



OPEN

## High-resolution transcriptional landscape of xeno-free human induced pluripotent stem cell-derived cerebellar organoids

Samuel Naylor<sup>1</sup>✉, Devika Agarwal<sup>2</sup>, Fabiola Curion<sup>3</sup>, Rory Bowden<sup>3,4</sup> & Esther B. E. Becker<sup>1,5</sup>✉

Current protocols for producing cerebellar neurons from human pluripotent stem cells (hPSCs) often rely on animal co-culture and mostly exist as monolayers, limiting their capability to recapitulate the complex processes in the developing cerebellum. Here, we employed a robust method, without the need for mouse co-culture to generate three-dimensional cerebellar organoids from hPSCs that display hallmarks of *in vivo* cerebellar development. Single-cell profiling followed by comparison to human and mouse cerebellar atlases revealed the presence and maturity of transcriptionally distinct populations encompassing major cerebellar cell types. Encapsulation with Matrigel aimed to provide more physiologically-relevant conditions through recapitulation of basement-membrane signalling, influenced both growth dynamics and cellular composition of the organoids, altering developmentally relevant gene expression programmes. We identified enrichment of cerebellar disease genes in distinct cell populations in the hPSC-derived cerebellar organoids. These findings ascertain xeno-free human cerebellar organoids as a unique model to gain insight into cerebellar development and its associated disorders.

The cerebellum has a major role in controlling sensorimotor functions but is also increasingly implicated in higher-order cognitive functions including language, emotion and reward. Accordingly, dysfunction of the cerebellum is associated with both motor diseases such as ataxia, dystonia and tremor, but also non-motor disorders<sup>1,2</sup>. Specifically, abnormal cerebellar development is increasingly linked to neurodevelopmental disorders including autism spectrum disorder and intellectual disability<sup>3,4</sup>. The human cerebellum is one of the first structures of the brain to differentiate and continues to develop until several years after birth. This protracted development makes the cerebellum especially vulnerable to insult. However, the molecular mechanisms underlying the physiological development of the human cerebellum and how these go awry in developmental disorders remain incompletely understood.

The cerebellum follows a highly stereotyped developmental programme to form a well-organized laminar structure, containing one of the largest and most metabolically active neuronal types (the Purkinje neuron (PN)), as well as the most abundant neuron in the brain (the granule cell (GC))<sup>5</sup>. During early development, the isthmus organizer at the mid-hindbrain (MHB) boundary in the dorsal region of rhombomere 1 induces the formation of the cerebellum through the expression of the lineage-specific transcription factors including OTX2, GBX2, EN1/2, and PAX2<sup>6,7</sup>. Unique to cerebellar development, precursors arise from two distinct germinal zones in rhombomere 1; the ventricular zone (VZ) gives rise to GABAergic neurons (PNs, interneurons, GABAergic deep cerebellar nuclei (DCN) neurons), whereas all glutamatergic neurons (GCs, unipolar brush cells, glutamatergic DCN neurons) are generated at the rhombic lip (RL). The fourth ventricle, roof plate and subsequent hindbrain choroid plexus are responsible for secretion of growth factors required for development of adjacent regions of the cerebellum, including bone morphogenetic proteins (BMPs) to direct RL formation and Sonic Hedgehog (SHH) for VZ precursor differentiation<sup>8</sup>. Regulatory mechanisms driving cerebellar cell-fate specification have

<sup>1</sup>Department of Physiology, Anatomy and Genetics, University of Oxford, Oxford OX1 3PT, United Kingdom. <sup>2</sup>Weatherall Institute for Molecular Medicine, University of Oxford, Oxford OX3 7BN, United Kingdom. <sup>3</sup>Wellcome Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, United Kingdom. <sup>4</sup>Present address: Walter and Eliza Hall Institute of Medical Research, Parkville, Victoria 3052, Australia. <sup>5</sup>Present address: Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford OX3 9DU, United Kingdom. ✉email: Samuel.Naylor@dpag.ox.ac.uk; Esther.Becker@ndcn.ox.ac.uk

been studied in various animal models; however, observing cerebellar development in humans poses several significant challenges, including acquisition of relevant material to study. Human induced pluripotent stem cells (hiPSCs) offer a powerful alternative model system to gain insight into human cerebellar development. Various methodologies have recently been developed to differentiate hPSCs into cerebellar neurons through the addition of growth factors that reproduce early patterning events in vitro, recapitulating MHB barrier establishment and subsequent formation of polarized cerebellar tissue characteristic of RL/VZ expansion<sup>9–14</sup>. Current models have mainly focussed on the differentiation of hPSC-derived PNs and subsequent maturation by co-culture with mouse cells, which hampers downstream analyses and has implications for any future potential clinical use. Moreover, the potential of current models to study other developing cerebellar cell populations has not been fully explored yet.

Here, we employ a reproducible method to generate hiPSC-derived self-organising three-dimensional organoids without the need for mouse co-culture. Using single-cell RNA sequencing (scRNA-seq), we demonstrate the presence of transcriptionally distinct populations representative of the majority of neuronal cell types in the developing cerebellum and confirm their identity and maturity through comparison with murine and human cerebellar scRNA-seq data<sup>15,16</sup>. We also explore the molecular consequences resulting from the addition of Matrigel to developing cerebellar organoids aimed to simulate basement membrane signalling. Basement membrane ligands have a well-known role in initiating developmentally-related signalling<sup>17</sup>. The addition of extra cellular matrix (ECM) molecules designed to recapitulate elements of the basement membrane is a popular method for improving culture conditions for cortical organoids<sup>18</sup>, but has not been investigated in cerebellar culture models. Speculation exists that the basement membrane may contribute to local, epithelial-mediated mechanical stiffness which allows basal constriction of boundary cells engaging in anterior versus posterior movement<sup>19</sup>. Laminin-dependent basal constriction has been shown to be involved in MHB specification<sup>20</sup>. Here, we show that embedding in Matrigel results in altered growth dynamics of the cerebellar organoids and influences organoid composition with a biased lineage commitment towards RL. Finally, we identify cell populations in the hiPSC-derived cerebellar organoids that are enriched for disease genes associated with cerebellar disorders. Together, our study shows that hiPSC-derived cerebellar organoids provide a xeno-free culture system that is ideally suited to model human cerebellar development and its associated disorders.

