Expression of genes in the 16p11.2 locus during development of the human fetal cerebral cortex

Expression of genes in the *16p11.2* locus during development of the human fetal cerebral cortex.

Sarah Morson[1,2], Yifei Yang[1,2], David J. Price[1,2], Thomas Pratt[1,2,3]

[1]Simons Initiative for the Developing Brain,

[2]Centre for Discovery Brain Sciences,

Hugh Robson Building, Edinburgh Medical School Biomedical Sciences, The University of Edinburgh, Edinburgh, United Kingdom, EH8 9XD, United Kingdom.

[3]Correspondence:

Thomas Pratt

t.pratt@ed.ac.uk

Tel: +44(0)1316503732

Running Title: 16p11.2 gene expression during human corticogenesis

Abstract

The 593 kbp *16p11.2* copy number variation (CNV) affects the gene dosage of 29 protein coding genes, with heterozygous *16p11.2* microduplication or microdeletion implicated in about 1% of autism spectrum disorder (ASD) cases. The *16p11.2* CNV is frequently associated with macrocephaly or microcephaly indicating early defects of neurogenesis may contribute to subsequent ASD symptoms, but it is unknown which *16p11.2* transcripts are expressed in progenitors and whose levels are likely, therefore, to influence neurogenesis. Analysis of human fetal gene expression data revealed that *KIF22*, *ALDOA*, *HIRIP3*, *PAGR1*, and *MAZ* transcripts are expressed in neural progenitors with *ALDOA* and *KIF22* significantly enriched compared to post-mitotic cells. To investigate the possible roles of *ALDOA* and *KIF22* proteins in human cerebral cortex development we used immunohistochemical staining to describe their expression in late first and early second trimester human cerebral cortex. KIF22 protein is restricted to proliferating cells with its levels increasing during the cell cycle and peaking at mitosis. ALDOA protein is expressed in all cell types and does not vary with cell-cycle phase. Our expression analysis suggests the hypothesis that altered neurogenesis in the cerebral cortex contributes to ASD in *16p11.2* CNV patients.

Keywords:

ALDOA, Autism, Cerebral Cortex, CNV, KIF22

Introduction.

Large, recurrent Copy Number Variations (CNVs) are implicated in many neuropsychiatric disorders including autism spectrum disorders (ASD), epilepsy, intellectual disability (ID) and schizophrenia (McCarthy *et al.*, 2009; Girirajan and Eichler, 2010; Levy *et al.*, 2011; Sanders *et al.*, 2011; Malhotra and Sebat, 2012). The *16p11.2* CNV (OMIM 611913) encompasses a 593 kb DNA sequence in the p11.2 region of human chromosome 16 (BP4-BP5). This region harbors 29 protein coding genes and is strongly linked to neurodevelopmental disorders (NDDs) including ASD (Kumar *et al.*, 2008; Bijlsma *et al.*, 2009; Rosenfeld *et al.*, 2010; Shinawi *et al.*, 2010; Zufferey *et al.*, 2012). This *16p11.2* region is flanked by two homologous 147kbp sequences that arose after the evolutionary divergence of humans from other primates, generating a hot-spot for mis-aligned recombination that explains the high frequency of the *16p11.2* CNV in the human population and also the high frequency of *de novo 16p11.2* CNV (Nuttle *et al.*, 2016). In humans, the *16p11.2* microdeletion is associated with transient infant brain overgrowth (macrocephaly) and focal thickening of the cerebral cortex, while the *16p11.2* microduplication is associated with reduced brain size (microcephaly) (Qureshi *et al.*, 2014; Blackmon *et al.*, 2018). The early manifestation of anatomical phenotype in newborns, along with the onset of ASD symptoms in infancy, suggests crucial roles for *16p11.2* genes during neural development. *16p11.2* is the most prevalent CNV associated with ASD, ~1% incidence, making this CNV particularly intriguing and providing motivation for investigating the role played by *16p11.2* genes in brain development and function (Weiss *et al.*, 2008). Available lines of evidence from *16p11.2* rodent models, *16p11.2* patient derived lymphoblastoid cell lines, and human induced pluripotent stem cells genetically engineered to harbor the *16p11.2* CNV indicate that all *16p11.2* mRNAs' levels reflect the altered gene dosage of *16p11.2* genes (50% in microdeletion and 150% in microduplication

heterozygotes) (Horev *et al.*, 2011; Blumenthal *et al.*, 2014; Pucilowska *et al.*, 2015; Tai *et al.*, 2016). This indicates that multiple *16p11.2* transcript levels are affected by the *16p11.2* CNV and that the pathology of the *16p11.2* CNV could stem from altered dosage of one or more them.

While none of the individual *16p11.2* genes have been identified as sole causative genes for the *16p11.2* phenotype, *MAPK3*, *QPRT, KCTD13*, *ALDO*A, *TAOK2*, and *KIF22* have each been individually associated with a variety of neural phenotypes in non-human models. These include cell proliferation, neuronal morphology, axonal projection and spine morphogenesis, altered head size, and behavioral phenotypes (Blaker-Lee *et al.*, 2012; de Anda *et al.*, 2012; Golzio *et al.*, 2012; Pucilowska *et al.*, 2015, 2018) (Escamilla et al., 2017) (Richter et al., 2019) (Yadav et al., 2017) (Ultanir et al., 2014).

The cellular mechanisms by which the *16p11.2* CNV cause the patient phenotype are poorly understood. One plausible hypothesis is that disrupted neurogenesis causes changes in neuronal output which produce a brain with abnormal cell number or composition and that this contributes to the *16p11.2* pathology. Consistent with this hypothesis, the *16p11.2* deletion mouse model exhibits proliferation defects in cortical progenitors during pre-natal brain development and subsequently develops ASD-like symptoms (Horev *et al.*, 2011; Pucilowska *et al.*, 2015). However, it is unknown which of the proteins produced by *16p11.2* CNV genes are expressed by progenitor cells in the developing human cerebral cortex and are therefore candidates for regulating neurogenesis.

Excitatory (glutamatergic) neurons in the human cerebral cortex are derived from progenitors that reside in the ventricular and subventricular zones of the dorsal telencephalon while inhibitory (GABAergic) interneurons are derived from progenitors that reside in the ventral telencephalon

and migrate into the cerebral cortex (Clowry et al., 2010) (Bystron et al., 2008) (Ma et al., 2013) Here we focused on the potential for the *16p11.2* CNV to affect neurogenesis of excitatory neurons in the developing human cerebral cortex by identifying *16p11.2* genes that are highly expressed in cerebral cortex progenitors in the ventricular zones and down-regulated as cells become post-mitotic. We analysed previously published human fetal cortex single cell RNA sequencing (scRNA-seq) data (Pollen *et al.*, 2015) (Zhong et al., 2018) to identify candidate genes and characterize their expression in sections of developing human fetal cerebral cortex from the late first and early second trimester.

Material and Methods

Human Tissue

Human embryos ranging in age from 12-16 post-conceptual weeks (PCW) were obtained from the MRC/Wellcome-Trust funded Human Developmental Biology Resource at Newcastle University (HDBR, http://www.hdbr.org/) with appropriate maternal written consent and approval from the Newcastle and North Tyneside NHS Health Authority Joint Ethics Committee. HDBR is regulated by the UK Human Tissue Authority (HTA; www.hta.gov.uk) and operates in accordance with the relevant HTA Codes of Practice.

For cryosections 12 PCW week brains were fixed in 4%PFA/PBS for 1 week then cryoprotected with 30% sucrose/PBS and then embedded in 50:50 30%sucrose:OCT, flash frozen and sectioned at 12µm using a Leica Cryostat.
Stages used for t study: 12 PCW (2 brains), 14 PCW (1 brain) and 16 PCW (1 brain).

scRNA-seq Analysis

The publicly available scRNA-seq data sets (Pollen *et al.*, 2015) (Zhong et al., 2018) were used to identify candidate genes. Prior to dataset publication the reads were aligned, we normalized the RPKMs as log(x+1). Analysis was performed using R studio. To determine genes with significant changes a Wilcox test by *FindAllMarkers* in Seurat package was used. Monocle2 R package was used to order cells in pseudotime. To identify cell-cycle phase specific transcripts we used function CellCycleScoring from Seurat R package.

Immunohistochemistry

Immunohistochemistry was carried out on paraffin sections obtained from HDBR. Antigen retrieval consisting of boiling sections in 10mM sodium citrate pH6 for 10 mins was used for all stains. Primary antibodies were diluted in 20% blocking serum in pH7.6 Tris buffered saline (TBS) and sections incubated overnight at 4°C. Primary antibodies used: KID 1/5000 DAB, 1/2000 fluorescent (Invitrogen PA5-29490), KI67 1/800 (Novus Biologicals NBP2-22112), ALDOA 1/100 (Sigma HPA004177).

For colourmetric stains, sections were incubated 1hr at room temperature with biotinylated secondary antibody (1/200) followed by incubation for 1 hour with ABC (Vector Labs) and developed with diaminobenzidine solution (Vector Labs), washed, counterstained with nuclear fast red, dehydrated and then mounted using DPX.

For immunofluorescence sections were incubated with secondary antibodies 1/200 1 hour room temperature, counterstained with 4',6-diamidino-2-phenylindole dihydrochloride (DAPI;ThermoFisher) and mounted with Vectashield H1400 Hardset Mounting Medium (Vector

Labs). Extensive TBS washes were carried out between each step.

In Situ Hybridisation

PCR primers used to clone *in situ* probes from human cDNA into pGEMTeasy for preparation of DIG labelled RNA were as follows: *ALDOA*, CTG TCA CTG GGA TCA CCT T, & GTG ATG GAC TTA GCA TTC AC*; KIF22*,: CGA GAG CGG ATG GTG CTA AT &: GAG ACC CAG GAT GTT TGC CT; *PAGR1*, ATG ATG AGC CAG TGA CAC CA & TCT GCC TCT CCC TTC AAG TG; *HIRIP3*, TGG TGC CCA TCG AAA CTA CA & TGG CCC AAA ATA CAG GAG GT; & *MAZ*, CAC GAG GAG AAA GTG CCA TG & GAG AGA AGA GGA CCG TCG AG.

In situ hybridisation was performed on cryosections of 12 PCW brain as described previously (Radonjić *et al.*, 2014). Briefly, 12µm cryosections were dried at 37°C for 3 hours then incubated overnight at 70°C in hybridization mix containing 1x salts (200mM NaCl, 10mM Tris HCl (pH 7.5), 1mM Tris Base, 5mM $NaH_2PO_42H_2O$, 5mM Na_2HPO_4, 0.5M EDTA: Sigma-Aldrich), 50% deionized formamide, 10% dextran sulfate, 1mg/ml rRNA, 1x Denhardt's, and DIG-labelled RNA probe. Next day sections were washed 3 times at 70°C in wash buffer comprising 1x SSC, 50% formamide, 0.1% Tween-20 and then 3 times at RT in 1x MABT (20mM Maleic acid, 30mM NaCl, 0.5% Tween-20 and pH adjusted to 7.5 with 10mM NaOH). Sections were incubated 1hr RT in 1x MABT blocking solution (20% sheep serum, 2% blocking reagent) and then incubated overnight with anti-DIG antibody 1:1500 in blocking solution at 4°C followed by colour reaction overnight at RT.

Microscopy and Imaging

DAB and *in situ* hybridisation images were taken using a Leica DMNB microscope with an attached Leica DFC480 Camera. Fluorescence images were obtained with a Leica DM5500B epifluorescence microscope with a DFC360FX camera. Confocal images were obtained using Nikon A1R FILM microscope and analysed in ImageJ.

Image Analysis and Quantification

For DAB stains and *in situ* hybridisation the images were stitched in ImageJ using the stitching plugin (Preibisch, Saalfeld and Tomancak, 2009).

For KIF22 analysis of DAB stains rectangular counting boxes (34x88µm) were overlaid across the section. Using ImageJ cell counting plugin cells in each box were counted and denoted KIF22+ (brown) or KIF22- (red). The distinction between the regions (VZ, SVZ, IZ/CP) was determined anatomically by cell density. The count for each box was averaged with other boxes in the region to provide the final value.

For analysis of KIF22/KI67 double staining counting boxes (20x145µm) were overlaid over the VZ and SVZ (determined based on cell density). For determining intensity cells were randomly selected on the DAPI channel, the nucleus outlined and intensity of KIF22 and KI67 recorded. 20 cells were selected per box and the counts from individual boxes combined to give final values.

For subcellular ALDOA analysis counting boxes (20x145µm) were overlaid over the VZ and SVZ. Cells were randomly selected on the DAPI channel, far enough apart to ensure their cytoplasm would not overlap, the Z plane through the center of the cell was used and the nucleus outlined. The KI67 and ALDOA intensity was measured constituting the nuclear value. To

obtain ALDOA cytoplasmic intensity the nuclear outline was duplicated and extended 4 pixels allowing a reading of just the cytoplasmic area to be obtained (see Fig.6e). This was performed for 10 cells in each box and the counts from individual boxes combined to give final values.

Data Analysis and Statistics

Where error bars are shown they are expressed as mean ± SEM. Boxplots show median and upper and lower quartiles. Statistical comparison between two groups was performed with a *t* test. Statistical comparison between more than two groups was performed with ANOVA followed by *post hoc* test. $P<0.05$ was considered statistically significant. Analysis was performed using GraphPad Prism.

