper provides an overview of basic helix-loop-helix proteins and their roles as regulators of embryonic development, which includes neurogenesis, myogenesis, heart development and hematopoiesis.

- Kala, K., Haugas, M., Lillevali, K., Guimera, J., Wurst, W., Salminen, M., & Partanen, J. (2009). Gata2 is a tissue-specific post-mitotic selector gene for midbrain GABAergic neurons. *Development*, 136(2), 253-262. doi: 10.1242/dev.029900

This paper discusses the role of GATA2 in the regulation of GABAergic neuron development in the mouse midbrain. This paper also explores the absence of GATA2 and its ability to switch from a GABAergic phenotype to an glutamatergic phenotype.

- Kelsom, C., & Lu, W. (2013). Development and specification of GABAergic cortical interneurons. *Cell & Bioscience*, 3(1), 3-19. doi: 10.1186/2045-3701-3-19

This paper reviews the transcriptional network of genes that play a role in the specification and maintenance of GABAergic interneuron fate.

- Kim, D., Langmead, B., & Salzberg, S. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nature Methods*, 12(4), 357-360. doi: 10.1038/nmeth.3317

This paper describes the function and the purpose of HISAT on RNA seq data experiments.

- Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., & Salzberg, S. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology*, 14(4). doi: 10.1186/gb-2013-14-4-r36

This paper discusses the function and purpose of TopHat2, and explores the new improvement from TopHat.

- Krueger, C., Stöcker, W., & Schlosser, M. (2007). Glutamic Acid Decarboxylase Autoantibodies. *Autoantibodies*, 369-378. doi: 10.1016/b978-044452763-9/50052-4

This paper investigates glutamic acid decarboxylase (GAD) and its involvement as an autoantigen in beta cell-specific autoimmunity diseases such as type 1 diabetes mellitus.

- Kumar, A., Pareek, V., Faiq, M., Kumar, P., Kumari, C., Singh, H., & Ghosh, S. (2020). Transcriptomic analysis of the signature of neurogenesis in human hippocampus suggests restricted progenitor cell progression post-childhood. *IBRO Reports*, 9, 224-232. doi: 10.1016/j.ibror.2020.08.003

This paper investigates adult hippocampal neurogenesis in humans. The study examines signature markers of neurogenesis and other related processes.

- Lamont, F., Tomlinson, D., Cooper, P., Shnyder, S., Chester, J., & Knowles, M. (2010). Small molecule FGF receptor inhibitors block FGFR-dependent urothelial carcinoma growth in vitro and in vivo. *British Journal Of Cancer*, 104(1), 75-82. doi: 10.1038/sj.bjc.6606016

This paper explores the use of FGF molecule inhibitors (PD173074, TKI-258 and SU5402) on bladder tumour cell lines with known FGFR expression levels and FGFR3. This resulted in all inhibitors preventing activation of FGFR3, and inhibited downstream MAPK pathway signalling. And the paper further proposes FGF inhibition as a method of therapeutic treatment.

- Le, T., Zhou, Q., Cobos, I., Zhang, S., Zagózewski, J., & Japoni, S. et al. (2017). GABAergic Interneuron Differentiation in the Basal Forebrain Is Mediated through Direct Regulation of Glutamic Acid

Decarboxylase Isoforms by Dlx Homeobox Transcription Factors. *The Journal Of Neuroscience*, 37(36), 8816-8829. doi: 10.1523/jneurosci.2125-16.2017

This paper demonstrates the role of both DLX1 and DLX2 homeoprotein in the activation of transcription of GAD genes.

- Liu, H., & Song, N. (2016). Molecular Mechanism of Adult Neurogenesis and its Association with Human Brain Diseases. *Journal Of Central Nervous System Disease*, 8, 5–11. doi: 10.4137/JCNSD.S32204

This paper discusses the molecular mechanism of adult neurogenesis and summarises the latest evidence regarding its role in brain diseases such as Alzheimer's disease, Parkinson's disease, Huntington's disease, and brain tumors.

- Lui, J., Hansen, D., & Kriegstein, A. (2011). Development and Evolution of the Human Neocortex. *Cell*, 146(1), 18-36. doi: 10.1016/j.cell.2011.06.030

This paper discusses how the proliferation of cells within the neocortex expands as the number of neurons increases. The paper also discusses related features to other mammalian species and focuses on the evolutionary significance.

- Liu, Y., Tsai, J., Chen, J., Yang, W., Chang, P., & Cheng, P. et al. (2017). Ascl1 promotes tangential migration and confines migratory routes by induction of Ephb2 in the telencephalon. *Scientific Reports*, 7(1), 1-2,5,10-13. doi: 10.1038/srep42895

This paper discusses the role of ASCL1 during development, specifically neurogenesis. The paper also identified Eph receptor B2 (Ephb2) as a direct target of Ascl1 and through knockdown of EphB2 disrupted the separation of the VZ/SVZ and IZ migratory routes.

- Long, Q., Luo, Q., Wang, K., Bates, A., & Shetty, A. (2017). Mash1-dependent Notch Signaling Pathway Regulates GABAergic Neuron-Like Differentiation from Bone Marrow-Derived Mesenchymal Stem Cells. *Aging And Disease*, 8(3), 301. doi: 10.14336/ad.2016.1018

This paper discuss the relationship between ASCL1 and NOTCH signaling and how it regulating the Regulates GABAergic neuron-Like differentiation from bone marrow-derived stem cells

- MacDonald, R., Pollack, J., Debiais-Thibaud, M., Heude, E., Coffin Talbot, J., & Ekker, M. (2013). The ascl1a and dlx genes have a regulatory role in the development of GABAergic interneurons in the zebrafish diencephalon. *Developmental Biology*, 381(1), 276-285. doi: 10.1016/j.ydbio.2013.05.025

This paper discusses ascl1a, Dlx and gad1b which are paralogues in the zebrafish brain, the paper demonstrated that ascl1a and Dlx genes in Zebrafish are involved in the differentiation process of GABAergic interneurons.

- Marín, O., Plump, A., Flames, N., Sánchez-Camacho, C., Tessier-Lavigne, M., & Rubenstein, J. (2003). Directional guidance of interneuron migration to the cerebral cortex relies on subcortical

Slit1/2-independent repulsion and cortical attraction. *Development*, 130(9), 1889-1901. doi: 10.1242/dev.00417

This paper looks at tangential migration from the basal telencephalon to the cortex is a highly directional process and explores the role of Slit1 and Slit2 in regulating the neuronal migration from the basal telencephalon.

- Marín, O. (2012). Interneuron dysfunction in psychiatric disorders. *Nature Reviews Neuroscience*, 13(2), 107-120. doi: 10.1038/nrn3155

This paper examines the neurological disorders, schizophrenia, autism and intellectual disabilities. Animal models demonstrate that the molecular basis of such disruption is linked to specific defects in the development and function of interneurons and showed that abnormal GABAergic function play a role in neurological disorders.