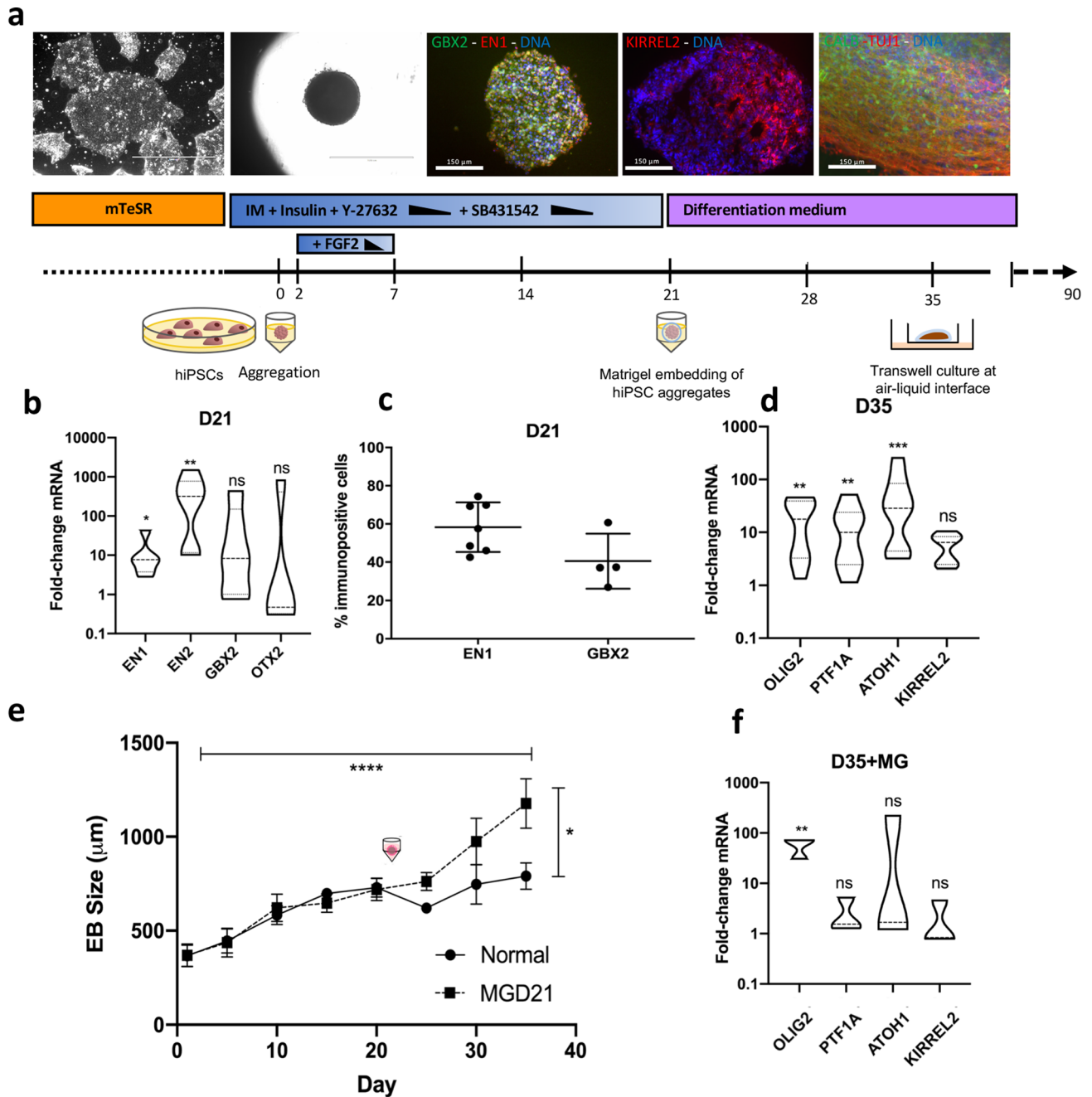
## Results

### Embryoid body differentiation recapitulates the formation of the isthmic organizer followed by RL/VZ expansion.

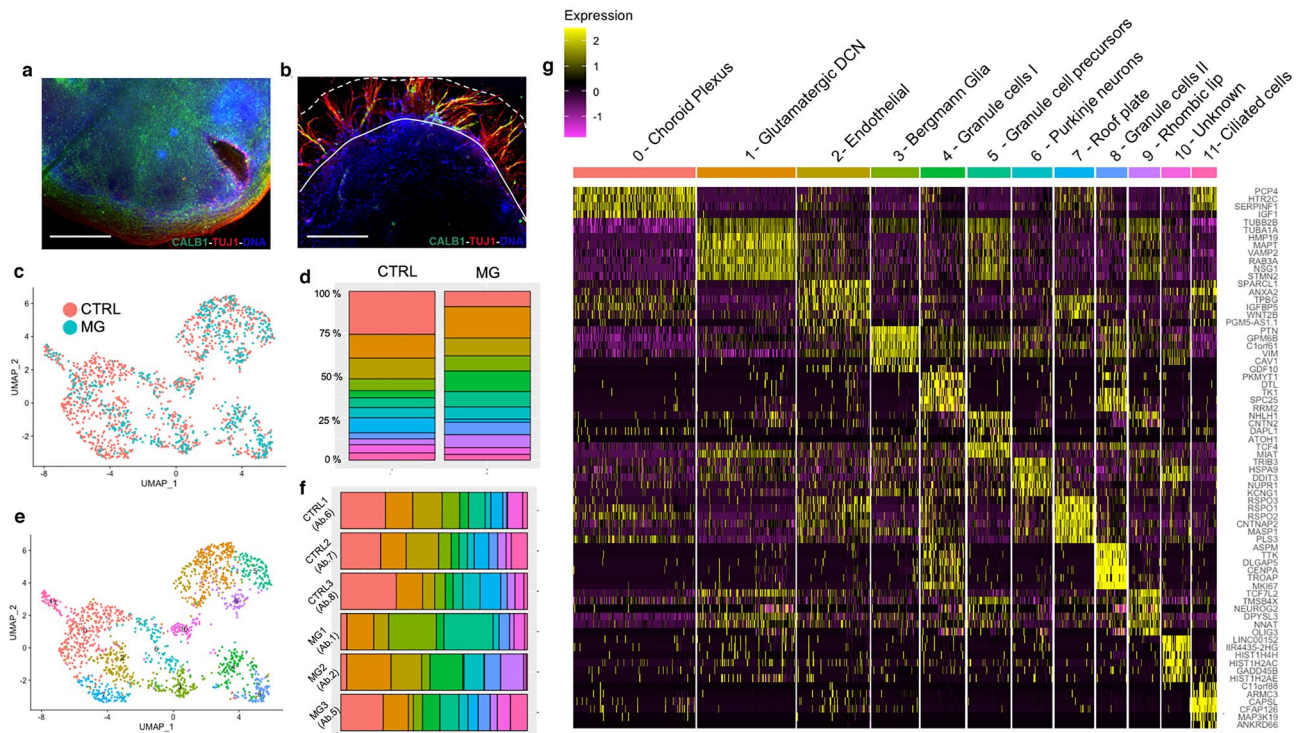
To generate cerebellar organoids from hiPSCs, we employed a previously established methodology aimed to mimic the self-inductive properties of the isthmic organizer<sup>9–11</sup>. However, instead of dissociating cultures at day 35 and subsequent co-culture with mouse progenitors, we continued the human-only cultures in a three-dimensional format at air–liquid interface (Fig. 1a). Re-aggregated hiPSCs were treated with a combination of fibroblast growth factor 2 (FGF2) and Insulin to mimic the self-inductive properties of the isthmic organizer and reproduce MHB development<sup>9,11</sup>. We refer to cellular aggregates as embryoid bodies (EBs) until the appearance of polarized neuroepithelial tissue and confirmation of MHB identity at day 21, from which point they are referred to as organoids. Successful acquisition of MHB fate acquisition was assessed by measuring expression of key lineage-specific transcription factors. At day 21, induction of the MHB markers *EN1* ( $7.66 \pm 5.05$ -fold-change, p-value 0.0346), *EN2* ( $318.91 \pm 183.45$ , p-value 0.0087), and *GBX2* ( $8.25 \pm 34.71$ , p-value 0.1147) was observed, at the expense of the anterior marker *OTX2* ( $0.47 \pm 73.66$ , p-value 0.6429) (Fig. 1b). Robust MHB acquisition was further confirmed by immunostaining; at day 21, 58.31  $\pm$  12.98% cells were positive for EN1 and 40.52  $\pm$  14.31% for GBX2 (Fig. 1a,c; Fig. 1a shows a representative dual-stained cryosection, where the majority of nuclei are immunoreactive for both markers). At day 35, acquisition of VZ/RL identity was confirmed by measurement of the VZ markers *OLIG2* ( $17.74 \pm 7.09$  fold-change, p-value 0.0023), *PTF1A* ( $9.96 \pm 5.36$ , p-value 0.0072), and *KIRREL2* ( $6.43 \pm 1.91$ , p-value 0.4211), and the RL marker *ATOH1* ( $28.53 \pm 19.39$ , p-value 0.0004) (Fig. 1d). Consistently, immunostaining of cryosectioned organoids at 35 revealed the emergence of extensive pockets of polarized neuroepithelium showing both broad-membrane and apical accumulation of KIRREL2, one of the earliest markers of VZ precursors, and PAX6, which is expressed in RL-derivatives (Fig. 1a, Supplementary Figure S1a,b,c). Following day 60, organoids exhibited robust expression of the neuronal markers TUJ1 and Calbindin, indicating the maturation of VZ-derived neurons (Fig. 1a, Supplementary Figure S2a). The majority of Calbindin-positive cells were observed at the periphery of the organoids and exhibited bipolar morphology consistent with maturing PN progenitors (Supplementary Figure S2a). We did not observe CALBINDIN-positive cells exhibiting the characteristic elaborate morphology of mature PNs that have been described in co-culture models<sup>9</sup>.