Results

Analysis of scRNA-seq data identifies *KIF22* and *ALDOA* as progenitor-enriched *16p11.2* transcripts in the developing human fetal cerebral cortex.

The *16p11.2* CNV involves microduplication or microdeletion of a 593kb locus on human chromosome 16 containing 29 protein coding genes (Fig.1a). The aim of the current study is to identify *16p11.2* genes which are potential candidates for being involved in neurogenesis in the developing human cerebral cortex (Fig.1b,c) and whose altered dosage in *16p11.2* microdeletion or microduplication patients may disrupt neurogenesis and contribute to the CNV phenotype. We reasoned that *16p11.2* genes important for neurogenesis would be highly expressed in proliferating progenitor cells and down-regulated as cells became post-mitotic.

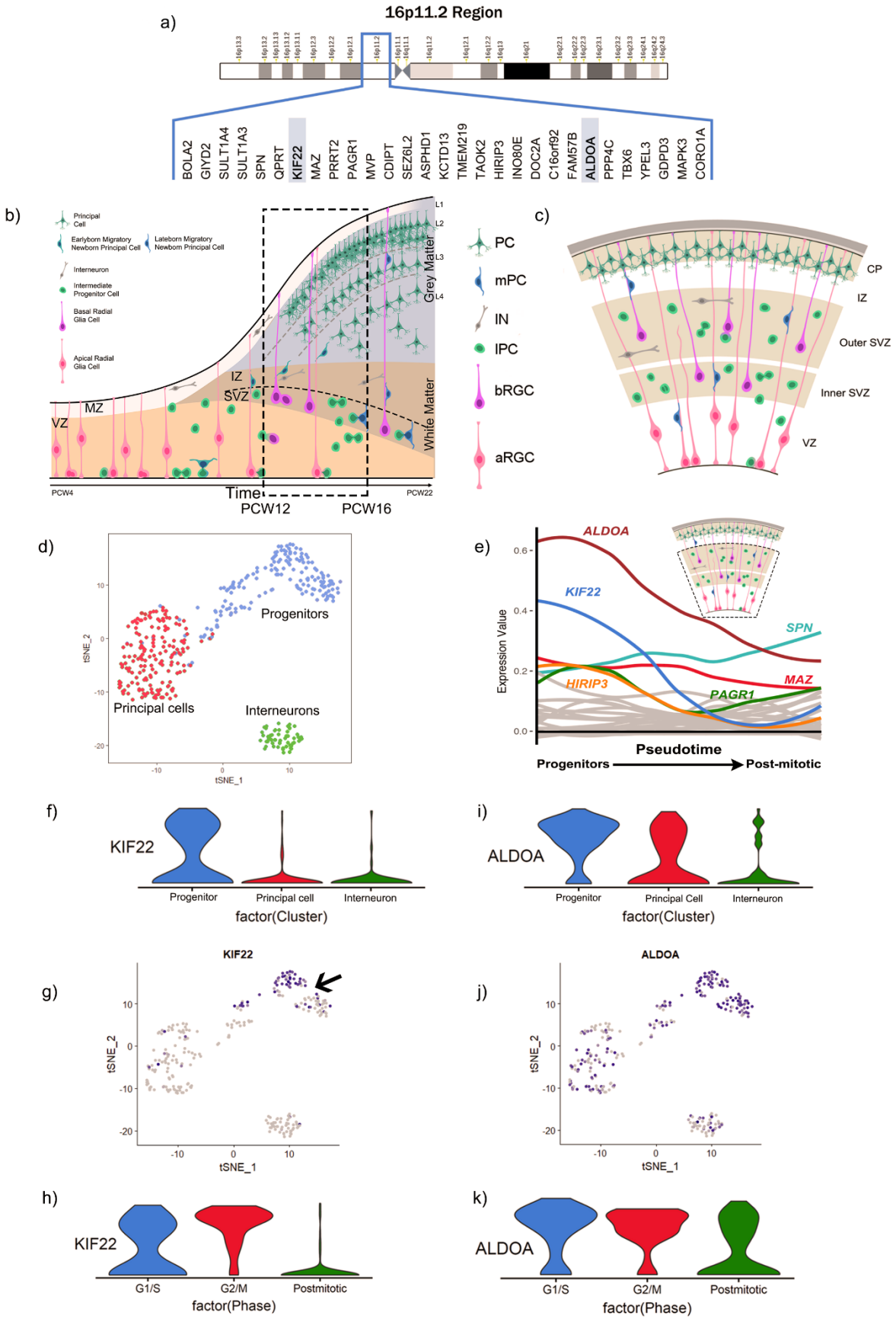

a)
16p11.2 Region
BOLA2
GIYD2
SULT1A4
SULT1A3
SPN
QPRT
KIF22
MAZ
PRRT2
PAGR1
MVP
CDIPT
SEZ6L2
ASPHD1
KCTD13
TMEM219
TAOK2
HIRIP3
INO80E
DOC2A
C16orf92
FAM57B
ALDOA
PPP4C
TBX6
YPEL3
GDPD3
MAPK3
CORO1A
b)
Principal Cell
Earlyborn Migratory Newborn Principal Cell
Lateborn Migratory Newborn Principal Cell
Interneuron
Intermediate Progenitor Cell
Basal Radial Glia Cell
Apical Radial Glia Cell
L1
L2
L3
L4
Grey Matter
White Matter
IZ
SVZ
MZ
VZ
PCW4
Time
PCW12
PCW16
PCW22
c)
PC
mPC
IN
IPC
bRGC
aRGC
CP
IZ
Outer SVZ
Inner SVZ
VZ
d)
Progenitors
Principal cells
Interneurons
tSNE_1
tSNE_2
e)
ALDOA
KIF22
SPN
HIRIP3
MAZ
PAGR1
Expression Value
Pseudotime
Progenitors
Post-mitotic
f)
KIF22
Progenitor
Principal cell
Interneuron
factor(Cluster)
i)
ALDOA
Progenitor
Principal Cell
Interneuron
factor(Cluster)
g)
KIF22
tSNE_1
tSNE_2
j)
ALDOA
tSNE_1
tSNE_2
h)
KIF22
G1/S
G2/M
Postmitotic
factor(Phase)
k)
ALDOA
G1/S
G2/M
Postmitotic
factor(Phase)

Figure 1: **Bioinformatics analysis of scRNA-seq from the VZ and SVZ of the 16-18GW human fetal cortex.** a) 16p11.2 region and genes. b) schematic adapted from Budday et al 2015 shows the process of human fetal cortex development over time. Dotted box indicates approximately the time-period of interest for our study; 12-16PCW. c) schematic of human cortical structure during development. d) tSNE clustering of cell types. e) changing mRNA expression levels of *16p11.2* genes as cells move from progenitors to neurons with *KIF22* and *ALDOA* transcripts identified as changing significantly. Schematic of human cortex with dotted box indicates the compartments of the brain that were used to generate this dataset – the germinative VZ and SVZ. f) Violin plots showing distribution of *KIF22* in different cell types. g) *KIF22* gradient plot (arrow indicates progenitors with a subset expressing high levels of KIF22 mRNA). h) Violin plots showing distribution of *KIF22* at different cell cycle stages. i) Violin plots showing distribution of *ALDOA* mRNA levels in different cell types. j) *ALDOA* gradient plot. k) Violin plots showing distribution of *ALDOA* mRNA at different cell cycle stages.

We took advantage of a published single cell RNA-sequencing (scRNA-seq) data-set acquired from 393 cells of the ventricular zone (VZ) and subventricular zone (SVZ) of gestational week (GW) 16-18 human fetal cerebral cortex (equivalent to post conception week (PCW) 14-16) to perform an unbiased screen to identify *16p11.2* transcripts that matched this expression profile (Pollen *et al.*, 2015). Dimensional reduction of the scRNA-seq data separated the cells into three clusters based on transcriptome similarity (Fig.1d – each dot on the tSNE plot represents an individual cell) that were subsequently identified as the three cardinal cell classes of progenitors (blue), post-mitotic neurons/ principal cells (red) and interneurons (green), by expression of cell-type specific transcripts. We next used the monocle2 R package to order the cells in pseudotime

using the normalized expression levels of selected differentially expressed genes (DEGs) as input to order the cells (Trapnell *et al.*, 2014; Qiu *et al.*, 2017) (Fig.1e) moving from the progenitor state (left) to post mitotic state (right) along the X-axis. We plotted the average expression of each *16p11.2* transcript at each pseudotime-point on the Y-axis. We found that two genes, *KIF22* (blue line) and *ALDOA* (brown line), were notable for having high expression in progenitors that declined as cells became post-mitotic. A Wilcox test identified *ALDOA* and *KIF22* as the only *16p11.2* transcripts that were significantly higher in progenitor than neuronal populations ($p<<0.05$). Although not significantly enriched in progenitors *HIRIP3* (orange line), *MAZ* (red line), *PAGR1* (green line) were expressed in progenitors at higher levels than the remaining *16p11.2* transcripts (shown as grey lines), many of which were barely expressed at all.

Violin plots of the numbers of cells expressing different levels of *KIF22* mRNA in the different cardinal cell classes show that *KIF22* is expressed predominantly in progenitors (Fig.1f) and mapping the expression level of *KIF22* onto the tSNE plot (Fig.1g) revealed that *KIF22* expression is highest in a subset of the progenitor cluster (arrow in Fig.1g) with a substantial proportion of progenitor cells expressing relatively low levels of *KIF22.* Very few post-mitotic neurons, both interneurons and principal cells, express appreciable levels of *KIF22* (Fig.1f,g). The expression of *KIF22* in a subset of progenitors prompted us to ask whether its expression was related to the cell-cycle phase. We used the expression of cell-cycle phase specific transcripts using function CellCycleScoring from Seurat R package to divide the cells into three classes (Macosko *et al.*, 2015; Tirosh *et al.*, 2016), G1/S, G2/M, and post-mitotic neurons, and compared *KIF22* transcript levels between these three groups using a violin plot (Fig.1h). We found that the majority of cells in G2/M phase expressed higher levels of *KIF22* (red plot), cells in G1/S expressed lower levels (blue plot) while the vast majority of post-mitotic cells expressed low levels of *KIF22* (green plot).

Similar analysis for *ALDOA* show that while a greater proportion of cells expressing the highest levels of *ALDOA* are progenitors (blue plot) there are also a substantial number of principal cells (red plot) expressing similarly high levels of *ALDOA* transcripts although very few interneurons (green plot). Mapping *ALDOA* expression level onto the tSNE plot (Fig.1j) shows cells expressing high levels of *ALDOA* are evenly distributed throughout the progenitor cluster with appreciable numbers of principal cells expressing high levels of *ALDOA* and a much lower proportion of interneurons. In contrast to *KIF22*, there is no clear difference in the partitioning of *ALDOA* expression level between different phases of the cell-cycle (Fig.1k). We performed the same analysis on another developing human cortex scRNAseq data set that spanned a wider developmental interval (PCW10-28) and also included cells from all layers of the cerebral cortex (Zhong et al., 2018). This analysis produced similar results. *KIF22* transcripts showing them to be enriched in progenitors (Sup Fig.1b,e) and in G2/M phase of the cell cycle (Sup Fig.1g). Although *ALDOA* transcript levels were highest in progenitors they did not decline as much as *KIF22* (Sup Fig.1b,f) in postmitotic cells and showed no clear difference between different phases of the cell-cycle (Sup Fig.1h).

We next used in situ hybridisation to visualise the expression of *KIF22*, *ALDOA*, *HIRIP3*, *PAGR1*, and *MAZ* transcripts in the different layers of the 12 PCW brain. We identified the VZ and SVZ based on cytoarchitecture, however, to validate our delineation we stained a section with PAX6, expressed by progenitor cells (Fig.2a-a'*). This confirmed the absence of progenitor cells in the IZ and CP and allowed us a guidance for estimating cellular location in other sections.

Consistent with the scRNAseq data *KIF22* and *ALDOA* transcripts are the most clearly differentially expressed between zones containing progenitors (VZ and SVZ) and the more superficial layers that are mainly composed of postmitotic cells (IZ and SP/CP). *KIF22* mRNA

expression is most prominent in the VZ and SVZ with a few expressing cells in the IZ and SP/CP (Fig.2b and 2b') While *ALDOA* mRNA expression is most prominent in the proliferative VZ and SVZ there are substantial numbers of *ALDOA* expressing cells in the SP/CP (Fig.2c and 2c'). *HIRIP3* (Fig.2d,d'), *PAGR1* (Fig.2e,e'), and *MAZ* (Fig.2f,f') mRNAs are expressed in the VZ/SVZ and also in cells of the SP/CP. In addition to being expressed in the proliferative zones of the cerebral cortex *KIF22*, *ALDOA*, *HIRIP3*, *PAGR1*, and *MAZ* are also expressed in the ventricular zone of the ganglionic eminences where interneuron progenitors reside suggesting the hypothesis that interneuron development may be affected by alterations in their dosage in the *16p11.2* CNV.