- Markers of neuronal precursor cells (NPCs) / neural stem cells (NSCs). Retrieved 6 May 2021, from <https://www.sanbio.nl/resources/blogs/blog-about-npcs-nscs/>

This website discuss what neuronal precursor cells (NPCs) / neural stem cells (NSCs) and the various examples of how neuronal precursor cells (NPCs) / neural stem cells (NSCs) are distinguished based on their protein expression, identified using Immunofluorescent analysis.

- Masoro, E., & Austad, S. (2011). *Handbook of the Biology of Aging* [Ebook] (7th ed.). London: Elsevier. Retrieved from <https://linkinghub.elsevier.com/retrieve/pii/C20090019679>

This book covers concepts related to ageing process and how it affects the human brain and the changes which are associated with it.

- Massari, M., & Murre, C. (2000). Helix-Loop-Helix Proteins: Regulators of Transcription in Eucaryotic Organisms. *Molecular And Cellular Biology*, 20(2), 429-440. doi: 10.1128/mcb.20.2.429-440.2000

This paper discusses the role and function of helix-loop-helix proteins and provides examples of cellular processes and demonstrates experimental examples to the role of HLH proteins.

- Matsui, T., & Hsieh, J. (2016). GABAergic Interneurons-in-a-Dish: High Five for Epilepsy. *Epilepsy Currents*, 16(3), 177-178. doi: 10.5698/1535-7511-16.3.177

This paper discusses the need to study epilepsy which is characterised by uncontrolled recurrent seizures and is often associated with an increase in excitatory circuits or a decrease in inhibitory circuits. The paper presents one approach to study epilepsy using GABAergic interneurons and describes the method to directly convert GABAergic interneurons from fibroblasts.

- Mikkonen, M., Soininen, H., Tapiola, T., Alafuzoff, I., & Miettinen, R. (1999). Hippocampal plasticity in Alzheimer's disease: changes in highly polysialylated NCAM immunoreactivity in the hippocampal formation. *European Journal Of Neuroscience*, 11(5), 1754-1764. doi: 10.1046/j.1460-9568.1999.00593.x

This article discusses neuronal marker, PSA-NCAM and its immunoreactivity in the hippocampal formation of Alzheimer's disease (AD) patients.

- Moffat, J., Ka, M., Jung, E., & Kim, W. (2015). Genes and brain malformations associated with abnormal neuron positioning. *Molecular Brain*, 8(1). doi: 10.1186/s13041-015-0164-4

This paper provides an overview of neural development and migration and relates to defects in neuronal positioning. This paper also discusses the recent progress in identifying genes and brain malformations associated with abnormal neuronal positioning during human brain development.

- Niklison-Chirou, M., Agostini, M., Amelio, I., & Melino, G. (2020). Regulation of Adult Neurogenesis in Mammalian Brain. *International Journal Of Molecular Sciences*, 21(14), 4869. doi: 10.3390/ijms21144869

This paper discusses adult neurogenesis and its role and its association to several human diseases, including cognitive impairment and neurodegenerative diseases.

- Noctor, S., Martínez-Cerdeño, V., Ivic, L., & Kriegstein, A. (2004). Cortical neurons arise in symmetric and asymmetric division zones and migrate through specific phases. *Nature Neuroscience*, 7(2), 136-144. doi: 10.1038/nn1172

This paper provides a new perspective of the cell division and migration of cortical neurons from the neuroepithelium of the embryonic forebrain into the adult cerebral cortex using time-lapse imaging of clonal cells in the rat cortex over several generations. And demonstrate that neurons are generated in two proliferative zones by distinct patterns of division. Neurons arise directly from radial glial cells in the ventricular zone (VZ) and indirectly from intermediate progenitor cells in the subventricular zone (SVZ).

- Pan, Z. (2012). Transcriptional control of Gad2. *Transcription*, 3(2), 68-72. doi: 10.4161/trns.19511

This paper reviews the regulatory mechanisms for Gad2 transcription and highlights the characteristics of TATA-less Gad2 promoters and regulation of Gad2 transcription by CREB.

- Paridaen, J., & Huttner, W. (2014). Neurogenesis during development of the vertebrate central nervous system. *EMBO Reports*, 15(4), 351-364. doi: 10.1002/embr.201438447

This paper discusses regulations of neurogenesis in the developing vertebrate central nervous system and explores the mechanisms of asymmetric and symmetrical cell division, transcriptional and epigenetic regulation, and signaling pathways.

- Poitras, L., Ghanem, N., Hatch, G., & Ekker, M. (2007). The proneural determinant MASH1 regulates forebrain Dlx1/2 expression through the I12b intergenic enhancer. *Development*, 134(9), 1755-1765. doi: 10.1242/dev.02845

This paper analyses of protein-DNA interactions provide evidence of a direct role for ASCL1 in Dlx1/2 regulation in the forebrain. This paper further discusses the role of DLX proteins in maintenance of their own expression and how they regulate other genes in the network.

- Quartu, M., Serra, M., Boi, M., Ibba, V., Melis, T., & Del Facco, M. (2008). Polysialylated-neural cell adhesion molecule (PSA-NCAM) in the human trigeminal ganglion and brainstem at prenatal and adult ages. *BMC Neuroscience*, 9(1). doi: 10.1186/1471-2202-9-108

The paper discusses the role of polysialylated neural cell adhesion molecules (PSA-NCAM) as a marker of developing and migrating neurons. PSA-NCAM persists in the mature normal brain in some regions. Using the immunohistochemical occurrence of PSA-NCAM in the human trigeminal ganglion (TG) and brainstem neuronal populations at prenatal and adult age.

- Quantification | RNA sequencing. Retrieved 28 November 2020, from <https://www.ebi.ac.uk/training-beta/online/courses/functional-genomics-ii-common-technologies-and-data-analysis-methods/rna-sequencing/performing-a-rna-seq-experiment/data-analysis/quanti-fication/>

This website provides information on qualification using RNA-seq data.

- Quiñones-Hinojosa, A., & Chaichana, K. (2007). The human subventricular zone: A source of new cells and a potential source of brain tumors. *Experimental Neurology*, 205(2), 313-324. doi: 10.1016/j.expneurol.2007.03.016

This paper discusses the subventricular zone as an area of stem-like cells which are capable of self-renewal and multipotentiality and provides a link to the development and synthesis of brain tumours.

- Raina, A., Mahajani, S., Bähr, M., & Kügler, S. (2019). Neuronal Trans-differentiation by Transcription Factors Ascl1 and Nurr1: Induction of a Dopaminergic Neurotransmitter Phenotype in Cortical GABAergic Neurons. *Molecular Neurobiology*, 57(1), 249-260. doi: 10.1007/s12035-019-01701-x

This paper discusses the transcription Factors Ascl1 and Nurr and their ability to Induce a Dopaminergic neurotransmitter phenotype in Cortical GABAergic Neurons.

- Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., & Vilo, J. (2019). g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Research*, 47(W1), W191-W198. doi: 10.1093/nar/gkz369

This paper discusses the function and use of g:Profiler as a method of identify GO terms and highlighting what genes are enriched.