### Encapsulation with Matrigel transforms growth of organoids.

Matrigel has previously been used to simulate basement-membrane signalling and improve the growth of cortical organoids<sup>18</sup>. However, its effects on the growth of cerebellar organoids have not been investigated. To evaluate this we encapsulated EBs following early neuronal lineage commitment, as marked by appearance of polarized neuroepithelia or ‘rosettes’, in an effort to supply local factors which influence MHB development<sup>20</sup>. Consistent with other reports<sup>9,10</sup>, we observed small, round rosettes and larger ovoid-shaped structures (Supplementary Figure S2b). Given the role of basement membrane molecules in MHB formation<sup>19,20</sup> we chose to embed at day 21 following the acquisition of MHB markers. Following embedding at day 21, the average diameter of embedded organoids at day 35 was  $1047.37 \pm 367.47$   $\mu$ m in contrast to  $752.71 \pm 168.73$   $\mu$ m in unencapsulated organoids, indicating that the addition of Matrigel to the culture setting results in enhanced organoid growth (p-value 0.0101) (Fig. 1e). In addition, excessive neuronal outgrowth was observed at the periphery of the organoids (Supplementary Figure S1a). The addition of Matrigel resulted in a significant increase in *OLIG2* mRNA expression (p-value 0.0044); however,



**Figure 1.** Cerebellar differentiation protocol recapitulating MHB formation and subsequent RL/VZ expansion. **(a)** Overview of cerebellar differentiation protocol. Representative images of key stages are shown above (from left to right); undifferentiated iPSCs, day 21 EBs, day 21 EB cryosections immunostained for GBX2 (green) and EN1 (red), with nuclear Hoechst staining (blue), Day 35 organoid cryosections immunostained for KIRREL2 (red), with nuclear Hoechst staining (blue), and Day 60 organoids immunostained for Calbindin (green), TUJ1 (red), with nuclear Hoechst staining (blue). Scale bars are 1000 µm (first two panels) and 150 µm (all other panels). **(b)** mRNA levels of day 21 EBs displayed as fold-change values relative to undifferentiated iPSCs. Samples were calibrated relative to *ACTB*. *EN1* (p-value 0.0346), *GBX2* (p-value 0.1147), *EN2* (p-value 0.0087) and *OTX2* (p-value 0.6429). Results are shown from 5–6 independent differentiation experiments, Mann–Whitney test. **(c)** Quantification of EN1 or GBX2 (single-channel stain) immunostaining in day 21 EB cryosections as assessed by percentage of immunoreactive nuclei divided by total nuclei. Graphs show mean with SD from 4–7 independent differentiation experiments. **(d)** Day 35 mRNA levels of *OLIG2* (p-value 0.0023), *PTF1A* (p-value 0.0072), *ATOH1* (p-value 0.0004), and *KIRREL2* (p-value 0.4211). Violin plots show median and interquartile range. Data is shown from 8–9 independent differentiation experiments tested for significance using unpaired Mann–Whitney test. **(e)** Growth rate of organoids was significant over the course of differentiation (p-value 0.0001) and significantly increased by the addition of Matrigel (MG) from day 21 onwards (p-value 0.0101) as assessed by two-way ANOVA, Dunnet multiple comparison test. Error-bars represent mean/SEM. **(f)** Matrigel encapsulation resulted in a significant increase in *OLIG2* mRNA (p-value 0.0044), but non-significant changes in *PTF1A* (p-value 0.0857), *ATOH1* (p-value 0.2527) and *KIRREL2* (p-value 0.4893). Results are shown from three differentiation experiments, unpaired Mann–Whitney test.