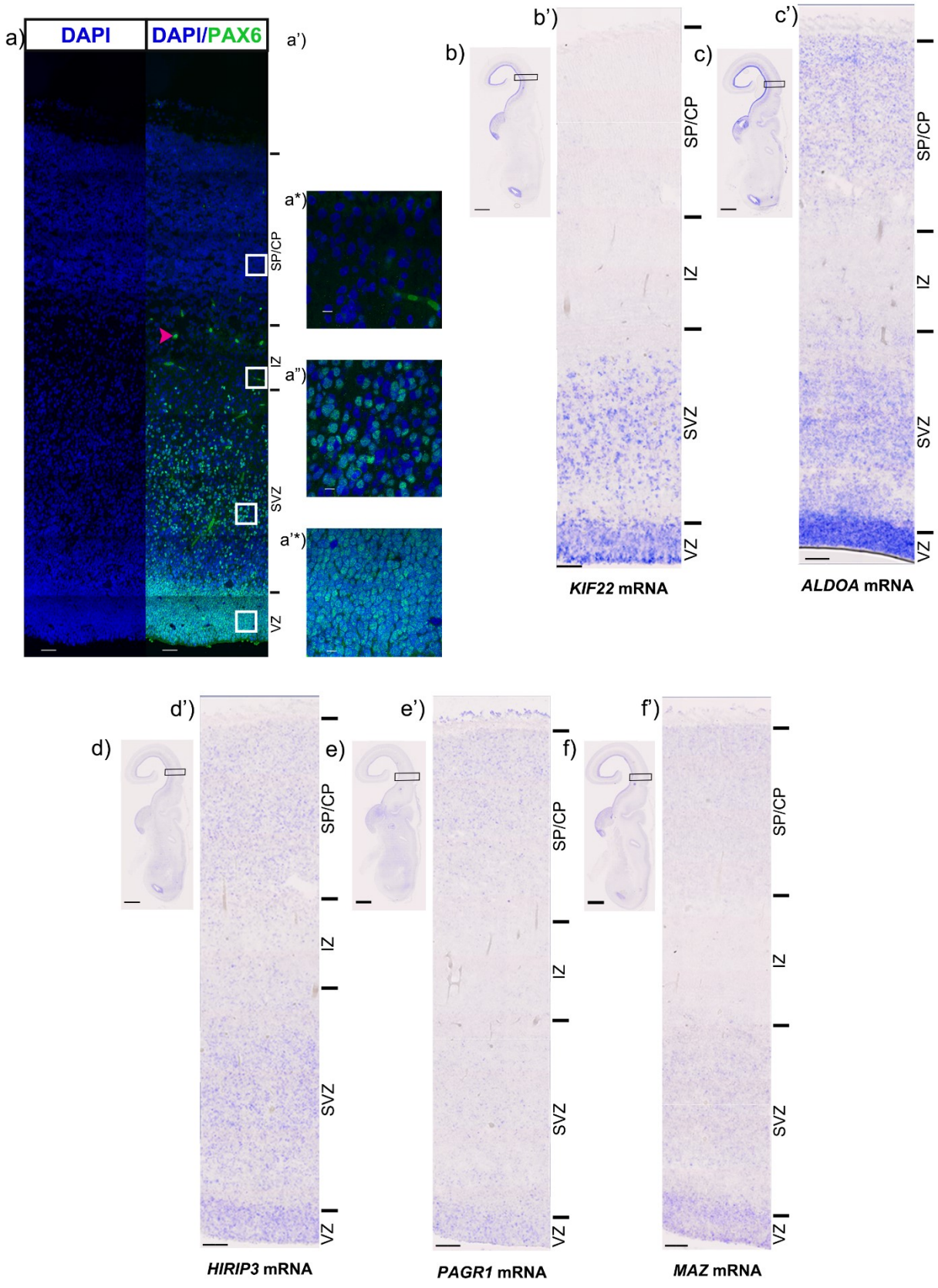

a)
DAPI
DAPI/PAX6
a')
a*)
a")
a'*)
SP/CP
IZ
SVZ
VZ
b)
b')
c)
c')
KIF22 mRNA
ALDOA mRNA
d)
d')
e)
e')
f)
f')
HIRIP3 mRNA
PAGR1 mRNA
MAZ mRNA

Figure 2: ***In situ*** **hybridisation of candidate genes.** a) PAX6 protein (green) at 12pcw. a-a'*) show high magnification images of PAX6 protein expression in the a') SP/CP, a*) IZ, a") SVZ and a'*) VZ. Low magnification scale bars = 50μm, high magnification scale bars = 10μm. b) Low magnification image of *KIF22* mRNA in the 12pcw human fetal cortex, b') High magnification showing *KIF22* mRNA (blue) to be predominantly expressed in the germinative zones. c) Low magnification image of *ALDOA* mRNA in the 12pcw human fetal cortex, b') High magnification showing *ALDOA* mRNA (blue) to be predominantly expressed in the germinative zones but also some expression in the IZ and CP. d) Low magnification image of *HIRIP3* mRNA in the 12pcw human fetal cortex, d') High magnification showing *HIRIP3* mRNA (blue) to be expressed throughout the telencephalic wall. e) Low magnification image of *PAGR1* mRNA in the 12pcw human fetal cortex, e') High magnification showing *PAGR1* mRNA (blue) to be expressed throughout the telencephalic wall. f) Low magnification image of *MAZ* mRNA in the 12pcw human fetal cortex, f') High magnification showing *MAZ* mRNA (blue) to be expressed throughout the telencephalic wall. For ISH, low magnification scale bars = 2mm and high magnification scale bars = 100μm.

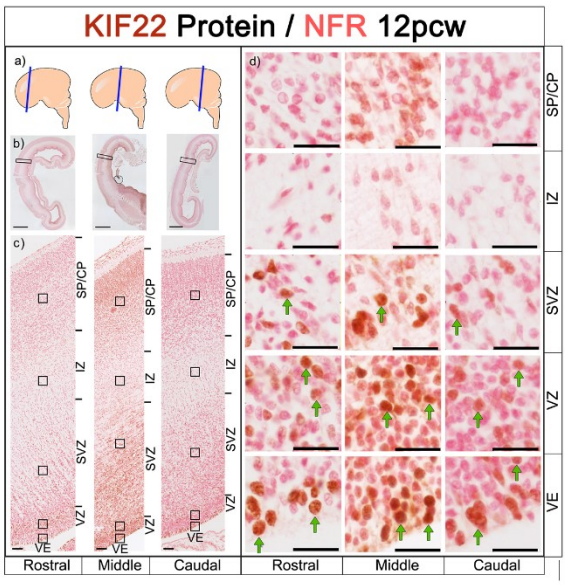
To conclude, of all the 29 *16p11.2* transcripts, five, *KIF22*, *ALDOA*, *HIRIP3*, *PAGR1*, and *MAZ*, are expressed in the ventricular and subventricular zones at higher levels than in post-mitotic cells. Of these only two*, KIF22* and *ALDOA*, are significantly enriched in progenitors compared to post-mitotic cells making them candidates for having specific roles in neurogenesis in the developing human fetal cerebral cortex. Although both are enriched in progenitors, *KIF22* and *ALDOA* transcript expression shows notable differences: *KIF22* transcripts are more restricted to

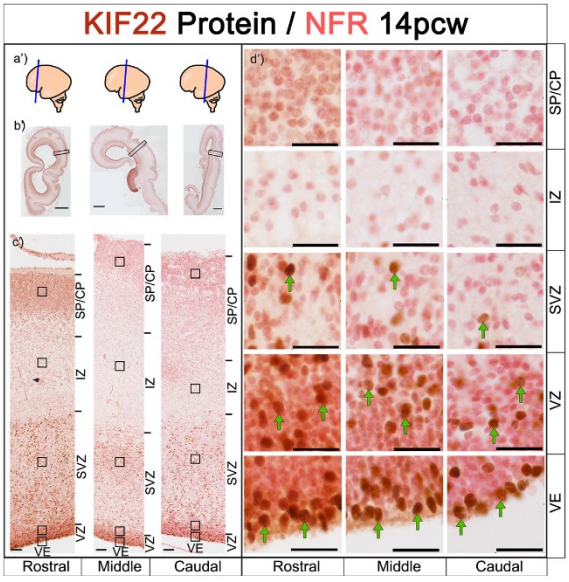
progenitors and their levels vary as the cell-cycle progresses. We next describe the expression of KIF22 and ALDOA protein over a range of developmental stages.

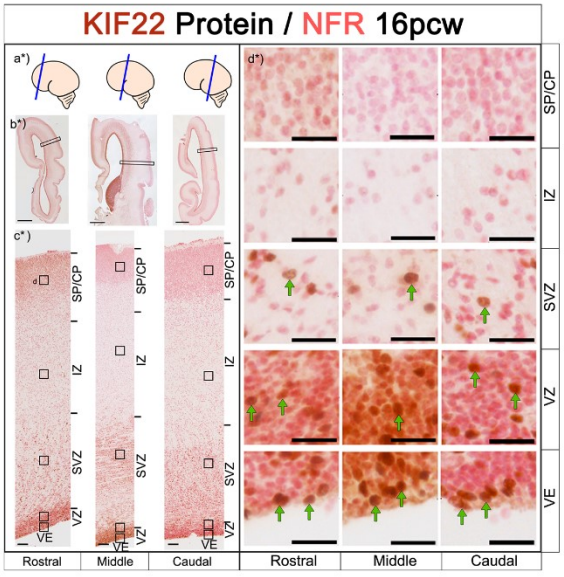
KIF22 protein is expressed in germinal zones of 12, 14 and 16 PCW cortex

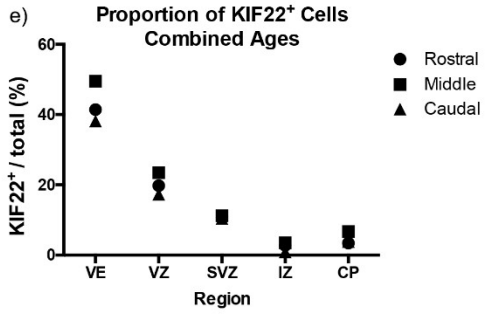
Here we characterize KIF22 protein expression during human corticogenesis. Coronal cortex sections spaced along the rostral-caudal axis were immunostained for KIF22 protein and counterstained with Nuclear Fast Red (NFR) to show cytoarchitecture. $KIF22^{+}$ (brown) and $KIF22^{-}$ (red) cells were counted for each region in the telencephalic wall (VE, VZ, SVZ, IZ and CP) (see methods for details of sampling) and lamination was identified by cell density (Bayer and Altman, 2002, 2005). These data are shown for three developmental stages, 12 PCW (Fig.3 a-d), 14 PCW (Fig.3 a'-d'), and 16 PCW (Fig.3 a*-d*). At all stages and rostro-caudal positions KIF22 expressing cells appear most abundant in the VE followed by the VZ and SVZ with the IZ and CP presenting a very low to complete absence of KIF22 (Fig.3 c, c', and c* with higher magnification of boxed regions from each zone shown in d, d', and d* respectively, green arrows indicate examples of individual $KIF22^{+}$ cells).

We next pooled $KIF22^{+}$ cell count data in two ways to compare between all ages (Fig.3e) and anatomical regions (Fig.3f) and found that the percentage of $KIF22^{+}$ cells in the VE (40-50%) was consistently higher than other regions, followed by the VZ (20-30%) and SVZ (10%), with even fewer cells (<10%) in the IZ, and CP (Fig.3e,f). This result describes KIF22 protein as predominantly restricted to a subset of cells in the germinal zones of the developing cortex at all stages studied.


KIF22 Protein / NFR 12pcw
a)
b)
c)
d)
SP/CP
IZ
SVZ
VZ
VE
Rostral
Middle
Caudal


KIF22 Protein / NFR 14pcw
a')
b')
c')
d')
SP/CP
IZ
SVZ
VZ
VE
Rostral
Middle
Caudal


KIF22 Protein / NFR 16pcw
a*)
b*)
c*)
d*)
SP/CP
IZ
SVZ
VZ
VE
Rostral
Middle
Caudal


e)
Proportion of KIF22+ Cells
Combined Ages
KIF22+ / total (%)
60
40
20
0
VE
VZ
SVZ
IZ
CP
Region
Rostral
Middle
Caudal

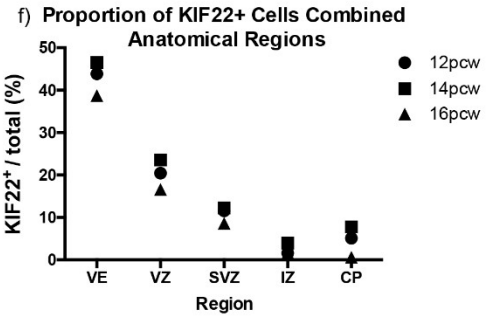

f) Proportion of KIF22+ Cells Combined
Anatomical Regions
KIF22+ / total (%)
50
40
30
20
10
0
VE
VZ
SVZ
IZ
CP
Region
12pcw
14pcw
16pcw

Figure 3: KIF22 protein expression levels in the cerebral cortex at 12, 14 and 16 PCW a, a', a*) schematic showing brain regions sectioned. b, b', b*) images of whole brain section, scale bars =2mm. c, c', c*) sections spanning the rostral-caudal axis showing KIF22 expression in the telencephalic wall, scale bars = 100μm. d, d', d*) high magnification images of different cortical zones rostral-caudal. KIF22+ cells in brown and examples indicated by green arrows, KIF22- cells in pink, scale bars =25 μm. e) Quantification of KIF22 expressing cells with all three ages combined. f) Quantification of KIF22 expressing cells with rostral, middle, caudal values combined.