- Read mapping or alignment | RNA sequencing. Retrieved 28 November 2020, from <https://www.ebi.ac.uk/training-beta/online/courses/functional-genomics-ii-common-technologies-and-data-analysis-methods/rna-sequencing/performing-a-rna-seq-experiment/data-analysis/read-mapping-or-alignment/>

This website discusses the purpose of read mapping or alignment into a reference genome in RNA-seq analysis.

- RNA-seq expression analysis hands-on tutorial: From FASTQ to differentially expressed genes. Retrieved 28 November 2020, from <https://research.csc.fi/rnaseq-tutorial>

This website discusses the workflow in processing the raw reads, checking the read quality and trimming the bad quality bases and how to align to a reference genome.

- Rogers, J. (1987). Calretinin: a gene for a novel calcium-binding protein expressed principally in neurons. *The Journal Of Cell Biology*, 105(3), 1343-1353. doi: 10.1083/jcb.105.3.1343

This paper discusses Calretinin and describes the evolutionary context of the novel gene. And explores its distribution of calretinin and calbindin mRNAs in chick tissues, mapped using RNA gel blots and in situ hybridisation.

- Schuurmans, C., Armant, O., Nieto, M., Stenman, J., Britz, O., & Klenin, N. et al. (2004). Sequential phases of cortical specification involve Neurogenin-dependent and -independent pathways. *The EMBO Journal*, 23(14), 2892-2902. doi: 10.1038/sj.emboj.7600278

The paper looks at the neocortical projection neurons and the role Ngn1 and Ngn2 in glutamatergic and GABAergic neuronal phenotype.

- Schwab, M. (Ed.). (2009). *Encyclopedia of cancer* [Ebook] (p. 90). Berlin, Heidelberg: Springer. Retrieved from <https://link.springer.com/referencework/10.1007/978-3-540-47648-1>

This book provides definitions on all of the key terms used in cancer research.

- Scott, J., Hjelmeland, A., Chinnaiyan, P., Anderson, A., & Basanta, D. (2014). Microenvironmental Variables Must Influence Intrinsic Phenotypic Parameters of Cancer Stem Cells to Affect Tumourigenicity. *Plos Computational Biology*, 10(1), e1003433. doi: 10.1371/journal.pcbi.1003433

This paper presents a mathematical/computational model of a tumour growing according to the canonical cancer stem-cell hypothesis with a simplified microenvironment. The paper concluded that future theoretical models of stem-cell driven tumours must include specific feedback from the microenvironment onto the individual cellular behavior.

- Seki, T. (2020). Understanding the Real State of Human Adult Hippocampal Neurogenesis From Studies of Rodents and Non-human Primates. *Frontiers In Neuroscience*, 14, 839. doi: 10.3389/fnins.2020.00839

This paper discusses adult hippocampal neurogenesis from the use and studies of rodents and Non-human primates in order to explore human adult hippocampal neurogenesis.

- Shahriyari, L., & Komarova, N. (2013). Symmetric vs. Asymmetric Stem Cell Divisions: An Adaptation against Cancer?. *Plos ONE*, 8(10), 2. doi: 10.1371/journal.pone.0076195

This paper proposes the idea that the symmetric stem cell divisions in mammals could be an adaptation which helps delay the onset of cancers.

- Socoli, G., Vezin, H., Ledolter, K., Kress, T., & Kurzbach, D. (2019). Conformational tuning of a DNA-bound transcription factor. *Nucleic Acids Research*, 47(10), 5429-5435. doi: 10.1093/nar/gkz291

This paper discusses the role of transcription factors involved in many cellular processes and demonstrates the function of the myc associated transcription factor X (MAX) through EPR, NMR, crystallographic analysis.

- Silver, J., Horn, K., Busch, S., & Yonkof, A. (2009). Axonal Regeneration: Role of the Extracellular Matrix and the Glial Scar. *Encyclopedia Of Neuroscience*, 1173-1180. doi: 10.1016/b978-008045046-9.00027-9

This paper presents an overview of axonal regeneration. The paper also discusses the role of the extracellular matrix and the glial scar, chondroitinase ABC, and other inhibitory ECM molecules.

- Slack, J. (2018). What is a stem cell?. *Wiley Interdisciplinary Reviews: Developmental Biology*, 7(5), e323. doi: 10.1002/wdev.323

This paper explores the modern definition of stem cells and explores the evidence in favor of a stochastic rather than an obligate asymmetric form of cell division.

- Stühmer, T., Anderson, S., Ekker, M., & Rubenstein, J. (2002). Ectopic expression of the *Dlx* genes induces glutamic acid decarboxylase and *Dlx* expression. *Development*, 129(1), 245-252. doi: 10.1242/dev.129.1.245

This paper discuss that ectopic expression of *Dlx2* and *Dlx5* induced the expression of glutamic acid decarboxylases (GADs), the enzymes that synthesize GABA and further displayed that *Dlx2* can induce *Dlx5* expression, and that *Dlx1*, *Dlx2* and *Dlx5* can induce expression from a *Dlx5/6-lacZ* enhancer/reporter construct.

- Stühmer, T., Puelles, L., Ekker, M., & L.R. Rubenstein, J. (2002). Expression from a *Dlx* Gene Enhancer Marks Adult Mouse Cortical GABAergic Neurons. *Cerebral Cortex*, 12(1), 75-85. doi: <https://doi.org/10.1093/cercor/12.1.75>

This paper analyses the expression pattern of a zebrafish *dlx4/6* enhancer reporter construct in embryonic transgenic mutant mice. It also demonstrated that the pattern of LacZ/β-galactosidase in cells that tangentially migrate from the ganglionic eminences to the cerebral cortex is identical to the subpallial markers *Dlx* and *GAD* genes.

- Szklarczyk, D., Gable, A., Nastou, K., Lyon, D., Kirsch, R., & Pyysalo, S. et al. (2020). The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Research*, 49(D1), D605-D612. doi: 10.1093/nar/gkaa1074

This journal discusses the purpose, function and utility of STRING in analysing protein-protein interactions.

- Taliano, R., Lu, S., Singh, K., Mangray, S., Tavares, R., & Noble, L. et al. (2013). Calretinin expression in high-grade invasive ductal carcinoma of the breast is associated with basal-like subtype and unfavorable prognosis. *Human Pathology*, 44(12), 2743-2750. doi: 10.1016/j.humpath.2013.07.021

This paper discusses Calretinin and its expression in breast carcinomas samples.

- Tang, B. (2017). The Potential of Targeting Brain Pathology with Ascl1/Mash1. *Cells*, 6(3), 26. doi: 10.3390/cells6030026

This paper discusses the role of ASCL1 in neuronal reprogramming neural progenitors and promotes neuronal differentiation by activating neuronal target genes. The potential development of using ASCL1 over-expression of Ascl1, alone or in combination with other factors as a treatment in brain injury and cancer.

- Tapscott, S. (2005). The circuitry of a master switch: Myod and the regulation of skeletal muscle gene transcription. *Development*, 132(12), 2685-2695. doi: 10.1242/dev.01874

This paper talks about the function of Myod as bHLH protein and its ability to induce the regulation of skeletal muscle gene transcription.