**Figure 2.** Characterisation of cerebellar organoids reveals population heterogeneity. **(a,b)** Representative images of unembedded (CTRL) organoid **(a)** and Matrigel (MG)-embedded organoid **(b)** prior to harvesting. Lines demarcate the zone of Matrigel, which surrounds the organoid. Organoids immunostained for Calbindin (green) and TUJ1 (red). Nuclei were stained with Hoechst (blue). Scale bars are 500 µm. **(c)** UMAP plot after canonical correlation analysis (CCA) illustrates representation of cells across treatment conditions (pink—control, teal—MG treatment from day 21). **(d)** Visual depiction of organoid heterogeneity, colored by population, shows representation of cell types across treatment groups. Left: control (CTRL), right: MG-embedded. **(e)** UMAP plot of all cells showing 12 populations following separation after CCA at resolution 0.8. **(f)** Visual depiction of cohort heterogeneity, per individually hash-labelled organoid, colored by population, shows representation of cell types across individually barcoded organoids. Further information is given in Supplementary Figure S5. **(g)** Heatmap depicting the top six marker genes per cell population across all groups, filtered by log fold-change values. Heatmap made with Seurat (v3.2.1).

it did not yield a significant difference in the expression of *PTF1A* (p-value 0.0857), *ATOH1* (p-value 0.2527) or *KIRREL2* (p-value 0.4893) (Fig. 1f). Extensive and uniform expression of *KIRREL2* and *PAX6* was observed in both control and embedded aggregates on day 35 of differentiation, indicating no obvious change to VZ or RL lineage-commitment upon Matrigel treatment (Supplementary Figure S1a). Upon culture of organoids past 35 days we observed necrotic or hollow pockets, possibly indicative of a lack of oxygen and/or nutrient diffusion. In order to maintain adequate oxygen supply, we continued to culture the organoids on transwell membranes. This is conventionally done with organotypic cerebellar slice cultures and was recently shown to improve neuronal survival in cortical organoids<sup>21</sup>.

### Individual organoid hashing and single-cell profiling of organoids reveals distinct cerebellar cell types.

Organoids were allowed to mature for an additional 55 days at the air–liquid interface on transwell membranes before transcriptional profiling using scRNA-seq. Six organoids, each labelled with unique oligonucleotide hashtags, were sequenced as two hash-pools (2526 total cells prior to QC). Each hash-pool consisted of three control and three Matrigel-embedded organoids, harvested on day 90 of culture. Representative images are shown of organoids following culture on transwells with and without Matrigel embedding at day 21 (Fig. 2a,b). Given recent findings regarding the disparate and rapidly changing cell type-specific metabolic activity across developing neurons<sup>22</sup>, conventional mitochondrial gene content QC cut-offs were discarded in favour of including all singlets regressed for mitochondrial content. Mitochondrial gene expression, unique molecular identifiers (UMIs) and gene counts are shown in Supplementary Figure S3, indicating near uniform mitochondrial and gene content across samples and hash pools, respectively.

Following Louvain clustering, representation of equivalent cell populations was evident in both treatment groups following integration (Fig. 2c). Optimal clustering to distinguish between cerebellar cell types was determined by visualizing separation of related clusters in pooled data at increasing resolutions using the ClusTree package (Supplementary Figure S4) as well as by manual curation of the cluster biomarkers with canonical cerebellar markers. Following integration of treatment cohorts to identify common sources of biological variation,

UMAP projection of canonical correlates at a resolution of 0.8 distinguished 12 distinct populations mutually occupying common transcriptional space (Fig. 2e,d).