KIF22 protein expression is restricted to proliferating cells

KIF22 protein expression is almost exclusively restricted to a subset of cells in the proliferative regions. From the scRNA-seq data, we expect these to be progenitor cells (Fig.1). To identify these KIF22 positive cells we performed double immunofluorescence for KIF22 and KI67 (a protein expressed in all proliferating cells (Scholzen and Gerdes, 2000; Miller *et al.*, 2018)). $KIF22^+$ cells were predominantly located in the VE, VZ and SVZ (Fig.3), therefore these regions were examined for analysis. Low magnification of KI67/KIF22 staining are shown in 12 PCW (Fig.4a) and 14 PCW (Fig.4b) with higher magnification showing individual cells in Fig.4c. Cell counts for $KIF22^+/KI67^+$ labelled cells show that the majority (80-90%) of $KI67^+$ cells also express KIF22 both at 12 (Fig.4d) and 14 PCW (Fig.4e) or across the rostral-caudal axis. Combining the data for anatomical locations and ages revealed significantly more $KIF22^+/KI67^+$ cells than $KIF22^+/KI67^-$ and $KIF22^-/KI67^+$ cells (Fig.4f).

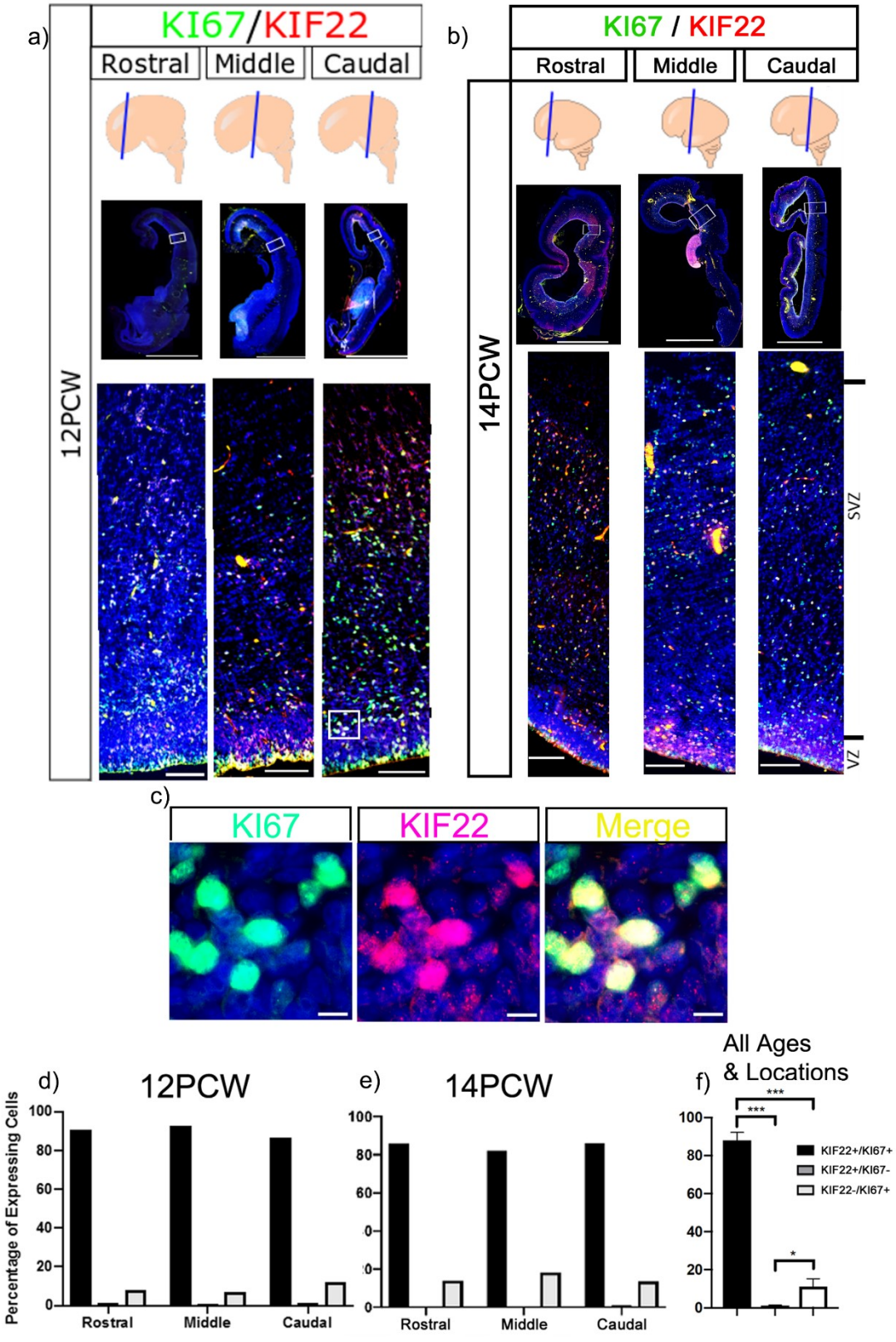

a)
KI67/KIF22
Rostral
Middle
Caudal
12PCW
b)
KI67 / KIF22
Rostral
Middle
Caudal
14PCW
SVZ
VZ
c)
KI67
KIF22
Merge
d)
12PCW
Percentage of Expressing Cells
100
80
60
40
20
0
Rostral
Middle
Caudal
e)
14PCW
100
80
60
0
0
0
Rostral
Middle
Caudal
f)
All Ages
& Locations
***
***
*
100
80
60
40
20
0
KIF22+/KI67+
KIF22+/KI67-
KIF22-/KI67+

Figure 4: Immunofluorescence of KIF22 and KI67 proteins in the cortex. a) KIF22 and KI67 at 12 PCW, low magnification scale bars = 4mm, high magnification scale bars = 100 μm. b) KIF22 and KI67 at 14 PCW, low magnification scale bars = 4mm, high magnification scale bars = 100 μm. c) high magnification of KI67/KIF22 expressing cells. Scale bars = 10 μm. d) Percentage of cells expressing KIF22, KI67 or both at 12 PCW. e) Percentage of cells expressing KIF22, KI67 or both at 14 PCW. f) Combined data of percentage of cells expressing KIF22, KI67 or both.

KIF22 levels vary with cell-cycle phase

From the scRNA-seq analysis, and the variable KIF22 protein levels in $KI67^+$ cells, we hypothesised that KIF22 protein levels change throughout the cell cycle. To test this, we quantified nuclear immunofluorescence intensity of KIF22 and KI67 in two 12 PCW brains (see methods for details of sampling procedure). KI67 protein levels vary during the cell cycle: lowest in G1 phase, increasing through S and G2 to peak in mitosis (Fig.5o) (Scholzen and Gerdes, 2000; Miller *et al.*, 2018). We found a strong correlation between KIF22 and KI67 intensity (Brain 1 rostral R^2=0.8095, middle R^2=0.8139, caudal R^2= 0.6691. Brain 2 rostral R^2 = 0.7489, middle R^2=0.7447, caudal R^2=0.7763) (Fig.5d-f). To ensure the correlation observed was not a result of nucleus size changing with cell cycle, we confirmed that KIF22 protein levels did not correlate with nuclear size by DAPI staining (Brain 1 rostral KIF22 R^2=0.104, middle KIF22 R^2=0.0874, caudal KIF22 R^2=0.2969. Brain 2 rostral KIF22 R^2=0.1183, middle KIF22 R^2=0.0512, caudal KIF22 R^2=0.2287). A strong positive correlation was also observed at 14 PCW (Brain 1 rostral R^2=0.7465, middle R^2=0.6668, caudal R^2= 0.634, Fig.5j-l). Again, we confirmed that KIF22

protein levels did not correlate with nuclear size (rostral KIF22 R^2=0.1239, middle KIF22 R^2=0.0599, caudal KIF22 R^2=0.0229). This demonstrates that the correlation between KI67 and KIF22 is consistent between ages and rostral-caudal location. Combining all values of KI67/KIF22 nuclear intensity values showed that KIF22 was expressed at significantly higher levels in $KI67^+$ cells (Fig.5m) with a strong correlation (R^2=0.7236) between KIF22 and KI67 levels (Fig.5n). Although KIF22 expressing cells were scattered throughout the VE, VZ and SVZ there was a general trend for the cells expressing the highest levels of KIF22 to be closest to the apical surface (yellow coloured dots on scatterplots Fig.5d-f, j-l,n) with lower expressing cells tending to be further from the apical surface (blue coloured dots on scatterplots Fig.5d-g, j-l,n). During interkinetic nuclear movement radial glial cell nuclei move to the apical surface to perform mitosis so this spatial distribution suggests KIF22 is expressed at high levels by radial glial cells undergoing mitosis at the apical surface of the VZ. Cerebral cortex progenitor that do not undergo interkinetic movement (intermediate progenitors and outer sub-ventricular zone progenitors) cannot be assigned to cell cycle phase using their position as we were able to do for radial glial cells which do undergo interkinetic movement. As *KIF22* transcript levels were at highest in all progenitors in M-phase (Fig.1h) we suspect that cells expressing high levels of KIF22 and KI67 further away from the apical surface are cortical progenitors in M-phase not undergoing interkinetic movement but further experiments would be needed to show this.

From these data we show that, for radial glial cells, KIF22 protein levels change throughout the cell cycle in positive correlation with KI67: KIF22 is present in G1 and increases through S and G2 phase to peak in mitosis (Fig.5p).

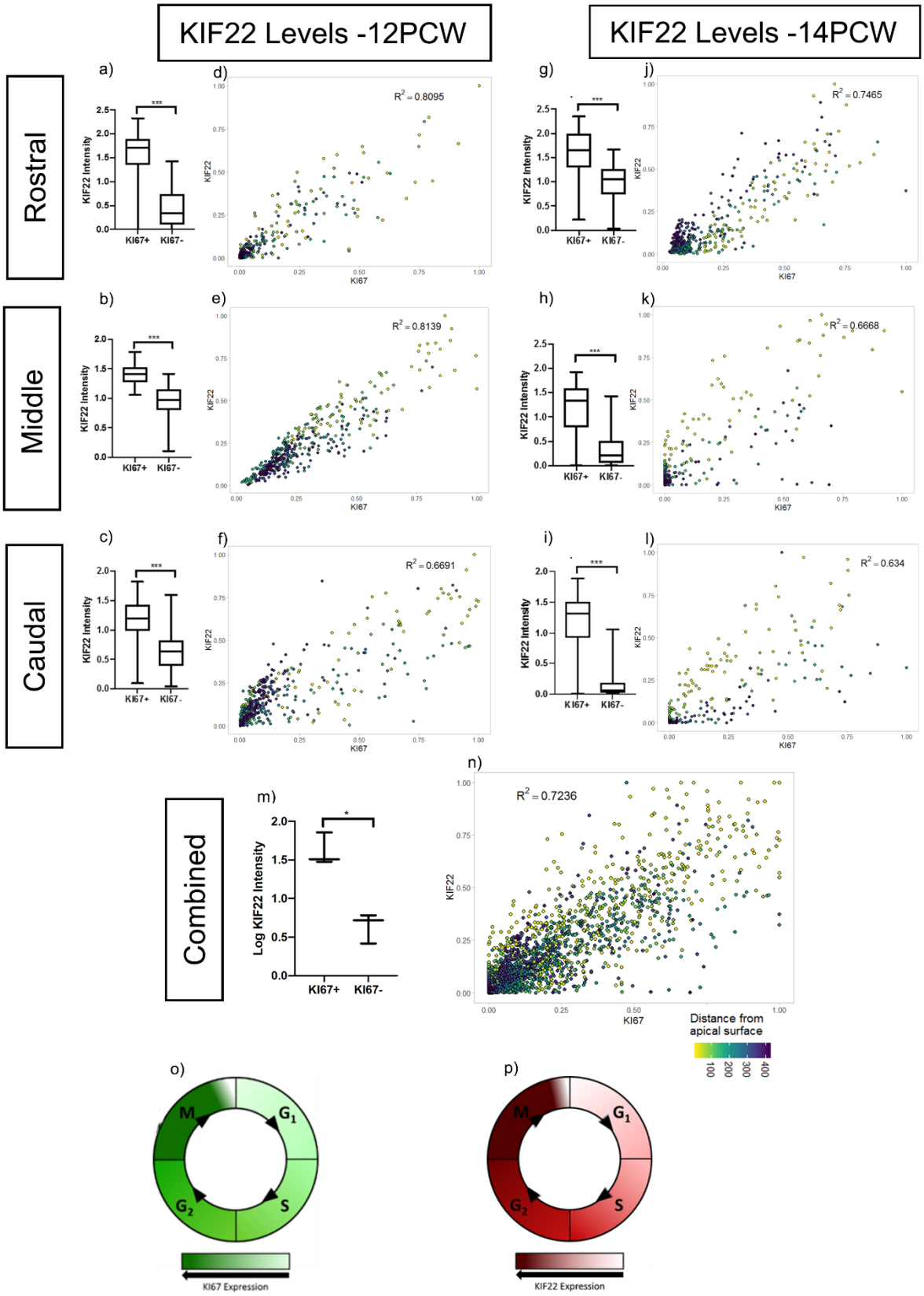
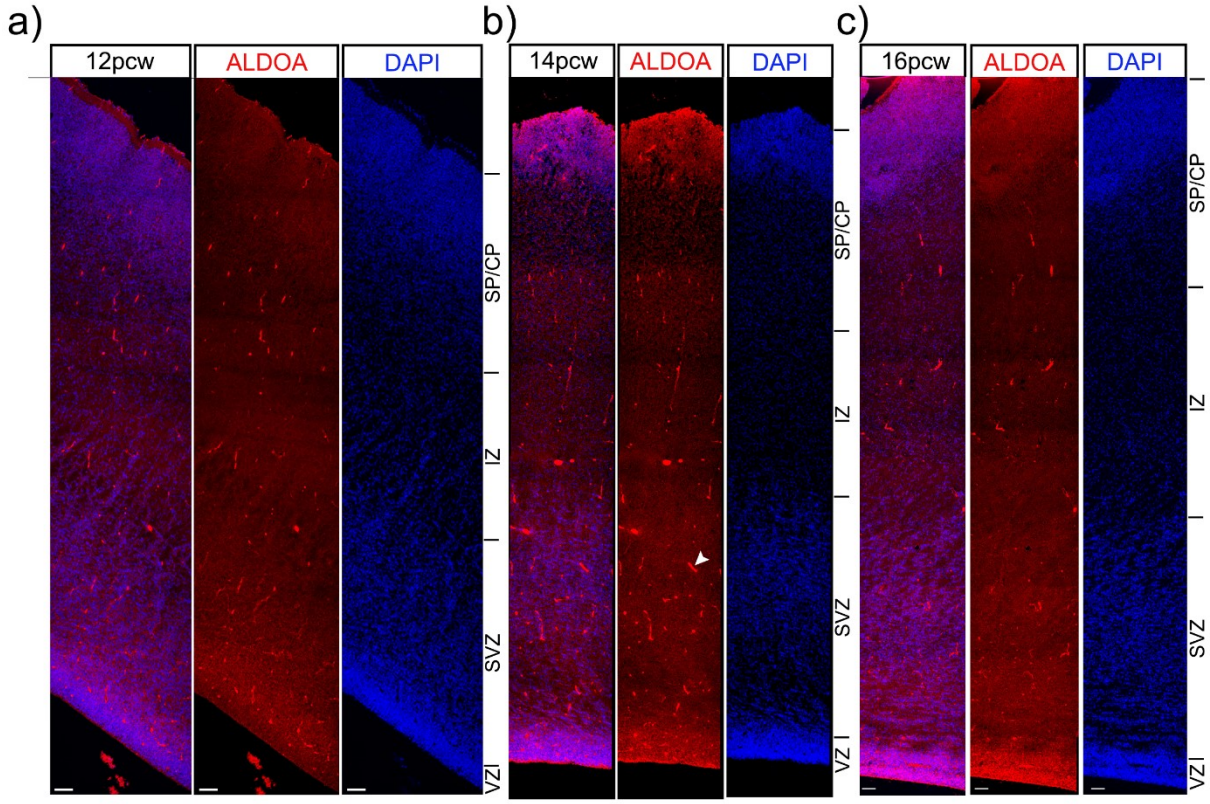