- Tong, M., Jun, T., Nie, Y., Hao, J., & Fan, D. (2019). The Role of the Slit/Robo Signaling Pathway. *Journal Of Cancer*, 10(12), 2694-2705. doi: 10.7150/jca.31877

This paper discuss the role of Slit/Robo signaling in angiogenesis neurogenesis, tumor cell migration, metastasis and various cell events and activities

- Touzot, A., Ruiz-Reig, N., Vitalis, T., & Studer, M. (2016). Molecular control of two novel migratory paths for CGE-derived interneurons in the developing mouse brain. *Development*, 143(10), 1753-1765. doi: 10.1242/dev.131102

This paper investigates GABAergic interneurons and their heterogeneity. The paper discovers two new CGE caudo-rostral migratory pathways; the lateral migratory stream and the medial migratory stream. Together with the caudal migratory stream, contribute to populate the neocortex, hippocampus and amygdala and propose the existence of spatial/temporal migratory paths in the subpallium in the adult brain.

- Tremblay, R., Lee, S., & Rudy, B. (2016). GABAergic Interneurons in the Neocortex: From Cellular Properties to Circuits. *Neuron*, 91(2), 260-292. doi: 10.1016/j.neuron.2016.06.033

This review investigates the diversity, classification, and properties of GABAergic inhibitory interneurons and the properties that distinguish cell types. And further discusses the morphology and the mechanisms by which GABAergic inhibition contributes to neural circuits operations.

- van Strien, M., van den Berge, S., & Hol, E. (2011). Migrating neuroblasts in the adult human brain: a stream reduced to a trickle. *Cell Research*, 21(11), 1523-1525. doi: 10.1038/cr.2011.101

This short paper discusses the breakthroughs in adult neurogenesis in rodents and humans and the identities of various cell types by their protein expression.

- Vasconcelos, F., & Castro, D. (2014). Transcriptional control of vertebrate neurogenesis by the proneural factor Ascl1. *Frontiers In Cellular Neuroscience*, 8. doi: 10.3389/fncel.2014.00412

This paper investigates the proneural transcription factor ASCL1 and the contrast in the oscillatory versus the sustained modes of expression, the functional interactions with other transcriptional networks and characterization of ASCL1 target genes.

- Waite, M., Skidmore, J., Billi, A., Martin, J., & Martin, D. (2011). GABAergic and glutamatergic identities of developing midbrain Pitx2 neurons. *Developmental Dynamics*, 240(2), 333-346. doi: 10.1002/dvdy.22532

This paper looks at Pitx2, and the relationship it has with GABAergic differentiation. The data suggest that PITX2 is present in regionally restricted subpopulations of midbrain neurons and may promote GABAergic and glutamatergic differentiation.

- Wang, B., Long, J., Flandin, P., Pla, R., Waclaw, R., Campbell, K., & Rubenstein, J. (2013). Loss of Gsx1 and Gsx2 function rescues distinct phenotypes in Dlx1/2 mutants. *Journal Of Comparative Neurology*, 521(7), 1561-1584. doi: 10.1002/cne.23242

This paper looks at the knockouts of Dlx1/2 and the effects on the subpallial differentiation. This used the knockouts of Dlx1/2 to investigate Gsx overexpression which contributes to the Dlx1/2 mutant phenotypes.

- Wang, B., Lufkin, T., & Rubenstein, J. (2011). Dlx6 regulates molecular properties of the striatum and central nucleus of the amygdala. *The Journal Of Comparative Neurology*, 519(12), 2320-2334. doi: 10.1002/cne.22618

This paper discusses the use of *Dlx6* RNA and β-galactosidase construct in prenatal telencephalic expression and to investigate the role in tangentially migrating interneurons.

- Wang, S., Zeng, X., Liu, Y., Liang, C., Zhang, H., & Liu, C. et al. (2012). Construction and characterization of a PDCD5 recombinant lentivirus vector and its expression in tumor cells. *Oncology Reports*, 28(1), 91-98. doi: 10.3892/or.2012.1756

The paper explores the function and mechanism of the programmed cell death 5 (PDCD5) gene through the use of PDCD5 recombinant lentiviral expression vector in tumour cells.

- Woodbury, M., & Ikezu, T. (2013). Fibroblast Growth Factor-2 Signaling in Neurogenesis and Neurodegeneration. *Journal Of Neuroimmune Pharmacology*, 9(2), 92-101. doi: 10.1007/s11481-013-9501-5

This paper looks at FGF2 and the role it has in adult neurogenesis in neural stem cells and neural progenitor cells and its importance in therapeutic development for neurodegenerative disorders.

- Wu, C., & Sun, D. (2014). GABA receptors in brain development, function, and injury. *Metabolic Brain Disease*, 30(2), 367-379. doi: 10.1007/s11011-014-9560-1

This review presents the role of the γ-aminobutyric acid (GABA) in both the developing and mature central nervous system. The aforementioned also covers the therapeutic role of GABA in various brain injuries.

- Yang, N., Chanda, S., Marro, S., Ng, Y., Janas, J., & Haag, D. et al. (2017). Generation of pure GABAergic neurons by transcription factor programming. *Nature Methods*, 14(6), 621-628. doi: 10.1038/nmeth.4291

This paper investigates methodologies to overcome problems faced by transforming differentiating pluripotent stem cells into neurons. The paper showed the transient expression of the transcription factors Ascl1 and Dlx2. Both factors induce the generation of GABAergic neurons from human PSCs which are fully functional.

- Yeh, M., Gonda, Y., Mommersteeg, M., Barber, M., Ypsilanti, A., & Hanashima, C. et al. (2014). Robo1 Modulates Proliferation and Neurogenesis in the Developing Neocortex. *Journal Of Neuroscience*, 34(16), 5717-5731. doi: 10.1523/jneurosci.4256-13.2014

This paper discusses the Robo family of receptors and the Slit proteins, which are expressed in the forebrain. They play a role in the migration of cortical interneurons and this study investigates their function in the development of pyramidal cells.

- Zhu, Y., Li, H., Zhou, L., Wu, J., & Rao, Y. (1999). Cellular and Molecular Guidance of GABAergic Neuronal Migration from an Extracortical Origin to the Neocortex. *Neuron*, 23(3), 473-485. doi: 10.1016/s0896-6273(00)80801-6

This paper talks about the migration of GABAergic inhibitory from interneurons to the lateral ganglionic eminence. The study used an explant assay to study GABAergic neuronal migration and found that the ventricular zone (VZ) of the LGE is repulsive to GABAergic neurons, due to the production of protein Slit which guides the direction of GABAergic neurons, and blockade of endogenous Slit signalling inhibits the repulsive activity in the VZ.

9. Acknowledgements

I would like to thank Dr Frank Schubert for his advice and support throughout the duration of this project.