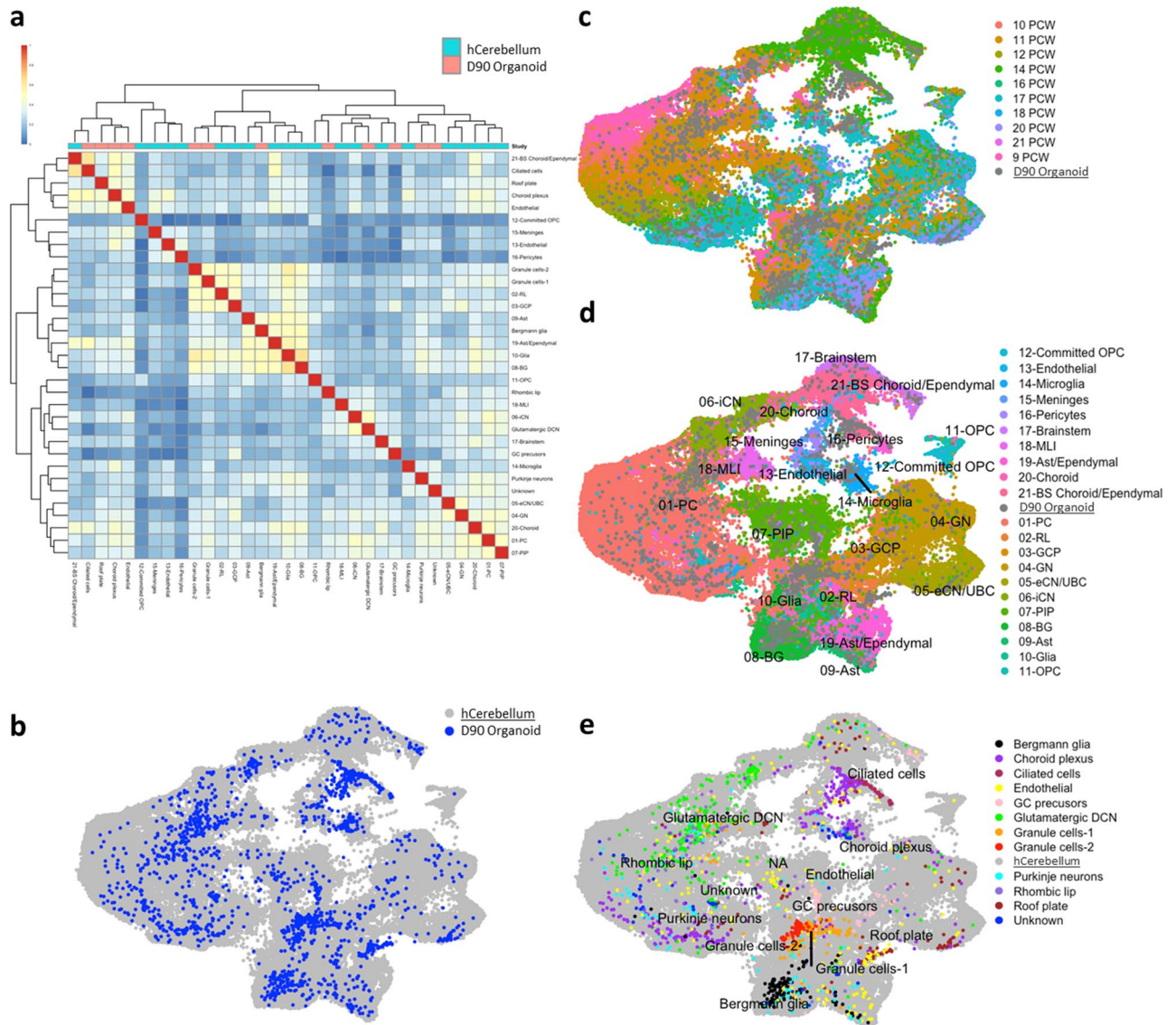
Individual organoid hash-labelling allowed us to chart the population composition per organoid. The majority of populations were represented across each organoid and treatment condition (Fig. 2d,f). However, population heterogeneity per organoid was greater in samples embedded in Matrigel compared to controls (Fig. 2f, Supplementary Figure S5).

We next employed differential expression testing to identify the top common genes driving separation of common populations across integrated cohorts (Fig. 2g, Supplementary Tables S1–12). Population 0 (Fig. 2g, pink; top genes *PCP4*, *HTR2C*, *SERPINF1*, *IGF1*) expressed markers consistent with fourth ventricle derivatives and the hindbrain choroid plexus including *TTR*, *ID1*, *SPARC* and also *BMP7*, which is known to be actively secreted from this region<sup>23</sup>. Population 1 (orange; top genes *HMP19*, *MAPT*, *VAMP2*, *RAB3A*, *NSG1* & *STMN2*) expressed *TCF7L2*, recently identified to mark RL-derived GCs and glutamatergic DCN<sup>24</sup>. This population also expressed *LHX9*, a marker of ventral RL-DCN, and *DCX*, which is expressed as DCN neurons undertake migration to the nuclear transitory zone<sup>25</sup>. Population 2 (beige; top genes *SPARCL1*, *ANXA2*, *TPBG*, *IGFBP5*, *WNT2B*, *PGM5-AS1.1*) expressed endothelial, astroglial and roof plate markers including *KRT18*, *IGFBP5*, *SPARCL1* and *PLTP*. Population 3 (lime; top genes *PTN*, *GPM6B*, *C1orf61*, *VIM*, *CAV1*, *GDF10*) exhibited expression of the characteristic Bergmann glia marker *FABP7*. We also noted expression within this population of *PTRPZ1*, recently shown to mark Bergmann glia<sup>15</sup>, and also *S100β*. Population 4 (green; top genes *PKMYT1*, *DYL*, *TK1*, *SPC23*, *RRM2*) also expressed *PCNA*, *MKI67* and *TOP2A* consistent with proliferative GCs. Population 5 (mint; top genes *NHLH1*, *CNTN2*, *DAPL1*, *ATOH1*, *TCF4*, *MIAT*) expressed GC precursor markers *NEUROD2*, *NEUROD6*, *PAX3*, *MEIS1*, *DCC* and *NFIB* consistent with RL derivatives. Population 6 (teal; top genes *TRIB3*, *HSPA9*, *DDIT3*, *NUPR1*, *EIF1* & *KCNG1*) expressed markers consistent with PNs including *PSAP*<sup>26</sup>, *CALCB*<sup>27</sup>, *ALDOA*, *FTL*, *FTH1*, and *AARS*. Population 7 (blue; top genes *RSPO3*, *RSPO1*, *RSPO2*, *CNTNAP2*, *MASP1*, *PLS3*) also strongly expressed the genes *SPARC*, *S100B*, *LMX1A*, *HES1* and *TTR*, consistent with roof plate cells. Population 8 (mauve; top genes *ASPM*, *TTK*, *DLGAP5*, *CENPA*, *TROAP*, *MKI67*) exhibited GC markers such as *PAX6*, *NFIB* and *TCF4* as well as *ASPM*, which sustains postnatal cerebellar neurogenesis<sup>28</sup>. Population 9 (purple; top genes *TCF7L2*, *TMSB4X*, *NEUROG2*, *DPYSL3*, *NNAT*, *OLIG3*) also expressed *NEUROD1* and *NHLH1*, consistent with cells of the RL. Population 10 (pink; top genes *LINC00152*, *MIR4435-2HG*, *HIST1H4H*, *HIST1H2AC*, *GADD45B*, *HIST1H2AE*) identity could not be conclusively determined, although expression of *PCP4*, *GRM8*, *TRPM3* and sub-threshold detectable levels of *PCP2* suggest this may be a second PN population. Population 11 (salmon; top genes *C11orf88*, *ARMC3*, *CAPSL*, *CFAP126*, *MAP3K19*, *ANKRD66*) also expressed cilia-related genes *DYNC112*, *DYNLRB2*, *IFT22*, *TMEM67*, *FUZ*<sup>29,30</sup>. Equivalent cell types speculated to represent the ciliated ependymal cells, which line the ventricular walls have been identified in other organoid models<sup>31,32</sup>. Expression of population-enriched markers is shown from in situ hybridisation data from the Allen Brain Atlas (ABA), where the majority of markers appeared restricted to specific cell types (Supplementary Figure S6). Moreover, a meta-comparison with reported markers from the literature provided further support for the identified cluster identities by revealing enrichment for canonical marker genes (Supplementary Figure S7, full gene lists for corresponding cell types in Supplementary Tables S1–S12). Together, our initial analysis identified that most cerebellar neuronal cell types are present in the organoids, including RL, GCPs, GCs, PNs, Bergmann glia, choroid plexus, ciliated cells, endothelial cells, roof plate and glutamatergic DCN.