KIF22 Levels -12PCW
KIF22 Levels -14PCW
Rostral
Middle
Caudal
Combined
a)
b)
c)
d)
e)
f)
g)
h)
i)
j)
k)
l)
m)
n)
o)
p)
KIF22 Intensity
Log KIF22 Intensity
KI67+
KI67-
KIF22
KI67
$R^2 = 0.8095$
$R^2 = 0.8139$
$R^2 = 0.6691$
$R^2 = 0.7465$
$R^2 = 0.6668$
$R^2 = 0.634$
$R^2 = 0.7236$
Distance from apical surface
M
G1
S
G2
KI67 Expression
KIF22 Expression

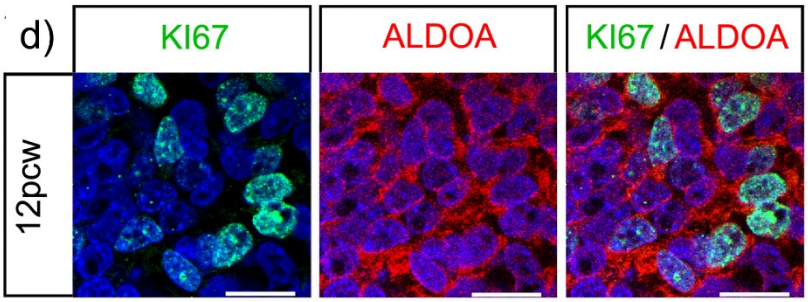
Figure 5: Quantification of KIF22 protein levels. a,b,c) 12 PCW quantification of KIF22 fluorescence intensity in KI67+/KI67- cells (raw data transformation =+1(log), unpaired t-test with Welch's correction, p=<0.001). d,e,f) intensity correlations of KIF22 and KI67 nuclear fluorescence intensity at 12 PCW. g, h, i) 14 PCW quantification of KIF22 fluorescence intensity in KI67+/KI67- cells (raw data transformation =+1(log), unpaired t-test with Welch's correction, p=<0.001). j, k, l) intensity correlations of KIF22 and KI67 nuclear fluorescence intensity at 14 PCW. m) quantification of KIF22 fluorescence intensity in KI67+/KI67- cells 12 and 14 weeks combined (raw data transformation =(log), paired t-test, p=0.0122). n) intensity correlations of KIF22 and KI67 nuclear fluorescence intensity for rostral-caudal points at 12 and 14 PCW with distance from apical surface indicated by dot colour. o) diagram of KI67 protein levels throughout the cell cycle. p) model based on our results of KIF22 protein levels throughout the cell cycle.

ALDOA protein is highest in the germinal zones of the cortex

Bioinformatics analysis and *in situ* hybridisation show *ALDOA* mRNA levels decrease as progenitor cells move towards a neuronal fate (Fig.1). Here we used immunofluorescence to characterize ALDOA protein expression across the telencephalic wall at 3 developmental time points; at 12, 14 and 16 PCW, ALDOA immunofluorescence is most intense in the VZ and SVZ before decreasing in the cortical plate (Fig.6a-c). Double immunofluorescence for KI67 and ALDOA viewed at high magnification shows that ALDOA protein is primarily localized outside $DAPI^+$ nuclei in the cytoplasm and that ALDOA is expressed by $KI67^+$ progenitor cells and also

by cells that do not express KI67 (Fig.6d). The schematic (Fig.6e and 6e') illustrates the areas used for quantification of nuclear and whole cell ALDOA fluorescence presented in Fig.6.


a)
12pcw
ALDOA
DAPI
b)
14pcw
ALDOA
DAPI
c)
16pcw
ALDOA
DAPI
SP/CP
IZ
SVZ
VZ


d)
KI67
ALDOA
KI67 / ALDOA
12pcw

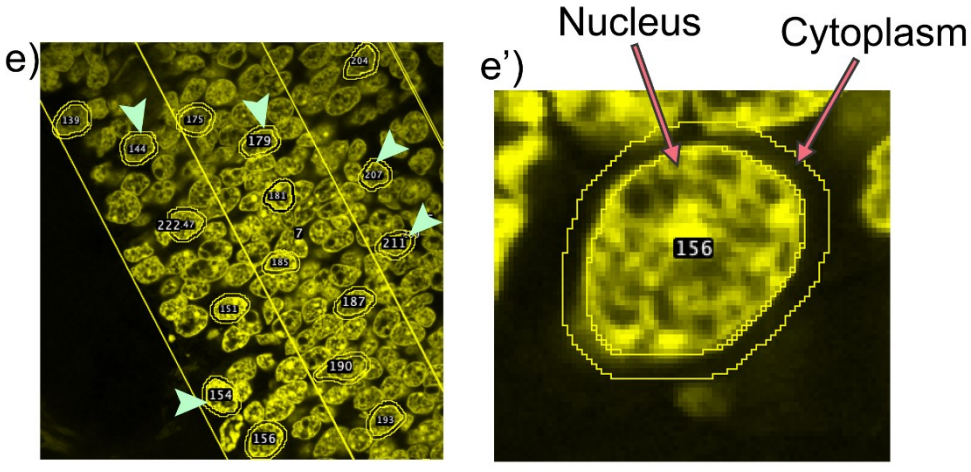

e)
e')
Nucleus
Cytoplasm
156

Figure 6: ALDOA protein expression in the cortex. ALDOA protein expression across the telencephalic wall at a) 12, b) 14 and c) 16 pcw. Scale bars = 100μm. White arrow indicates non-specific binding to blood vessels. d) high magnification immunofluorescence of ALDOA and KI67 proteins, scale bar = 10μm. e) low power image showing how cells were randomly selected for analysis using the DAPI channel. e') high power image showing how the nucleus and cytoplasm were delineated for analysis

ALDOA protein levels do not correlate with proliferation

Although examination of ALDOA mRNA expression indicated it was enriched in progenitors we were unable to find a significant difference in ALDOA protein levels between $KI67^+$ and $KI67^-$ cells at 12 (Fig.7a), 14 (Fig.7b) and 16 PCW (Fig.7c) in the human cortex. To look for any fluctuation in ALDOA levels with the cell cycle we quantified immunofluorescence for KI67 and cell body ALDOA ''(nucleus and adjacent cell body) using the same analysis as that described above for KIF22, and found no correlation or discernible pattern at 12 (Fig.7e) (R^2= 0.018), 14 (Fig.7f) (R^2= 2e-4) or 16 PCW (Fig.7g) (R^2= 0.00992). These data show that in human cortex development, cellular ALDOA protein levels do not correlate with proliferation or fluctuate with cell cycle.

Previous work in different models demonstrated nuclear ALDOA level is greater in proliferating cells (Mamczur, Mazurek and Rakus, 2010; Mamczur *et al.*, 2013). To see if this was the case in human cortex development, we quantified nuclear ALDOA and KI67 (Fig.6e) but found no significant difference in nuclear ALDOA fluorescence between $KI67^+$ and $KI67^-$ cells at 12

(Fig.7i), 14 (Fig.7j) or 16 PCW (Fig.7k). We next tested if nuclear ALDOA levels in proliferating cells varied with cell cycle. Analysis of ALDOA and KI67 nuclear intensity established no correlation or pattern at 12 (Fig.7m) (R^2= 5e-05), 14 (Fig.7n) (R^2= 0.0365) or 16 PCW (Fig.7o) (R^2= 0.0723). This shows nuclear ALDOA levels do not increase with proliferation, nor fluctuate with cell cycle. We combined results across the 12, 14 and 16 PCW. There was no significant difference between $KI67^+$ and $KI67^-$ cells when examining ALDOA protein intensity in the whole cell (Fig.7d) or the nucleus (Fig.7i). Using the pooled data, there was no correlation or discernible pattern when nuclear ALDOA intensity was graphed against nuclear KI67 level for the whole cell (Fig.7h) (R^2= 0.0049) or the nucleus (Fig.7p) (R^2= 6e-04).