10. Appendix

I. The list of references genomes used the experiment

Human Release 37 (GRCh38.p13):

ftp://ftp.ebi.ac.uk/pub/databases/gencode/Gencode_human/release_37/gencode.v37.annotation.gtf.gz

II. The datasets used the experiment

E-MTAB-4840 - RNA-seq of coding RNA: Human Developmental Biology Resource (HDBR) expression resource of prenatal human brain development:

<https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-4840/samples/>

Telencephalon Left Sample 1 Source HDBR251 Development stage - Carnegie Stage 23:

ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR147/000/ERR1473520/ERR1473520_1.fastq.gz

Telencephalon Left Sample 2 Source HDBR251 Development stage - Carnegie Stage 23:

ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR147/000/ERR1473520/ERR1473520_2.fastq.gz

Telencephalon Right Sample 3 Source HDBR253 Development stage - Carnegie Stage 23:

ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR147/001/ERR1473521/ERR1473521_1.fastq.gz

Telencephalon Right Sample 4 Source HDBR253 Development stage - Carnegie Stage 23:

ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR147/001/ERR1473521/ERR1473521_2.fastq.gz

Telencephalon Right Sample 5 Source HDBR256 Development stage- Carnegie Stage 23:

ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR147/002/ERR1473522/ERR1473522_1.fastq.gz

Telencephalon Right Sample 6 Source HDBR256 Development stage - Carnegie Stage 23:

ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR147/002/ERR1473522/ERR1473522_2.fastq.gz

Telencephalon Left Sample 7 Source HDBR259 Development stage - Carnegie Stage 23:

ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR147/003/ERR1473523/ERR1473523_1.fastq.gz

Telencephalon Left Sample 8 Source HDBR259 Development stage - Carnegie Stage 23:

ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR147/003/ERR1473523/ERR1473523_2.fastq.gz

Cerebral Cortex Left Sample 1 Source HDBR265 Development stage - Carnegie Stage 23:

ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR147/001/ERR1473151/ERR1473151_1.fastq.gz

Cerebral Cortex Left Sample 2 Source HDBR265 Development stage - Carnegie Stage 23:

ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR147/001/ERR1473151/ERR1473151_2.fastq.gz

Cerebral Cortex Left Sample 3 Source HDBR283 Development stage - Carnegie Stage 23:

ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR147/004/ERR1473154/ERR1473154_1.fastq.gz

Cerebral Cortex Left Sample 4 Source HDBR283 Development stage - Carnegie Stage 23:

ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR147/004/ERR1473154/ERR1473154_2.fastq.gz

Cerebral Cortex Right Sample 5 Source HDBR328 Development stage - 9 post conception weeks cerebral cortex: ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR147/003/ERR1473183/ERR1473183_1.fastq.gz

Cerebral Cortex Right Sample 6 Source HDBR328 Development stage - 9 post conception weeks cerebral cortex: ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR147/003/ERR1473183/ERR1473183_2.fastq.gz

Cerebral Cortex Right Sample 7 Source HDBR333 Development stage - 9 post conception weeks cerebral cortex: ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR147/007/ERR1473237/ERR1473237_1.fastq.gz

Cerebral Cortex Right Sample 8 Source HDBR333 Development stage - 9 post conception weeks cerebral cortex: ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR147/007/ERR1473237/ERR1473237_2.fastq.gz

Series GSE31635 - Isolation of a Novel Rat Neural Progenitor Clone that Expresses Dlx Family Transcription Factors and Gives Rise to Functional GABAergic Neurons in Culture:
<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE31635>

Series GSE29985 - Identification by ChIP-on-Chip of ARX target genes, a transcription factor implicated in mental retardation and epilepsy: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE29985>

Series GSE78949 - Induction of cells expressing markers of GABAergic neurons by transcription factors (mRNA): <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE78949>

Series GSE46791 - An epigenetic signature of developmental potential in neural stem cells and early neurons [expression]: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE46791>

III. Table used for GO enrichment analysis for molecular function

Source	Term Name	Term ID	Adjusted P Value	Negative log10 of adjusted P Value	Term Size	Query Size	Intersection Size	Effective Domain Size	Intersections
GO:MF	calcium ion binding involved in regulation of cytosolic calcium ion concentration	GO:0099510	0.000572333	3.242351249	2	20	2		18694 CALB1,CALB2
GO:MF	calcium ion binding involved in regulation of presynaptic cytosolic calcium ion concentration	GO:0099534	0.000572333	3.242351249	2	20	2		18694 CALB1,CALB2
GO:MF	identical protein binding	GO:0042802	0.028050833	1.55205424	2080	184	39		18694 DBI,UGDH,GOT2
GO:MF	calmodulin-dependent cyclic-nucleotide phosphodiesterase activity	GO:0004117	0.029196851	1.534663982	3	81	2		18694 PDE1A,PDE1C
GO:MF	calcium- and calmodulin-regulated 3',5'-cyclic-GMP phosphodiesterase activity	GO:0048101	0.029196851	1.534663982	3	81	2		18694 PDE1A,PDE1C
GO:MF	ion channel inhibitor activity	GO:0008200	0.040453705	1.393041693	39	109	4		18694 WNK1,CAV1,CAMK2D,PACSN3
GO:MF	channel inhibitor activity	GO:0016248	0.044747587	1.349230375	40	109	4		18694 WNK1,CAV1,CAMK2D,PACSN3