Given the well-defined biology regarding cell-cycle progression in the developing cerebellum<sup>33</sup>, we anticipated a strong cell-cycle presentation in GC populations. Indeed, the strongest patterns of cell-cycle progression emerged in clusters ascribed in identity as GCs (Supplementary Figure S8a–d). Regression of cell-cycle did not result in appreciable differences in assignment of clusters (not shown). Consolidation of these groups as one cluster showed strong expression of canonical GC markers *PAX6*, *NEUROD1*, *TCF7L2*, *NFIB* and *NCAM1* (Supplementary Figure S8e), which were also observed in parental RL/GCP populations. It is likely these two distinct clusters represent distinct stages of GC maturation and migration, which have been shown to be instrumental in foliation of the cerebellum<sup>34</sup>.

**An unbiased classification strategy confirms assignment of cerebellar cell types.** Following putative identification of cluster identities, we next sought to validate our findings using an independent and unbiased approach. For this, we employed recently published datasets from the developing mouse and human cerebellum that were also generated using the 10× platform<sup>15,16</sup> and carried out data integration based on pairwise correspondences across datasets.

Integration of human cerebellar organoid data with a library of 70,000 cells across human prenatal cerebellar development from 9–22 postconceptional weeks (PCW) (Fig. 3) demonstrated that the majority of cerebellar neuronal cell types that are present during early human cerebellar development are found in the hiPSC-derived cerebellar organoids at day 90 and grossly localize to their tissue counterparts, overlapping in UMAP space (Fig. 3b,d,e). Human tissue clusters represent pooled data from multiple time points (UMAP metadata labels are shown in Fig. 3d). Hierarchical clustering of average expression profiles per population showed clustering of organoid-derived ciliated cells, roof plate, choroid plexus and endothelial cells with the tissue choroid/ependymal cluster. Bergmann glia clustered proximally to tissue clusters from astrocytes, glia and Bergmann glia (Fig. 3a,e). To examine whether gene expression patterns were obfuscated by pooling of timed data, we examined average expression values of cell type clusters according to developmental age (integrated object plotted according to maturity are shown in Fig. 3c), which revealed stronger and more effective clustering once average gene expression signatures were batch-corrected and devolved to their temporal components. Primarily, we found that organoid-derived PNs clustered strongly with 20 PCW PNs (Supplementary Figure S9). Organoid-derived ciliated cells, choroid plexus, endothelial and roof plate populations all clustered most strongly with the 14PCW



**Figure 3.** Human organoids comprise major cerebellar neuron types. (a) 69,174 human cells from ten developmental stages were integrated with 1653 human day 90 organoid-derived cells. Heatmap of Pearson correlation of average logged expression of highly-variable genes per cluster (cell type) from human cerebellar data<sup>16</sup> against human organoids (D90), with non-parametric/ComBat batch correction, clustering by cell type. Heatmap made with Pheatmap (v1.0.12). (b) UMAP plot of integrated, human organoid data (blue) with human developmental data (10-21PCW, grey) shows extensive and discrete localisation to the majority of identified human developmental clusters. (c) UMAP plot of integrated, human organoid data color-coded according to human developmental stage (9-21 PCW) shows complex cell-type and epoch-stage clustering. Human organoid data are represented as NA entry for visualisation purposes (grey). (d) Integrated object highlighting human organoid cell type identity reveals close proximity of human organoid and human developmental tissue cell types. Developmental human tissue data are coloured according to identity and human organoid data are represented as NA entry (grey) for visualisation purposes. (e) Integrated object highlighting human developmental tissue type identity reveals close proximity of human organoid and human developmental cell types. Human organoid data are represented as coloured while tissue data is shown as NA entry (grey) for visualisation purposes. The authors' original cell metadata labels have been preserved for consistency. Extended clustering of pseudo-bulk cluster with time course subdivisions are shown in Supplementary Figure S9.

choroid plexus tissue population. Organoid GC populations most strongly clustered alongside 21/14PCW GC progenitors. Bergmann glia most strongly resembled 18 PCW Bergmann glia. In agreement with this, principle component analysis with Brainspan data showed clustering of pseudo-bulk organoid samples with mid-to-late prenatal samples (19-38 PCW) (Supplementary Figure S10a). To further explore this, we deconvolved bulk organoid data to generate average expression profiles from each cluster and compared these to the top 500

region-based differentially expressed genes in the Brainspan Data (p-value cutoff of < 0.05). All neuronal populations showed strongest similarity matching to the cerebellum (CB), and in particular at 16PCW (Supplementary Figure S10b). Together, these analyses show that hiPSC-derived cerebellar organoids recapitulate early processes in the developing human cerebellum with the majority of cerebellar cell types present.