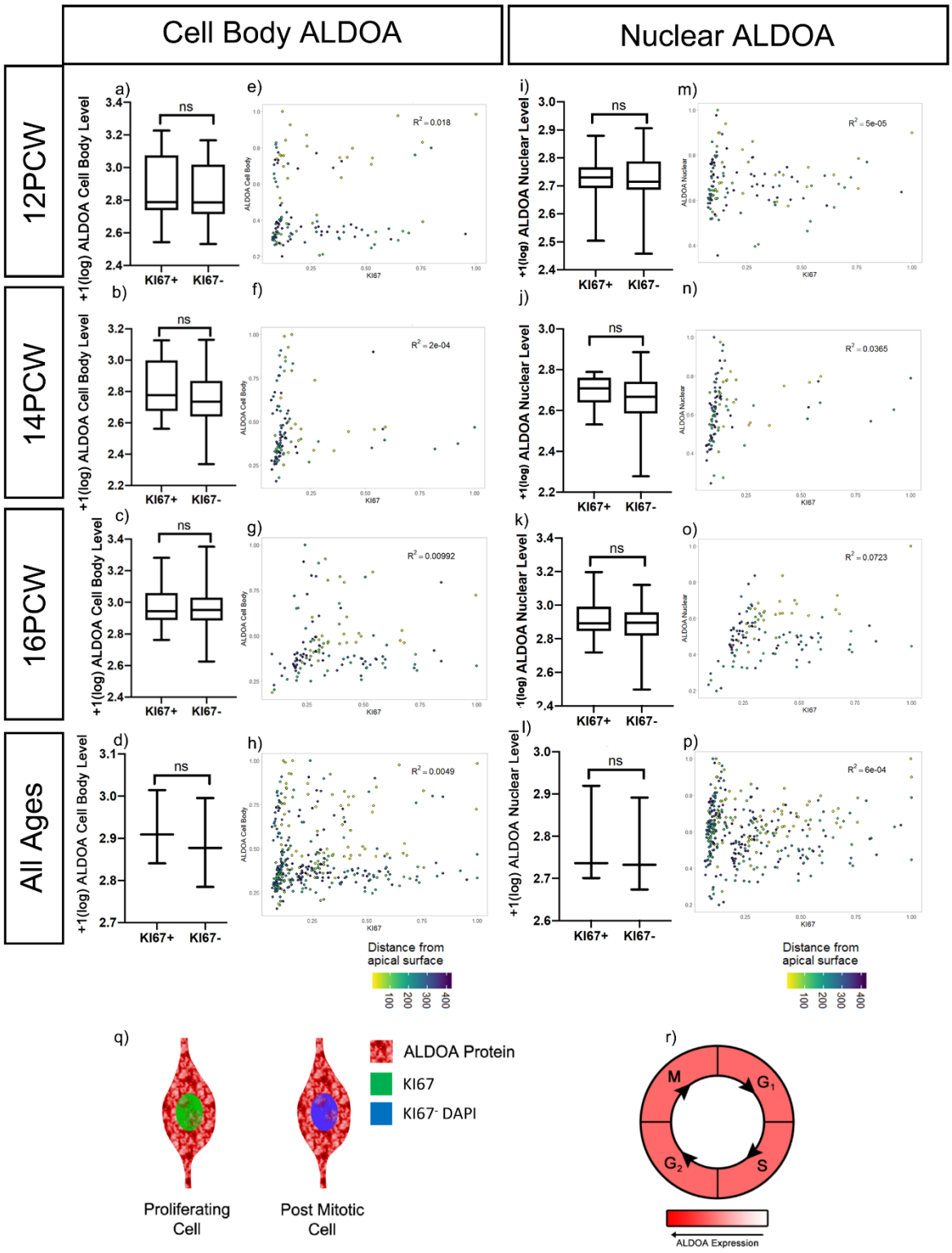

Cell Body ALDOA
Nuclear ALDOA
12PCW
14PCW
16PCW
All Ages
a)
b)
c)
d)
e)
f)
g)
h)
i)
j)
k)
l)
m)
n)
o)
p)
+1(log) ALDOA Cell Body Level
+1(log) ALDOA Nuclear Level
ns
KI67+
KI67-
R² = 0.018
R² = 2e-04
R² = 0.00992
R² = 0.0049
R² = 5e-05
R² = 0.0365
R² = 0.0723
R² = 6e-04
ALDOA Cell Body
ALDOA Nuclear
KI67
Distance from apical surface
100
200
300
400
q)
Proliferating Cell
Post Mitotic Cell
ALDOA Protein
KI67
KI67- DAPI
r)
M
G1
S
G2
ALDOA Expression

Figure 7: ALDOA protein quantification. a-d) Cell body ALDOA protein fluorescent intensity in KI67+ and KI67- cells at a) 12 PCW (raw data transformation =+1(log), bimodal distribution, Mann-Whitney test, p = 0.3702), b) 14 PCW (raw data transformation =+1(log), normal distribution, unpaired t-test with Welch's correction, p = 0.2032), c) 16 PCW (raw data transformation =+1(log), normal distribution, unpaired t-test with Welch's correction, p = 0.3523). d) ALDOA cell body protein fluorescent intensity in KI67+ and KI67- cells, 12, 14 and 16 PCW individual datasets averaged, (raw data transformation =+1(log)), paired t-test, p = 0.0836. e-h) ALDOA cellular protein intensity levels ls correlated to nuclear KI67 protein intensity at e)12, f)14 and g)16 PCW with distance from ventricular edge indicated. h) ALDOA whole cell protein intensity levels correlated to nuclear KI67 protein intensity pooled 12, 14, 16 PCW. i-l) Nuclear ALDOA protein fluorescent intensity in KI67+ and KI67- cells at i) 12 PCW (raw data transformation =+1(log), normal distribution, unpaired t-test with Welch's correction, p = 0.7543), j) 14 PCW (raw data transformation =+1(log), normal distribution, unpaired t-test with Welch's correction, p = 0.0694), k) 16 PCW (raw data transformation =+1(log), normal distribution, unpaired t-test with Welch's correction, p = 0.0772). l) ALDOA nuclear protein fluorescent intensity in KI67+ and KI67- cells, 12, 14 and 16 PCW individual datasets averaged, (raw data transformation =+1(log)), paired t-test, p = 0.1330. m-p) ALDOA nuclear protein intensity levels is correlated to nuclear KI67 protein intensity at m)12, n)14 and o)16 PCW with distance from ventricular edge indicated. p) ALDOA nuclear protein intensity levels correlated to nuclear KI67 protein intensity pooled 12, 14, 16 PCW. q) schematic demonstrating ALDOA protein is predominantly in the cytoplasm and lower in the nucleus in both KI67+ proliferating cells and KI67- post mitotic cells. r) model showing ALDOA levels do not change with the cell cycle.

Discussion

16p11.2 transcript expression during human neurogenesis.

The *16p11.2* CNV is a polygenic mutation that causes NDDs and the current study identified a number of the 29 *16p11.2* transcripts expressed in progenitor cells of the cerebral cortex. In addition to *ALDOA* and *KIF22* that are significantly enriched in progenitors, several other transcripts (e.g. *HIRIP3*, *PAGR1*, and *MAZ*) are also expressed in progenitors albeit at lower levels and are not significantly down-regulated as cells become post-mitotic. The simultaneous expression of multiple *16p11.2* genes in cells undergoing neurogenesis suggests that these cells may be particularly vulnerable to simultaneous alteration in their dosage as a consequence of the *16p11.2* microdeletion or microduplication. This lends support to the hypothesis that neurogenesis is disrupted in *16p11.2* CNV patients and that this contributes to subsequent development of NDDs.

In this study we focused on the expression of *16p11.2* genes in progenitors of the developing cerebral cortex that will give rise to excitatory neurons. However, *KIF22*, *ALDOA*, *HIRIP3*, *PAGR1*, and *MAZ* transcripts are also expressed in the ventricular zone of the ganglionic eminences where the progenitors of the inhibitory interneurons that subsequently migrate into the cerebral cortex reside. Simultaneous increase or decrease of these transcripts in interneuron progenitors resulting from the *16p11.2* CNV may therefore have an impact on the neurogenesis or differentiation of interneurons.

KIF22

KIF22 is a multifunctional protein that can regulate cell proliferation through at least two distinct mechanisms. First, KIF22 is a kinesin-like microtubule-based motor that binds microtubules and

chromosomes during mitosis and regulates mitotic spindle microtubule stability and symmetric/asymmetric cell division (Tokai *et al.*, 1996; Tokai-Nishizumi *et al.*, 2005; Sun and Hevner, 2014). Second, KIF22 regulates the expression of the cell-cycle regulator CDC25C. During cell division, CDC25C dephosphorylates CDK1, thus activating the CDK1-cyclinB complex while the CDK1-cyclin B complex phosphorylates CDC25C, causing an amplification loop to drive cells to mitosis (Nilsson and Hoffmann, 2000). KIF22 directly transcriptionally represses CDC25C and inhibits mitosis; this transcriptional repression of CDC25C is dependent on KIF22 being phosphorylated on Thr463 (Ohsugi *et al.*, 2003; Yu *et al.*, 2014). KIF22 depletion in a tumor cell line accelerates the G2/M transition and slows M/G1 transition (Yu *et al.*, 2014).

Overall, it therefore appears that KIF22 can act at several different points in the cell cycle making it difficult to predict how increased or decreased dosage of KIF22 in the *16p11.2* microduplication or microdeletion respectively would impact cell cycle in the specific context of cerebral cortex neural progenitors especially in light of the concomitant altered dosage of other *16p11.2* genes co-expressed with *KIF22* in progenitors. Our observation that *KIF22* mRNA and KIF22 protein levels both increase during the cell cycle to achieve highest levels in G2/M phase that drop as cells enter G1 phase implies that KIF22 protein does not persist for long after it is translated and is degraded at the end of M-phase suggesting that both transcriptional and post-transcriptional mechanisms regulate its levels. A clear outcome of our study is that KIF22 levels positively correlate with KI67 in neural progenitors and steadily rise as the cell progresses through G1>S>G2>M phases culminating in the maximum level during M-phase. One possibility is that KIF22 is required to reach a threshold for mitosis to occur, after which its levels must decrease sufficiently to allow mitotic exit. Whether cells undertake proliferative or neurogenic divisions is a process heavily controlled by cell cycle length (Borrell and Calegari, 2014). Perturbing *KIF22* gene dosage as a

consequence of the *16p11.2* CNV might affect the timing of KIF22 protein reaching this threshold in neural progenitors and therefore affect cell-cycle kinetics and perturb neurogenesis and neuronal output. Our results suggest the hypothesis that KIF22 regulates neurogenesis in the human developing cortex through cell-cycle regulation.

ALDOA

The process of brain development requires a vast and consistent supply of energy. Glucose is the predominant energy substrate for the fetal brain (Gustafsson, 2009), therefore efficient and controlled glycolysis is essential for normal brain development. ALDOA is required for the fourth step of glycolysis, conversion of fructose 1,6-biphospate to dihyroxyacetone phosphate and gluteraldehyde 3-phosphate. The metabolic role of cytoplasmic ALDOA is well established, and ALDOA also has other non-glycolytic "moonlighting" roles such as regulating mitochondrial function and cytoskeleton stability (Orosz, Christova and Ovadi, 1988; Pagliaro and Taylor, 1992; Kao *et al.*, 1999; Jewett and Sibley, 2003; Buscaglia *et al.*, 2006). In addition to its cytoplasmic roles, ALDOA has been observed in the nucleus (Mamczur and Dzugaj, 2008; Mamczur, Mazurek and Rakus, 2010; Mamczur *et al.*, 2013) where it has been suggested to impact cell cycle by positively regulating cyclin D1 expression to mediate G1/S progression (Ritterson Lew and Tolan, 2012; Fu *et al.*, 2018). Cell-culture studies show ALDOA sub-cellular localisation depends on the availability of energetic substrates, with addition of glucose driving ALDOA protein to the cytoplasm (Mamczur *et al.*, 2013). Therefore, it is likely the primary role for ALDOA is metabolic when cells require, and have available to them, large amounts of energy. The majority of ALDOA studies have used highly abnormal cancer tissue, or artificial cell culture systems in which

glycolytic enzymes have been shown to be increased (Ritterson Lew and Tolan, 2012; Mamczur *et al.*, 2013; Fu *et al.*, 2018; Pollen *et al.*, 2019). How these observations of ALDOA in a variety of systems relate to its role in human cerebral cortex development is unclear.

Altering ALDOA dosage in the developing brain will likely impact energy metabolism by altering the flow of metabolites through the glycolytic pathway and impacting subsequent pathways which feed on outputs of glycolysis. Disruption to energy metabolism during development has previously been linked to ASD and ADHD (Rash *et al.*, 2018). The offspring of hyperglycemic mice presented microcephaly, a phenocopy of the microcephaly observed in *16p11.2* microduplication patients (Rash *et al.*, 2018) and disruptions to energy metabolism may contribute to the microcephaly seen in the offspring of Zika infected mothers (Gilbert-Jaramillo *et al.*, 2019). No homozygous null ALDOA patients have been identified suggesting it is essential for life, but patients with changes to ALDOA levels have been identified; one patient with reduced ALDOA activity presented microcephaly (Kreuder *et al.*, 1996) and another presented intellectual disability (Beutler *et al.*, 1973). Of particular interest is the finding of schizophrenia patients with upregulated cortical ALDOA levels (Beasley *et al.*, 2006) and *16p11.2* microduplication is strongly associated with risk of schizophrenia. This information, coupled with our results that ALDOA is expressed in all cell types, make it clear that any changes to ALDOA dose will perturb energy metabolism at many stages in the brain, impacting its development.

ALDOA is much more abundant in the cytoplasm and we also found no clear relationship between cell body ALDOA levels and cell proliferation status. Nuclear ALDOA has been linked to cell proliferation (Mamczur, Mazurek and Rakus, 2010; Mamczur *et al.*, 2013; Fu *et al.*, 2018) but we found no clear relationship between nuclear ALDOA protein levels and cell proliferation status. While *ALDOA* mRNA levels are higher in proliferating cells compared to non-proliferating cells,

quantitative analysis of ALDOA protein revealed that ALDOA protein persists once cells exit mitosis. Therefore, while ALDOA protein is abundant in progenitor cells of the developing human cerebral cortex and so may play a role in neurogenesis phenotypes, the persistent expression of ALDOA protein as cells become post-mitotic argues against a specific role in neurogenesis and raises the additional possibility that ALDOA also plays roles in differentiated neurons.