IV. Table used for GO enrichment analysis for biological process

Source	Term Name	Term ID	Adjusted P Value	Negative log10 of adjusted P Value	Term Size	Query Size	Intersection Size	Effective Domain Size	Intersections
GO:BP	blood circulation	GO:0008015	0.001606678	2.794071123	562	133	17	18092	HDAC4,POSTN,WNK1,SLT2,PDE4B,ACTC1,CAV1,CACNB2,ENPEP,PRCP,CASQ1,CAMK2D,TH,KCNH2,K,CNP3,ATP1A1,MKKS
GO:BP	regulation of localization	GO:0032879	0.002306425	2.637060751	2903	134	44	18092	HDAC4,POSTN,WNK1,SNCA,SLT2,VP54B,MTUS1,LDB2,PDE4B,GRIA4,NUP54,PARD3,CAV1,CLL20,ACN2B,LMNA,YRDC,ZDHHC2,GHRHR,PRCP,PKN2,CDK5R2,NTNG1,PIK3R3,CASQ1,NPY5R,PLXNA2,NPVF,CAMK2D,COL3A1,RIM51,PASCIN3,KCNH2,UCP2,SCARB1,NOL3,EFNA1,KCNIP3,MKD,FABP3,TRADD,A,TP1A1,MKKS,FGFR1
GO:BP	regulation of cellular component movement	GO:0051270	0.003249738	2.488151714	1156	173	29	18092	HDAC4,POSTN,WNK1,SLT2,MTUS1,LDB2,PDE4B,CAV1,CLL20,LMNA,PRCP,PKN2,NTNG1,PIK3R3,PLXNA2,CAMK2D,COL3A1,SCARB1,EFNA1,MKD,TRADD,ATP1A1,MKKS,FGFR1,RET,CD151,NBL1,ITGA6,GA,DK,BMF,FGFR1,RET,SURF4,TSC1,FHD1,MAP2K7,IGF2,CD151,ITGA6,F1S1,GATA2,ACTR3B,NCK2
GO:BP	positive regulation of cellular component organization	GO:0051130	0.007202153	2.142537672	1182	176	29	18092	HDAC4,SNCA,SLT2,VPS4B,GPMP6A,TPPP3,CAV1,LMNA,CUX1,ZDHHC2,ARPC3,IRRN1,PLXNA2,RPS3,M,DK,BMF,FGFR1,RET,SURF4,TSC1,FHD1,MAP2K7,IGF2,CD151,ITGA6,F1S1,GATA2,ACTR3B,NCK2
GO:BP	regulation of system process	GO:0044057	0.008736688	2.058653152	632	134	17	18092	HDAC4,WNK1,PDE4B,PARD3,CAV1,CACNB2,LMNA,CASQ1,NPVF,CAMK2D,TH,RIM51,KCNH2,NOL3,K,CNP3,ATP1A1,FGFR1
GO:BP	circulatory system process	GO:0003013	0.010536125	1.977319068	646	133	17	18092	HDAC4,POSTN,WNK1,SNCA,SLT2,MTUS1,LDB2,CAV1,CLL20,LMNA,PRCP,PKN2,NTNG1,PIK3R3,PLXNA2,CAMK2D,COL3A1,SCARB1,EFNA1,MKD,TRADD,MKKS,FGFR1,RET,CD151,NBL1,ITGA6,GATA2
GO:BP	regulation of locomotion	GO:0040012	0.01572787	1.803330097	1118	173	27	18092	HDAC4,POSTN,WNK1,SNCA,SLT2,MTUS1,LDB2,CAV1,CLL20,LMNA,PRCP,PKN2,NTNG1,PIK3R3,PLXNA2,CAMK2D,COL3A1,SCARB1,EFNA1,MKD,TRADD,MKKS,FGFR1,RET,CD151,NBL1,ITGA6,GATA2
GO:BP	striated muscle contraction	GO:0006941	0.015875978	1.799259515	183	132	9	18092	HDAC4,PDE4B,ACTC1,CAV1,CACNB2,CASQ1,CAMK2D,KCNH2,ATP1A1
GO:BP	cardiac muscle contraction	GO:0060048	0.021260532	1.672425864	144	132	8	18092	HDAC4,PDE4B,ACTC1,CAV1,CACNB2,CASQ1,CAMK2D,TH,KCNH2,KCNIP3,ATP1A1
GO:BP	heart contraction	GO:0060047	0.021675228	1.664036316	295	132	11	18092	HDAC4,PDE4B,ACTC1,CAV1,CACNB2,CASQ1,CAMK2D,KCNH2,ATP1A1
GO:BP	regulation of cell motility	GO:2000145	0.023442638	1.629993524	1076	173	26	18092	HDAC4,POSTN,WNK1,SLT2,MTUS1,LDB2,CAV1,CLL20,LMNA,PRCP,PKN2,NTNG1,PIK3R3,PLXNA2,COL3A1,SCARB1,EFNA1,MKD,TRADD,MKKS,FGFR1,RET,CD151,NBL1,ITGA6,GATA2
GO:BP	regulation of presynaptic cytosolic calcium ion concentration	GO:0099509	0.028952387	1.538315627	16	55	3	18092	CALB1,CALB2,CACNB2
GO:BP	heart process	GO:0003015	0.029594194	1.528793481	305	132	11	18092	HDAC4,PDE4B,ACTC1,CAV1,CACNB2,CASQ1,CAMK2D,TH,KCNH2,KCNIP3,ATP1A1
GO:BP	cardiac muscle cell contraction	GO:0086003	0.038723822	1.412021789	76	132	6	18092	PDE4B,CAV1,CACNB2,CAMK2D,KCNH2,ATP1A1 POSTN,WNK1,OLFM3,SLT2,GPMP6A,IRX3,PARD3,PTX2,CUX1,OPCML,CDK5R2,NTNG1,PLXNA2,NEUR,OD6,TH,COL3A1,RIM51,EFNA1,MOK,HDAC11,FGFR1,CEND1,RET,TSC1,OMP,NOTCH2,NCDN,RRN1L,OLFM1,CTDSP1,NBL1,ITGA6,GATA2,NCK2,TSKU,UGDH
GO:BP	neurogenesis	GO:0022008	0.039717167	1.401021741	1722	183	36	18092	POSTN,WNK1,OLFM3,SLT2,GPMP6A,IRX3,PARD3,PTX2,CUX1,OPCML,NTNG1,PLXNA2,NEUR,OD6,TH,COL3A1,RIM51,EFNA1,MOK,HDAC11,FGFR1,CEND1,RET,TSC1,NOTCH2,NCDN,RRN1L,OLFM1,CTDSP1,NBL1,ITGA6,GATA2,A2,NCK2,TSKU,UGDH
GO:BP	neuron differentiation	GO:0030182	0.040767293	1.389688122	1445	183	32	18092	HDAC4,PDE4B,CAV1,CACNB2,CASQ1,CAMK2D,TH,KCNH2,KCNIP3,ATP1A1
GO:BP	regulation of heart contraction	GO:0008016	0.041039868	1.386794045	259	132	10	18092	HDAC4,PDE4B,CAV1,CACNB2,CASQ1,CAMK2D,TH,KCNH2,KCNIP3,ATP1A1
GO:BP	regulation of apoptotic signaling pathway	GO:2001233	0.041107397	1.386080026	378	176	14	18092	IVNS1ABP,GNAA1,USP42,CAV1,LMNA,RPS3,NOL3,TRADD,BMF,FGFR1,RET,ITGA6,F1S1,NCK2 POSTN,WNK1,GTPE2I,OLFM3,CALB1,SLT2,GPMP6A,IRX3,MAB21L2,PALMD,COL11A1,FREM2,ACTC1,PA,RD3,PTX2,CAV1,ENPEP,CUX1,PRCP,DUSP2,COL9A1,CDK5R2,NTNG1,PIK3R3,CASQ1,NPY5R,GAS2,PLXNA2,TH,COL3A1,RIM51,EFNA1,MOK,SERPINH1,MKKS,FGFR1,CEND1,RET,TSC1,IGF2,NOTCH2,RRN1L,OLFM1,CD151,NBL1,S100A13,ITGA6,F1S1,GATA2
GO:BP	anatomical structure morphogenesis	GO:0009653	0.042358977	1.373054532	2865	173	49	18092	IVNS1ABP,GNAA1,USP42,CAV1,LMNA,RPS3,NOL3,TRADD,BMF,FGFR1,RET,ITGA6,F1S1,NCK2 POSTN,WNK1,GTPE2I,OLFM3,CALB1,SLT2,GPMP6A,IRX3,MAB21L2,PALMD,COL11A1,FREM2,ACTC1,PA,RD3,PTX2,CAV1,ENPEP,CUX1,PRCP,DUSP2,COL9A1,CDK5R2,NTNG1,PIK3R3,CASQ1,NPY5R,GAS2,PLXNA2,TH,COL3A1,RIM51,EFNA1,MOK,SERPINH1,MKKS,FGFR1,CEND1,RET,TSC1,IGF2,NOTCH2,RRN1L,OLFM1,CD151,NBL1,S100A13,ITGA6,F1S1,GATA2