Recently published  $10 \times$  scRNA-seq data obtained from the developing mouse cerebellum (embryonic day E10 to postnatal day P10) represents the most current and comprehensive atlas of cerebellar cell types across murine cerebellar development, identifying 15 distinct cell types and comprising 48 clusters across ten developmental timepoints<sup>15</sup>. To identify corresponding cell populations in the cerebellar organoids, we utilized a strategy to identify pairwise correspondences across murine data and project equivalent cell types across species (Supplementary Figure S11). Following one-to-one human to mouse orthologue conversion and anchor integration, UMAP projection showed clustering of human organoid and murine cerebellar cell types (Supplementary Figure S11b, d,e), showing that particular organoid-derived cells co-occupied communal transcriptional space with murine cerebellar cells. This is further illustrated by hierarchical clustering of cell type showing the highest degree of clustering across species in organoid-derived glutamatergic DCNs, ciliated cells and roof plate clusters from both studies (Supplementary Figure S11a). Organoid choroid plexus and endothelial cells clustered with the murine meninges/pia membrane cluster. While Carter et al., did not identify specifically a group of Bergmann glia, the organoid Bergmann glia cluster alongside the murine astrocyte and glia clusters. Of note, both organoid RL and GC precursor groups clustered most closely to the mouse DCN group, reflecting their close developmental origin. UMAP projection showing the total integrated population with murine cell types coloured according to temporal maturity, reveals changing representation of cell types across time (Supplementary Figure S11c). Clustering with developmental age reveals a high degree of similarity between organoid-derived human ciliated cells and murine tissue-derived ciliated cell populations at E17 (Supplementary Figure S12). Human organoid-derived roof plate clustered most strongly with the E13 murine roof plate, which also clustered proximally to E13 glia and astrocyte groups. The human organoid choroid plexus and endothelial clusters also localized proximally to these groups, indicating shared transcriptomic similarity.

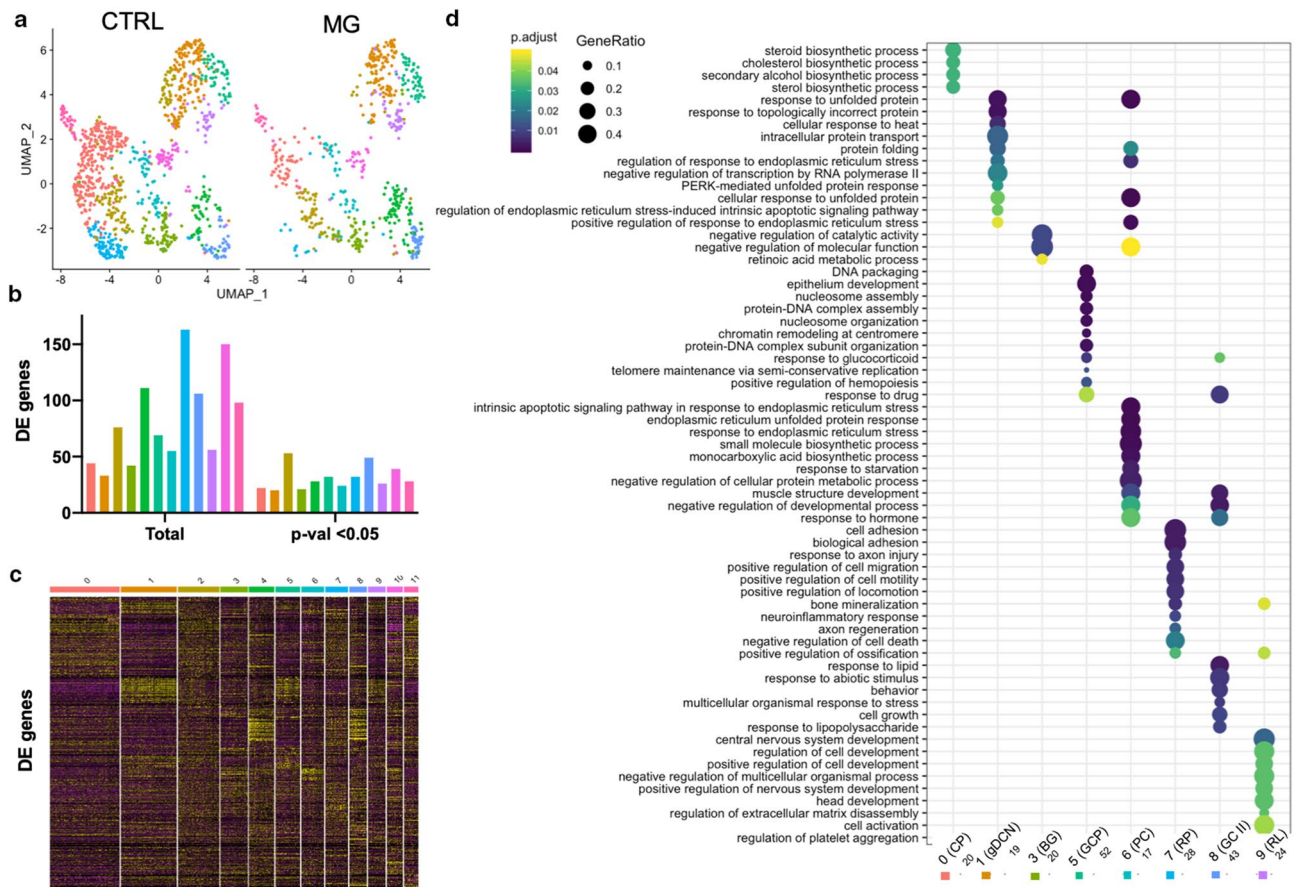
Thus, using the murine and human cerebellum as a close developmental blueprint, most transcriptional signatures indicate a mixture of mid-to-late embryonic temporal maturity, suggesting that the human cerebellar organoids recapitulate discrete developmental stages of the normally developing cerebellum and constitute the majority of cerebellar cell types present during this period.

### Differential expression and trajectory reconstruction analysis reveals gross effect of Matrigel encapsulation on organoid differentiation.

Following assignment of cell type clusters, we performed pseudo-time reconstruction to examine whether we could identify a developmental progression of maturing cell types. Trajectory reconstruction in pooled data showed a pattern reminiscent of the developmental cellular phylogeny of the cerebellum; cellular trajectory colour-coded by cell-type shows progression from primitive choroid plexus/roof plate cell types to RL/VZ precursors and subsequently to committed neuronal progeny bifurcating along the major glutamatergic/GABAergic lineages (Supplementary Figure S13a). Major drivers of branching included a number of recently identified markers involved in cerebellar specification<sup>22,35</sup>, including *WLS*, *MEIS1*, *MEIS2*, *TOP2A*, *MKI67*, *NEUROG1*, *ATOH1* and *HES3* (Supplementary Figure S13b). Pseudo-time reconstruction was largely consistent following lineage reconstruction in embedded samples (Supplementary Figure S14). Interactive pseudo-time reconstructions are available at <https://plotly.com/~SamN1985/3> and <https://plotly.com/~SamN1985/1> (control and Matrigel, respectively).