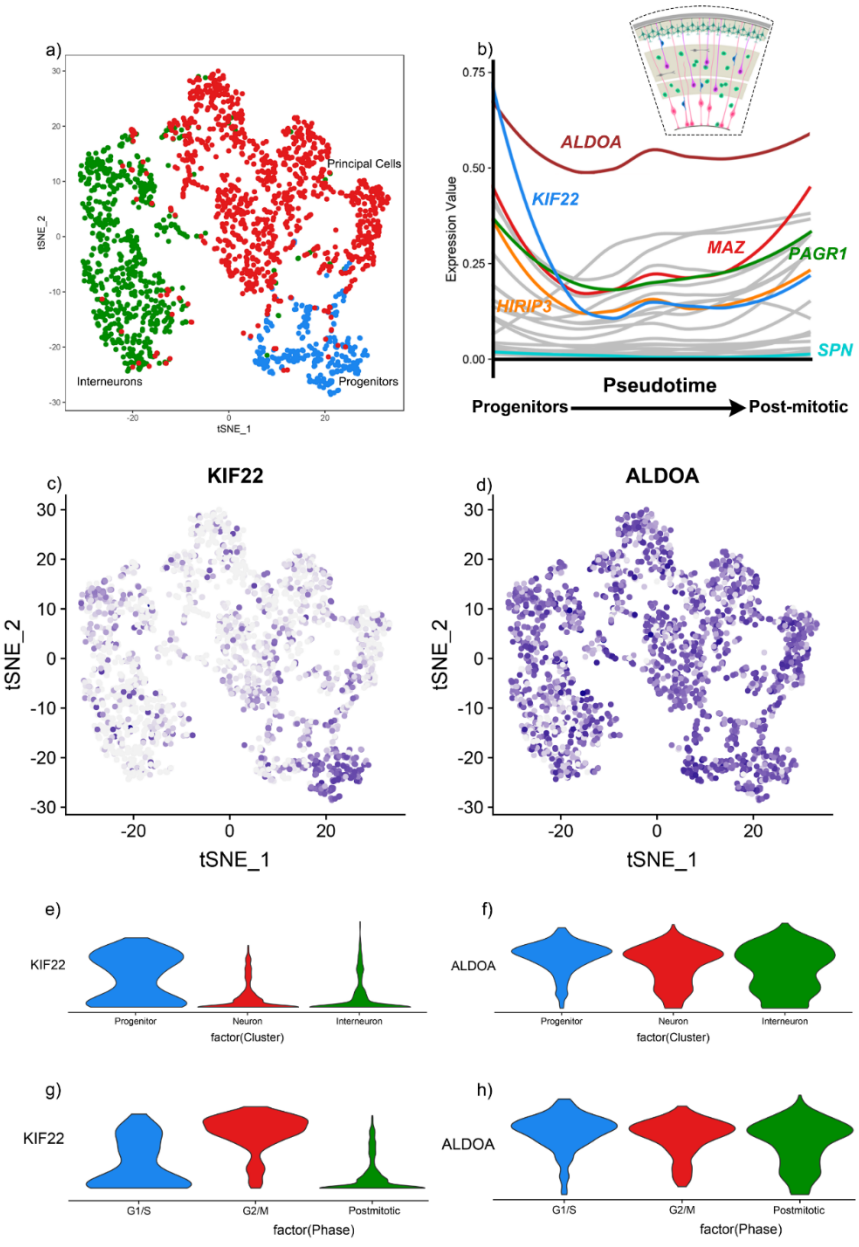
Conclusion

Our study of *16p11.2* gene expression in developing human fetal cerebral cortex indicates that altered dosage of *KIF22*, *ALDOA*, *HIRIP3*, *PAGR1*, and *MAZ* caused by the *16p11.2* microduplication or microdeletion may impact on neurogenesis in the developing human cortex and we identified *KIF22* being a strong candidate for having a specific role in neurogenesis. Further studies are required to unpick the mechanisms involved, but given the nature of the tissue, the scope for studying this *in vivo* is currently limited. However, growth of new model systems in which *16p11.2* gene expression can be manipulated such as human cerebral organoids will provide the opportunity address these questions.

Acknowledgements

This work was funded by BBSRC grant (BB/M00693X/1) and the Simons Initiative for the Developing Brain (SFARI - 529085). SM was supported by a BBSRC EASTBIO PhD studentship. The human embryonic and fetal material was provided by the Joint MRC / Wellcome (MR/R006237/1) Human Developmental Biology Resource (www.hdbr.org).

We thank Katherine Howe for her assistance with designing *in situ* probes and undergraduate student Emma Fowler (EF funded by a WR Henderson Scholarship) for her assistance with quantification



Supplementary Figure 1: Bioinformatics analysis of Zhong et al scRNA-seq dataset which included all regions of the telencephalic wall in samples from 8-36GW. a) tSNE clustering of cell types. b) changing mRNA expression levels of 16p11.2 genes as cells move from progenitors to neurons. Schematic of human cortex with dotted box indicates that all the regions of the cortex were used in this dataset. c) *KIF22* gradient plot. d) *ALDOA* gradient plot. e) Violin plots showing distribution of *KIF22* in different cell types. f) Violin plots showing distribution of *ALDOA* in different cell types. g) Violin plots showing distribution of *KIF22* at different cell cycle stages. h) Violin plots showing distribution of *ALDOA* at different cell cycle stages.

References

Bayer SA, Altman J. 2002. The Human Brain During the Late First Trimester. Taylor & Francis Group.

Bayer SA, Altman J. 2005. The Human Brain During The Second Trimester. CRC Press Taylor & Francis Group.

Beasley CL, Pennington K, Behan A, Wait R, Dunn MJ, Cotter D. 2006. Proteomic analysis of the anterior cingulate cortex in the major psychiatric disorders: Evidence for disease-associated changes. Proteomics. 6:3414–3425.

Beutler E, Scott S, Bishop A, Margolis N, Matsumoto F, Kuhl W. 1973. Red cell aldolase deficiency and hemolytic anemia: a new syndrome. Trans Assoc Am Physicians. 86:154–166.

Bijlsma EK, Gijsbers AC, Schuurs-Hoeijmakers JH, van Haeringen A, Fransen van de Putte DE, Anderlid BM, Lundin J, Lapunzina P, Perez Jurado LA, Delle Chiaie B, Loeys B, Menten B, Oostra A, Verhelst H, Amor DJ, Bruno DL, van Essen AJ, Hordijk R, Sikkema-Raddatz B, Verbruggen KT, Jongmans MC, Pfundt R, Reeser HM, Breuning MH, Ruivenkamp CA. 2009. Extending the phenotype of recurrent rearrangements of 16p11.2: deletions in mentally retarded patients without autism and in normal individuals. Eur J Med Genet. 52:77–87.

Blackmon K, Thesen T, Green S, Ben-Avi E, Wang X, Fuchs B, Kuzniecky R, Devinsky O. 2018. Focal Cortical Anomalies and Language Impairment in 16p11.2 Deletion and Duplication Syndrome. Cereb Cortex. 28:2422–2430.

617 Blaker-Lee A, Gupta S, McCammon JM, De Rienzo G, Sive H. 2012. Zebrafish homologs of
618 genes within 16p11.2, a genomic region associated with brain disorders, are active during
619 brain development, and include two deletion dosage sensor genes. Dis Model Mech. 5:834–
620 851.

621 Blumenthal I, Ragavendran A, Erdin S, Klei L, Sugathan A, Guide JR, Manavalan P, Zhou JQ,
622 Wheeler VC, Levin JZ, Ernst C, Roeder K, Devlin B, Gusella JF, Talkowski ME. 2014.
623 Transcriptional consequences of 16p11.2 deletion and duplication in mouse cortex and
624 multiplex autism families. Am J Hum Genet. 94:870–883.

625 Borrell V, Calegari F. 2014. Mechanisms of brain evolution: regulation of neural progenitor cell
626 diversity and cell cycle length. Neurosci Res. 86:14–24.

627 Budday S, Steinmann P, Ellen K. 2015. Physical biology of human brain development. Frontiers
628 in Cellular Neuroscience 9:257

629 Buscaglia CA, Penesetti D, Tao M, Nussenzweig V. 2006. Characterization of an aldolase-
630 binding site in the Wiskott-Aldrich syndrome protein. J Biol Chem. 281:1324–1331.

631 Bystron, I., Blakemore, C. and Rakic, P. 2008. Development of the human cerebral cortex:
632 Boulder Committee revisited. *Nat Rev Neurosci* 9, 110-122.
633 Clowry, G., Molnar, Z. and Rakic, P. (2010). Renewed focus on the developing human
634 neocortex. *J Anat* 217, 276-288.
635 de Anda FC, Rosario AL, Durak O, Tran T, Graff J, Meletis K, Rei D, Soda T, Madabhushi R,
636 Ginty DD, Kolodkin AL, Tsai LH. 2012. Autism spectrum disorder susceptibility gene
637 TAOK2 affects basal dendrite formation in the neocortex. Nat Neurosci. 15:1022–1031.

638 Escamilla, C. O., Filonova, I., Walker, A. K., Xuan, Z. X., Holehonnur, R., Espinosa, F., Liu, S.,
639 Thyme, S. B., Lopez-Garcia, I. A., Mendoza, D. B., et al. (2017). Kctd13 deletion
640 reduces synaptic transmission via increased RhoA. *Nature* 551, 227-231.
641
642 Fu H, Gao H, Qi X, Zhao L, Wu D, Bai Y, Li H, Liu X, Hu J, Shao S. 2018. Aldolase A
643 promotes proliferation and G1/S transition via the EGFR/MAPK pathway in non-small cell
644 lung cancer. Cancer Commun. 38:18.
645 Gilbert-Jaramillo J, Garcez P, James W, Molnár Z, Clarke K. 2019. The potential contribution of
646 impaired brain glucose metabolism to congenital Zika syndrome. J Anat. 235:468-480.
647 Girirajan S, Eichler EE. 2010. Phenotypic variability and genetic susceptibility to genomic
648 disorders. Hum Mol Genet. 19:R176–R187.
649 Golzio C, Willer J, Talkowski ME, Oh EC, Taniguchi Y, Jacquemont S, Reymond A, Sun M,
650 Sawa A, Gusella JF, Kamiya A, Beckmann JS, Katsanis N. 2012. KCTD13 is a major driver
651 of mirrored neuroanatomical phenotypes of the 16p11.2 copy number variant. Nature.
652 485:363–367.
653 Gustafsson J. 2009. Neonatal energy substrate production. Indian J Med Res. 130:618–623.
654 Horev G, Ellegood J, Lerch JP, Son YE, Muthuswamy L, Vogel H, Krieger AM, Buja A,
655 Henkelman RM, Wigler M, Mills AA. 2011. Dosage-dependent phenotypes in models of
656 16p11.2 lesions found in autism. Proc Natl Acad Sci U S A. 108:17076–17081.
657 Jewett TJ, Sibley LD. 2003. Aldolase forms a bridge between cell surface adhesins and the actin
658 cytoskeleton in apicomplexan parasites. Mol Cell. 11:885–894.

Kao AW, Noda Y, Johnson JH, Pessin JE, Saltiel AR. 1999. Aldolase mediates the association of F-actin with the insulin-responsive glucose transporter GLUT4. J Biol Chem. 274:17742–17747.

Kreuder J, Borkhardt A, Repp R, Pekrun A, Göttsche B, Gottschalk U, Reichmann H, Schachenmayr W, Schlegel K, Lampert F. 1996. Brief report: inherited metabolic myopathy and hemolysis due to a mutation in aldolase A. N Engl J Med. 334:1100–1104.

Kumar RA, KaraMohamed S, Sudi J, Conrad DF, Brune C, Badner JA, Gilliam TC, Nowak NJ, Cook EH Jr, Dobyns WB, Christian SL. 2008. Recurrent 16p11.2 microdeletions in autism. Hum Mol Genet. 17:628–638.

Levy D, Ronemus M, Yamrom B, Lee YH, Leotta A, Kendall J, Marks S, Lakshmi B, Pai D, Ye K, Buja A, Krieger A, Yoon S, Troge J, Rodgers L, Iossifov I, Wigler M. 2011. Rare de novo and transmitted copy-number variation in autistic spectrum disorders. Neuron. 70:886–897.

Ma T, Wang C, Wang L, Zhou X, Tian M, Zhang Q, Zhang Y, Li J, Liu Z, Cai Y, Liu F, You Y, Chen C, Campbell K, Song H, Ma L, Rubenstein JL, Yang Z. 2013. Subcortical origins of human and monkey neocortical interneurons. Nat Neurosci. 16:1588-97.

Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, Tirosh I, Bialas AR, Kamitaki N, Martersteck EM, Trombetta JJ, Weitz DA, Sanes JR, Shalek AK, Regev A, McCarroll SA. 2015. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell. 161:1202–1214.

Malhotra D, Sebat J. 2012. CNVs: harbingers of a rare variant revolution in psychiatric genetics. Cell. 148:1223–1241.

681 Mamczur P, Dzugaj A. 2008. Aldolase A is present in smooth muscle cell nuclei. Acta Biochim
682 Pol. 55:799–805.

683 Mamczur P, Gamian A, Kolodziej J, Dziegiel P, Rakus D. 2013. Nuclear localization of aldolase
684 A correlates with cell proliferation. Biochim Biophys Acta. 1833:2812–2822.

685 Mamczur P, Mazurek J, Rakus D. 2010. Ubiquitous presence of gluconeogenic regulatory
686 enzyme, fructose-1,6-bisphosphatase, within layers of rat retina. Cell Tissue Res. 341:213–
687 221.