V. Table used for GO enrichment analysis for cellular component

Source	Term Name	Term ID	Adjusted P Value	Negative log10 of adjusted P Value	Term Size	Query Size	Intersection Size	Effective Domain Size	Intersections
GO:CC	dendritic spine	GO:0043197	0.000648607	3.188018335	185	39	6	18963	CALB1,CTTNBP2,GPM6A,PDE4B,GRIA4,PALMD
GO:CC	neuron spine	GO:0044309	0.000712068	3.147478759	188	39	6	18963	CALB1,CTTNBP2,GPM6A,PDE4B,GRIA4,PALMD
GO:CC	cytoplasm	GO:0005737	0.002152827	2.666990783	11888	185	144	18963	BET1,ASZ1,HDAC4,POSTN,WNK1,GTF21,OLFM3,SCFD2,CALB1,TRDMT1,MRPL48,SNCAC,CTTNBP2,IVNS1ABP,SLT2,VP54B,CALB2,HSPA4L,PDCL,RX3,GSTK1,LRBA,MTUS1,PDE1A,CLCC1,PDE4B,MAB21L2,SCOC,COMMDS8,SV2B,NDUFAS5,GRIA4,PALMD,PRSS23,COL11A1,GNA13,PIGY,HOOK1,ACTC1,CRYM,T,PRCP,DUSP2,NENF,HMCN1,PKN2,COLQ1,ARP3,CDK5R2,RASGRP1,PDE1C,PIK3R3,PPP2CB,CASQ1,PRIP2,POU2F1,NPY5R,NOL6,GAS2,CHTS10,CAMK2D,PPPAR2,TH,COL3A1,SEC61A2,SMOX,PLEKH5,PPFBP2,RIMS1,YTHDF3,RPS3,QTTR1,PACsin3,KCNH2,SURP1,UCP2,SCAR81,NOL3,KCNIP3,ODF2,FABP3,PPOX,SERF2,SERPINH1,TRADD,DHDOS,BMF,NSUN4,DP7,ATP1A1,MKKS,FGFR1,FBX19,HSFBB,CEND1,RET,DEDD,ATP11B,SURF4,C03EAP,TSCLFHOD1,MAP2K7,IGF2,OMP,NOTCH2,NCND,GTBP6,AMPD2,OLFM1,RLP35,QDPR,CD151,GAD2,PRKA81,LRC23,S100A13,RESP18,FS11,DIDO1,GATA2,ACTR3B,DBI,NCK2,LRPA1,MRT04,SLC25A22,ETFB,UGDH,GTG2,CDH15
GO:CC	synapse	GO:0045202	0.002503926	2.601378495	1382	107	22	18963	HDAC4,POSTN,OLFM3,CALB1,SNCAC,CTTNBP2,CALB2,GPM6A,PDE4B,SV2B,GRIA4,PALMD,ACTC1,CA,CNB2,ZDHHC2,NTNG1,NPY5R,TH,PLEKH5,PPFBP2,RIMS1,RPS3
GO:CC	actin cytoskeleton	GO:0015629	0.003490912	2.457061064	513	20	6	18963	HDAC4,CALB1,SNCAC,CTTNBP2,LSP1,CALB2,HDAC4,POSTN,OLFM3,CALB1,SNCAC,CTTNBP2,CALB2,GPM6A,PDE4B,SV2B,GRIA4,PALMD,ACTC1,PA,PKN2,ARP3,CDK5R2,PDE1C,NPY5R,CAMK2D,TH,RPS3,SCAR81,KCNIP3,ODF2,ATP1A1,MKKS,CDH15,ATP1A1,TSCLFHOD1,OMP,NOTCH2,NCND,NNR11,OLFM1,CD151,GAD2,ITGA6
GO:CC	cell junction	GO:0030054	0.006822943	2.166028285	2131	169	38	18963	HDAC4,CALB1,SNCAC,CTTNBP2,CALB2,GPM6A,PDE4B,SV2B,GRIA4,PALMD,ACTC1,PA,PKN2,ARP3,CDK5R2,PDE1C,NPY5R,CAMK2D,TH,RPS3,SCAR81,KCNIP3,ODF2,ATP1A1,MKKS,CDH15,ATP1A1,TSCLFHOD1,OMP,NOTCH2,NCND,NNR11,OLFM1,CD151,GAD2,ITGA6
GO:CC	presynapse	GO:0098793	0.009175714	2.037360149	543	116	13	18963	CALB1,SNCAC,CTTNBP2,CALB2,GPM6A,PDE4B,SV2B,CACNB2,NTNG1,TH,PPFBP2,RIMS1,KCNIP3
GO:CC	terminal bouton	GO:0043195	0.009748887	2.011044982	58	20	3	18963	CALB1,SNCAC,CALB2
GO:CC	plasma membrane bounded cell projection	GO:0120025	0.010353215	1.984924749	2255	169	39	18963	CALB1,SNCAC,CTTNBP2,CALB2,GPM6A,PDCL,RX3,PDE4B,SV2B,GRIA4,PALMD,ACTC1,PAT6,PKN2,ARP3,CDK5R2,PDE1C,NPY5R,CAMK2D,TH,RPS3,SCAR81,KCNIP3,ODF2,ATP1A1,MKKS,CDH15,ATP1A1,TSCLFHOD1,OMP,NOTCH2,NCND,NNR11,OLFM1,CD151,GAD2,ITGA6
GO:CC	cell projection	GO:0042995	0.01051443	1.978214251	2352	174	41	18963	CALB1,SNCAC,CTTNBP2,CALB2,GPM6A,PDCL,RX3,PDE4B,SV2B,GRIA4,PALMD,ACTC1,PAT6,PKN2,ARP3,CDK5R2,PDE1C,NPY5R,CAMK2D,TH,RMS1,RP53,SCAR81,KCNIP3,ODF2,ATP1A1,MKKS,CDH15,ATP1A1,TSCLFHOD1,OMP,NOTCH2,NCND,NNR11,OLFM1,CD151,GAD2,ITGA6
GO:CC	cuticular plate	GO:0032437	0.012377961	1.907305884	9	20	2	18963	CALB1,CALB2
GO:CC	endoplasmic reticulum	GO:0005783	0.02084487	1.681000804	1975	177	36	18963	BET1,SNCAC,LRBA,CLCC1,PRSS23,COL11A1,GNA13,PIGY,CAV1,ZDHHC2,RPL21,AGPAT5,NENF,COL9A1,RASGRP1,CASQ1,POU2F1,CAMK2D,TH,COL3A1,SEC61A2,RPS3,NOL3,KCNIP3,ODF2,ATP1A1,TSCLFHOD1,OMP,NOTCH2,NCND,NNR11,OLFM1,CD151,GAD2,ITGA6
GO:CC	somatodendritic compartment	GO:0036477	0.021135452	1.674988458	893	39	9	18963	CALB1,SNCAC,CTTNBP2,CALB2,GPM6A,PDE1A,PDE4B,GRIA4,PALMD
GO:CC	distal axon	GO:0150034	0.024462444	1.611500156	318	158	11	18963	CALB1,SNCAC,CALB2,GPM6A,PARD3,ARP3,CDK5R2,TH,KCNIP3,TSCLFHOD1,OMP,NOTCH2,NCND,NNR11,OLFM1,CD151,GAD2,ITGA6
GO:CC	neuromuscular junction	GO:0031594	0.032576183	1.4870998	78	4	2	18963	HDAC4,POSTN
GO:CC	cytosol	GO:0005829	0.034405345	1.463374084	5231	168	69	18963	HDAC4,WINK1,CALB1,SNCAC,IVNS1ABP,VPS4B,CALB2,HSPA4L,PDCL,GSTK1,LRBA,PDE1A,PDE4B,SCOC,COMMDS8,HOOK1,ACTC1,CRYM,PARD3,USP47,LMNA,CUX1,RPL21,ZFAND6,PKN2,ARP3,RASGRP1,PDE1C,PIK3R3,PPP2CB,RPK2,GAS2,CAMK2D,TH,SEC61A2,SMOX,PLEKH5,PPFBP2,RIMS1,YTHDF3,RPS3,PACsin3,NOL3,KCNIP3,ODF2,FABP3,SERF2,TRADD,BMF,DP7,MKKS,FGFR1,FBX19,HSFBB,SRF4,C03EAP,TSCLFHOD1,MAP2K7,OMP,NCND,AMPD2,RPL35,CDPR,CD151,GAD2,PRKAB1,LRRK23,S100A13
GO:CC	organelle	GO:0043226	0.038115586	1.418897403	13974	53	50	18963	BET1,ASZ1,HDAC4,POSTN,GTf21,OLFM3,SCFD2,CALB1,TRDMT1,MRPL48,SNCAC,CTTNBP2,IVNS1ABP,SLT2,LSP1,VP54B,CALB2,HSPA4L,ELAVL2,PDCL,IRX3,GSTK1,LRBA,MTUS1,LDB2,PDE1A,CLCC1,PDE4B,MAB21L2,SCOC,COMMDS8,SV2B,NDUFAS5,GRIA4,PRSS23,COL11A1,FREM2,GNA13,PIGY,HOOK1,NUP54,ACTC1,CRYM,TPPP3,PAR3,USP47,PTX2,CAV1