688 McCarthy SE, Makarov V, Kirov G, Addington AM, McClellan J, Yoon S, Perkins DO, Dickel
689 DE, Kusenda M, Krastoshevsky O, Krause V, Kumar RA, Grozeva D, Malhotra D, Walsh
690 T, Zackai EH, Kaplan P, Ganesh J, Krantz ID, Spinner NB, Roccanova P, Bhandari A,
691 Pavon K, Lakshmi B, Leotta A, Kendall J, Lee Y-H, Vacic V, Gary S, Iakoucheva LM,
692 Crow TJ, Christian SL, Lieberman JA, Stroup TS, Lehtimäki T, Puura K, Haldeman-Englert
693 C, Pearl J, Goodell M, Willour VL, Derosse P, Steele J, Kassem L, Wolff J, Chitkara N,
694 McMahon FJ, Malhotra AK, Potash JB, Schulze TG, Nöthen MM, Cichon S, Rietschel M,
695 Leibenluft E, Kustanovich V, Lajonchere CM, Sutcliffe JS, Skuse D, Gill M, Gallagher L,
696 Mendell NR, Wellcome Trust Case Control Consortium, Craddock N, Owen MJ,
697 O'Donovan MC, Shaikh TH, Susser E, Delisi LE, Sullivan PF, Deutsch CK, Rapoport J,
698 Levy DL, King M-C, Sebat J. 2009. Microduplications of 16p11.2 are associated with
699 schizophrenia. Nat Genet. 41:1223–1227.

700 Miller I, Min M, Yang C, Tian C, Gookin S, Carter D, Spencer SL. 2018. Ki67 is a Graded
701 Rather than a Binary Marker of Proliferation versus Quiescence. Cell Rep. 24:1105–1112..

702 Nilsson I, Hoffmann I. 2000. Cell cycle regulation by the Cdc25 phosphatase family. Prog Cell

Cycle Res. 4:107–114.

Nuttle X, Giannuzzi G, Duyzend MH, Schraiber JG, Narvaiza I, Sudmant PH, Penn O, Chiatante G, Malig M, Huddleston J, Benner C, Camponeschi F, Ciofi-Baffoni S, Stessman HAF, Marchetto MCN, Denman L, Harshman L, Baker C, Raja A, Penewit K, Janke N, Tang WJ, Ventura M, Banci L, Antonacci F, Akey JM, Amemiya CT, Gage FH, Reymond A, Eichler EE. 2016. Emergence of a Homo sapiens-specific gene family and chromosome 16p11.2 CNV susceptibility. Nature. 536:205–209.

Ohsugi M, Tokai-Nishizumi N, Shiroguchi K, Toyoshima YY, Inoue J, Yamamoto T. 2003. Cdc2-mediated phosphorylation of Kid controls its distribution to spindle and chromosomes. EMBO J. 22:2091–2103.

Orosz F, Christova TY, Ovádi J. 1988. Modulation of phosphofructokinase action by macromolecular interactions. Quantitative analysis of the phosphofructokinase-aldolase-calmodulin system. Biochim Biophys Acta. 957:293–300.

Pagliaro L, Taylor DL. 1992. 2-Deoxyglucose and cytochalasin D modulate aldolase mobility in living 3T3 cells. J Cell Biol. 118:859–863.

Pollen AA, Bhaduri A, Andrews MG, Nowakowski TJ, Meyerson OS, Mostajo-Radji MA, Di Lullo E, Alvarado B, Bedolli M, Dougherty ML, Fiddes IT, Kronenberg ZN, Shuga J, Leyrat AA, West JA, Bershteyn M, Lowe CB, Pavlovic BJ, Salama SR, Haussler D, Eichler EE, Kriegstein AR. 2019. Establishing Cerebral Organoids as Models of Human-Specific Brain Evolution. Cell. 176:743–756..

Pollen AA, Nowakowski TJ, Chen J, Retallack H, Sandoval-Espinosa C, Nicholas CR, Shuga J, Liu SJ, Oldham MC, Diaz A, Lim DA, Leyrat AA, West JA, Kriegstein AR. 2015.

725 Molecular identity of human outer radial glia during cortical development. Cell. 163:55–67.

726 Preibisch S, Saalfeld S, Tomancak P. 2009. Globally optimal stitching of tiled 3D microscopic
727 image acquisitions. Bioinformatics. 25:1463–1465.

728 Pucilowska J, Vithayathil J, Pagani M, Kelly C, Karlo JC, Robol C, Morella I, Gozzi A,
729 Brambilla R, Landreth GE. 2018. Pharmacological Inhibition of ERK Signaling Rescues
730 Pathophysiology and Behavioral Phenotype Associated with 16p11.2 Chromosomal
731 Deletion in Mice. J Neurosci. 38:6640–6652.

732 Pucilowska J, Vithayathil J, Tavares EJ, Kelly C, Karlo JC, Landreth GE. 2015. The 16p11.2
733 deletion mouse model of autism exhibits altered cortical progenitor proliferation and brain
734 cytoarchitecture linked to the ERK MAPK pathway. J Neurosci. 35:3190–3200.

735 Qiu X, Hill A, Packer J, Lin D, Ma Y-A, Trapnell C. 2017. Single-cell mRNA quantification and
736 differential analysis with Census. Nat Methods. 14:309–315.

737 Qureshi AY, Mueller S, Snyder AZ, Mukherjee P, Berman JI, Roberts TP, Nagarajan SS, Spiro
738 JE, Chung WK, Sherr EH, Buckner RL. 2014. Opposing brain differences in 16p11.2
739 deletion and duplication carriers. J Neurosci. 34:11199–11211.

740 Radonjić NV, Ortega JA, Memi F, Dionne K, Jakovcevski I, Zecevic N. 2014. The complexity of
741 the calretinin-expressing progenitors in the human cerebral cortex. Front Neuroanat. 8:82.

742 Rash BG, Micali N, Huttner AJ, Morozov YM, Horvath TL, Rakic P. 2018. Metabolic regulation
743 and glucose sensitivity of cortical radial glial cells. Proc Natl Acad Sci U S A. 115:10142–
744 10147.

745 Richter M, Murtaza N, Scharrenberg R, White SH, Johanns O, Walker S, Yuen RKC, Schwanke

B, Bedürftig B, Henis M, Scharf S, Kraus V, Dörk R, Hellmann J, Lindenmaier Z, Ellegood J, Hartung H, Kwan V, Sedlacik J, Fiehler J, Schweizer M, Lerch JP, Hanganu-Opatz IL, Morellini F, Scherer SW, Singh KK, Calderon de Anda F. Altered TAOK2 activity causes autism-related neurodevelopmental and cognitive abnormalities through RhoA signaling. 2019. Mol Psychiatry. 24:1329-1350.

Ritterson Lew C, Tolan DR. 2012. Targeting of several glycolytic enzymes using RNA interference reveals aldolase affects cancer cell proliferation through a non-glycolytic mechanism. J Biol Chem. 287:42554–42563.

Rosenfeld JA, Coppinger J, Bejjani BA, Girirajan S, Eichler EE, Shaffer LG, Ballif BC. 2010. Speech delays and behavioral problems are the predominant features in individuals with developmental delays and 16p11.2 microdeletions and microduplications. J Neurodev Disord. 2:26–38.

Sanders SJ, Ercan-Sencicek AG, Hus V, Luo R, Murtha MT, Moreno-De-Luca D, Chu SH, Moreau MP, Gupta AR, Thomson SA, Mason CE, Bilguvar K, Celestino-Soper PB, Choi M, Crawford EL, Davis L, Wright NR, Dhodapkar RM, DiCola M, DiLullo NM, Fernandez TV, Fielding-Singh V, Fishman DO, Frahm S, Garagaloyan R, Goh GS, Kammela S, Klei L, Lowe JK, Lund SC, McGrew AD, Meyer KA, Moffat WJ, Murdoch JD, O'Roak BJ, Ober GT, Pottenger RS, Raubeson MJ, Song Y, Wang Q, Yaspan BL, Yu TW, Yurkiewicz IR, Beaudet AL, Cantor RM, Curland M, Grice DE, Gunel M, Lifton RP, Mane SM, Martin DM, Shaw CA, Sheldon M, Tischfield JA, Walsh CA, Morrow EM, Ledbetter DH, Fombonne E, Lord C, Martin CL, Brooks AI, Sutcliffe JS, Cook EH Jr, Geschwind D, Roeder K, Devlin B, State, M. W. 2011. Multiple recurrent de novo CNVs, including

768 duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism.
769 Neuron. 70:863–885.

770 Scholzen T, Gerdes J. 2000. The Ki-67 protein: from the known and the unknown. J Cell
771 Physiol. 182:311–322.

772 Shinawi M, Liu P, Kang SH, Shen J, Belmont JW, Scott DA, Probst FJ, Craigen WJ, Graham
773 BH, Pursley A, Clark G, Lee J, Proud M, Stocco A, Rodriguez DL, Kozel BA, Sparagana S,
774 Roeder ER, McGrew SG, Kurczynski TW, Allison LJ, Amato S, Savage S, Patel A,
775 Stankiewicz P, Beaudet AL, Cheung SW, Lupski JR. 2010. Recurrent reciprocal 16p11.2
776 rearrangements associated with global developmental delay, behavioural problems,
777 dysmorphism, epilepsy, and abnormal head size. J Med Genet. 47:332–341.

778 Sun T, Hevner RF. 2014. Growth and folding of the mammalian cerebral cortex: from molecules
779 to malformations. Nat Rev Neurosci. 15:217–232.

780 Tai DJC, Ragavendran A, Manavalan P, Stortchevoi A, Seabra CM, Erdin S, Collins RL,
781 Blumenthal I, Chen X, Shen Y, Sahin M, Zhang C, Lee C, Gusella JF, Talkowski ME.
782 2016. Engineering microdeletions and microduplications by targeting segmental
783 duplications with CRISPR. Nat Neurosci. 19:517–522.

784 Tirosh I, Izar B, Prakadan SM, Wadsworth MH 2nd, Treacy D, Trombetta JJ, Rotem A, Rodman
785 C, Lian C, Murphy G, Fallahi-Sichani M, Dutton-Regester K, Lin J-R, Cohen O, Shah P, Lu
786 D, Genshaft AS, Hughes TK, Ziegler CGK, Kazer SW, Gaillard A, Kolb KE, Villani A-C,
787 Johannessen CM, Andreev AY, Van Allen EM, Bertagnolli M, Sorger PK, Sullivan RJ,
788 Flaherty KT, Frederick DT, Jané-Valbuena J, Yoon CH, Rozenblatt-Rosen O, Shalek AK,
789 Regev A, Garraway LA. 2016. Dissecting the multicellular ecosystem of metastatic

melanoma by single-cell RNA-seq. Science. 352:189–196.

Tokai N, Fujimoto-Nishiyama A, Toyoshima Y, Yonemura S, Tsukita S, Inoue J, Yamamota T. 1996. Kid, a novel kinesin-like DNA binding protein, is localized to chromosomes and the mitotic spindle. EMBO J. 15:457–467.

Tokai-Nishizumi N, Ohsugi M, Suzuki E, Yamamoto T. 2005. The chromokinesin Kid is required for maintenance of proper metaphase spindle size. Mol Biol Cell. 16:5455–5463.

Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, Lennon NJ, Livak KJ, Mikkelsen TS, Rinn JL. 2014. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotechnol. 32:381–386.

Ultanir, S. K., Yadav, S., Hertz, N. T., Oses-Prieto, J. A., Claxton, S., Burlingame, A. L., Shokat, K. M., Jan, L. Y. and Jan, Y. N. (2014). MST3 kinase phosphorylates TAO1/2 to enable Myosin Va function in promoting spine synapse development. *Neuron* 84, 968-982.

Weiss LA, Shen Y, Korn JM, Arking DE, Miller DT, Fossdal R, Saemundsen E, Stefansson H, Ferreira MA, Green T, Platt OS, Ruderfer DM, Walsh CA, Altshuler D, Chakravarti A, Tanzi RE, Stefansson K, Santangelo SL, Gusella JF, Sklar P, Wu BL, Daly MJ. 2008. Association between microdeletion and microduplication at 16p11.2 and autism. N Engl J Med. 358:667–675.

Yadav, S., Oses-Prieto, J. A., Peters, C. J., Zhou, J., Pleasure, S. J., Burlingame, A. L., Jan, L. Y. and Jan, Y. N. (2017). TAOK2 Kinase Mediates PSD95 Stability and Dendritic Spine Maturation through Septin7 Phosphorylation. Neuron 93, 379-393

Yu Y, Wang XY, Sun L, Wang YL, Wan YF, Li XQ, Feng YM. 2014. Inhibition of KIF22 suppresses cancer cell proliferation by delaying mitotic exit through upregulating CDC25C

expression. Carcinogenesis. 35:1416–1425.

Zhong, S., Zhang, S., Fan, X., Wu, Q., Yan, L., Dong, J., Zhang, H., Li, L., Sun, L., Pan, N. , Xu X, Tang F, Zhang J, Qiao J, Wang X.. (2018). A single-cell RNA-seq survey of the developmental landscape of the human prefrontal cortex. *Nature* 555, 524-528.

Zufferey F, Sherr EH, Beckmann ND, Hanson E, Maillard AM, Hippolyte L, Mace A, Ferrari C,Kutalik Z, Andrieux J, Aylward E, Barker M, Bernier R, Bouquillon S, Conus P, Delobel B, Faucett WA, Goin-Kochel RP, Grant E, Harewood L, Hunter JV, Lebon S, Ledbetter DH, Martin CL, Mannik K, Martinet D, Mukherjee P, Ramocki MB, Spence SJ, Steinman KJ, Tjernagel J, Spiro JE, Reymond A, Beckmann JS, Chung WK, Jacquemont S. 2012. A 600 kb deletion syndrome at 16p11.2 leads to energy imbalance and neuropsychiatric disorders. J Med Genet. 49:660–668