VI. Table of the gene overlaps between datasets: GSE29985, GSE31635 and GSE46791

[GSE46791] and [GSE29985] and [GSE31635]: 195 Genes	Actc1	Chst10	Cd3eap
	Crym	Camk2d	Tsc1
Bet1	Tppp3	Ppp4r2	Fhod1
Pot1a	Pard3	Neurod6	Map2k7
Asz1	Usp47	Th	Ccdc97
Hdac4	Pitx2	Col3a1	Igf2
Postn	Cav1	Sec61a2	Omp
Wnk1	Ccl20	Smox	Notch2
Gtf2i	Cacnb2	Plekha5	Ncdn
Olfm3	Enpep	Ppfibp2	Gtpbp6
Scfd2	Lmna	Rims1	
Calb1	Pldn	Ythdf3	Plekha6
Trdmt1	Yrdc	Rps3	Ampd2
Mrpl48	Slc19a2	Qtrt1	Surf2
Snca	Cux1	Pacsin3	Nrn1l
Gatad1	Zdhhc2	Kcnh2	Prkcdbp
Plxdc2	Rpl21	Surf1	Zfp84
Ctnbp2	Baz2b	Ucp2	Olfm1
lvns1abp	Ghrhr	Scarb1	Accn4
Slit2	Agpat5	Nol3	Tmem141
Lsp1	Zfand6	Efna1	Ctdsp1
Vps4b	Spata6	Kcnip3	Rpl35
Calb2	Galnt5	Odf2	Qdpr
Hspa4l	Prcp	Mdk	Cd151
Gpm6a	Dusp2	Fabp3	Nbl1
Elavl2	Nenf	Ttyh3	Gad2
Pdcl	Hmcn1	Ppox	
Irx3	Zfhx4	Serf2	Prkab1
Gstk1	Opcml	Serpinh1	Lrrc23
Lrba	Zranb2	Gper	S100a13
Mtus1	Pkn2	Tradd	Itga6
Ldb2	Col9a1	Ufc1	Resp18
Pde1a	Arpc3	Dhdds	Fis1
Clcc1	Cdk5r2	Hdac11	Dido1
Pde4b	Rasgrp1	Cldn15	Gata2
Mab21l2	Pde1c	Bmf	Actr3b
Scoc	Ntng1	Nsun4	Dbi
Commd8	Pik3r3	Dpp7	Nck2
Sv2b	Ppp2cb	Atp1a1	Lrpap1
Ndufa5	Cugbp2	Mkks	Tsku
Gria4	Casq1	Fgfr1	Gpatch1
Palmd	Ripk2	Fbxl19	Mrto4
Pxmp3	Lrrn1	Hspb8	Slc25a22
Prss23	Cryga	Ppm1j	
Col11a1	Pou2f1	Cend1	Mesdc2
Frem2	Npy5r	Ret	Etfb
Gnai3	Nol6	Dedd	Ugdh
Pigy	Gas2	Emp3	Got2
Hook1	Plxna2	Atp11b	Cdh15
Nup54	Npvf	Surf4	Zfp238

VII. The processed excel gene datasets

GSE29985 Identification by ChIP-on-Chip of ARX target genes, a transcription factor implicated in mental retardation and epilepsy:

<https://drive.google.com/file/d/1AaK7leSAadEqI9SR3VKerfOuo0BtkGT5/view?usp=sharing>

GSE31635 Isolation of a Novel Rat Neural Progenitor Clone that Expresses Dlx Family Transcription Factors and Gives Rise to Functional GABAergic Neurons in Culture:

<https://drive.google.com/file/d/1GiD7bzvXEb49p326WbcdwJKwooWytyFR/view?usp=sharing>

E-MTAB-4840 RNA-seq of coding RNA: Human Developmental Biology Resource (HDBR) expression resource of prenatal human brain development:

<https://drive.google.com/file/d/1bJSjxhLaxWi835zF2TDwB8TzitB2ZKlg/view?usp=sharing>

GSE78949 Induction of cells expressing markers of GABAergic neurons by transcription factors (mRNA):

<https://drive.google.com/file/d/1cOAx5UTYOOOnAg74M-buaKpw88VF58arE/view?usp=sharing>

GSE46791 An epigenetic signature of developmental potential in neural stem cells and early neurons [expression]: https://drive.google.com/file/d/1dZ1pDrqtApCo7qhTxCVmi_eleVUrkfdD/view?usp=sharing

11. Total number of words

Total number of words: 5,008

12. Signed declaration

- I hereby declare that this written report is substantially my own work;
- I do consent to my written report in this attributed format (not anonymous), subject to final approval by the Board of Examiners, being made available electronically in the Library Dissertation Repository and/or Department/School/Subject Group digital repositories. Such reports will normally be kept for a maximum of ten years;
- I understand that if I consent, this written report will be accessible only to staff and students for reference only;
- This permission may be revoked at any time by emailing data-protection@port.ac.uk

I declare that the work submitted here is my own. Any material used from other sources (e.g. data, ideas, images) is clearly identified as such with the author(s) of the material cited.

Date: 21/05/2021