To explore embedding-related differences, we performed differential expression testing between control and Matrigel-embedded cohorts. 46 genes were significantly differentially-regulated following Matrigel encapsulation (36 genes FDR < 0.05; 6 up, 30 down) relative to control organoids (Supplementary Figure S13c,d, Supplementary Table S16). Gene-set overrepresentation analysis returned significantly altered Gene Ontology (GO) pathways including *carboxylic acid biosynthetic process* (FDR 5.16E-05), *organic acid biosynthetic process* (FDR 5.24E-05), *regulation of angiogenesis* (FDR 6.21E-05), *regulation of vasculature development* (FDR 1.1E-04), and *tissue regeneration* (FDR 1.2E-04). Consistent with the transformation of growth and peripheral neuronal outgrowth following encapsulation, a number of genes concordant with survival, differentiation, and neurogenesis were among those significantly upregulated including *BIRC5*/Survivin, *HES6*, which promotes neurogenesis through the Notch pathway<sup>36</sup>, *DCN*, which has been shown to regulate endothelial cell–matrix interactions during angiogenesis<sup>37</sup>, and *PKM*, which inhibits proliferation during postnatal cerebellar neurogenesis<sup>38</sup> (Fig. 4d). Furthermore, a number of hormones, chemokines and neuropeptides were amongst the genes differentially regulated with Matrigel treatment. These include *TRH*, a hormone acting on receptors normally expressed in GCs and molecular layer interneurons<sup>39</sup>, *IGF1*, which is known to have pleiotropic effects in all neural cells including neurons, oligodendrocytes and astrocytes<sup>40</sup>, *CXCL14*, a chemokine secreted by PNs directing GC migration<sup>41</sup>, and *CRABP2*, which regulates RA production<sup>42</sup>. In addition, several cell cycle-related genes were amongst the list of those differentially expressed (*CENPF*, *CCNB1*, *CCNB2*, *BIRC5*, *DCN*), suggesting that Matrigel encapsulation transforms organoids through modulation of cell cycle-dependent pathways. Furthermore, cell type-specific markers *TTR*, *IDI*, *DCN*, *PCP4*, *CXCL14*, *RSP02* and *HES6* were differentially regulated following encapsulation. *IDI* (downregulated in Matrigel-embedded organoids, p-value 0.04755947) is transcriptionally active in embryonic ventricular and subependymal zones as well as being involved with postnatal development<sup>43</sup>. *ID* family members are known to regulate differentiation and their presence supports the detection of heterogeneity across treatment conditions. *TTR* and *PCP4* distinguish VZ precursors, while *DCN* is normally restricted to RL derivatives.

Following several independent lines of cluster identity confirmation and marker enrichment, we observed a consistent bias in population composition in the Matrigel-encapsulated samples, towards cell type identities



**Figure 4.** Cerebellar populations show distinct responses to Matrigel encapsulation. **(a)** UMAP projections outlining equivalent populations from control (CTRL) and Matrigel (MG)-embedded cell groups for differential-expression testing. **(b)** Graph depicting numbers of total (left) and significantly changing genes (right) per cell population. P-value < 0.05 accepted as significant. **(c)** Heatmap depicting all differentially-regulated genes between untreated and MG-embedded organoids. Heatmap made with Seurat (v3.2.1) **(d)** Dot plots showing the top differentially expressed GO pathways per cell population. The number of genes considered for each population is listed on the x-axis below cell labels.

arising from the RL (Fig. 2f & Supplementary Figure S5). To more clearly illustrate the relative contribution of RL/VZ derivatives, we performed retrospective re-clustering, distinguishing clusters comprising RL/EGL/VZ populations (Supplementary Figure S15a,b,c). Control organoids comprised  $68.99 \pm 0.29\%$  VZ derivatives compared with Matrigel-embedded organoids, which showed  $46.97 \pm 10.43\%$  VZ derivatives (Supplementary Figure S15d), indicating a considerable bias away from the VZ following embedding. Chi-square testing revealed consistent VZ/RL bias following treatment at both clustering resolutions (Supplementary Figures S16 and S17). Validation of RL/EGL/VZ clustering is shown by enrichment of genes driving separation at this resolution (Supplementary Figure S15b), top markers (Supplementary Figure S15c) and genes distinguishing populations (Supplementary Tables S13–15). This includes *TRPM3*, *PCP4*, *SOX2*, *TTR* and *ID1* (VZ), *NCAM1*, *TCF7L2*, *NHLH1*, *NEUROD6* (RL) and *TOP2A*, *MKI67*, *PCNA*, *FOXM1* (EGL). The EGL cluster also expressed *DLGAP5*, *CDC48*, *KIF2C*, *CDC20*, and *AURKA*. In summary, significant lineage-bias towards RL derivatives was observed upon Matrigel embedding, which was confirmed by iterated clustering at multiple resolutions.

### Population-level differential expression testing reveals cell-type specific responses to Matrigel.

To evaluate the effect that Matrigel encapsulation had on specific cerebellar populations (Fig. 4a), we performed gene over-representation tests on differentially expressed genes (Fig. 4b,c, FDR < 0.05) across treatment conditions in equivalent populations. A list of differentially regulated genes is provided for each population comparison in Supplementary Tables S17–S28. GO-enriched pathways (Fig. 4d) were examined for each population following treatment. Population 0 (choroid plexus) following Matrigel encapsulation exhibited significantly altered regulation of ‘Steroid biosynthetic process’ (FDR  $7.44E-05$ ) and ‘Cholesterol biosynthetic process’ (FDR  $7.44E-05$ ). Notably, cholesterol release from epithelial cells of the choroid plexus epithelia has been described to play an important role in cholesterol homeostasis in the cerebrospinal fluid<sup>44</sup>. Population 3 (Bergmann glia) showed significant changes in ‘Retinoic acid metabolic process’ (FDR  $1.68E-04$ ). Release of retinoic acid by radial glial cells has shown to be instrumental for numerous developmental processes including formation of the blood–brain barrier<sup>45</sup>, while cholesterol has been demonstrated to be vital for maturation

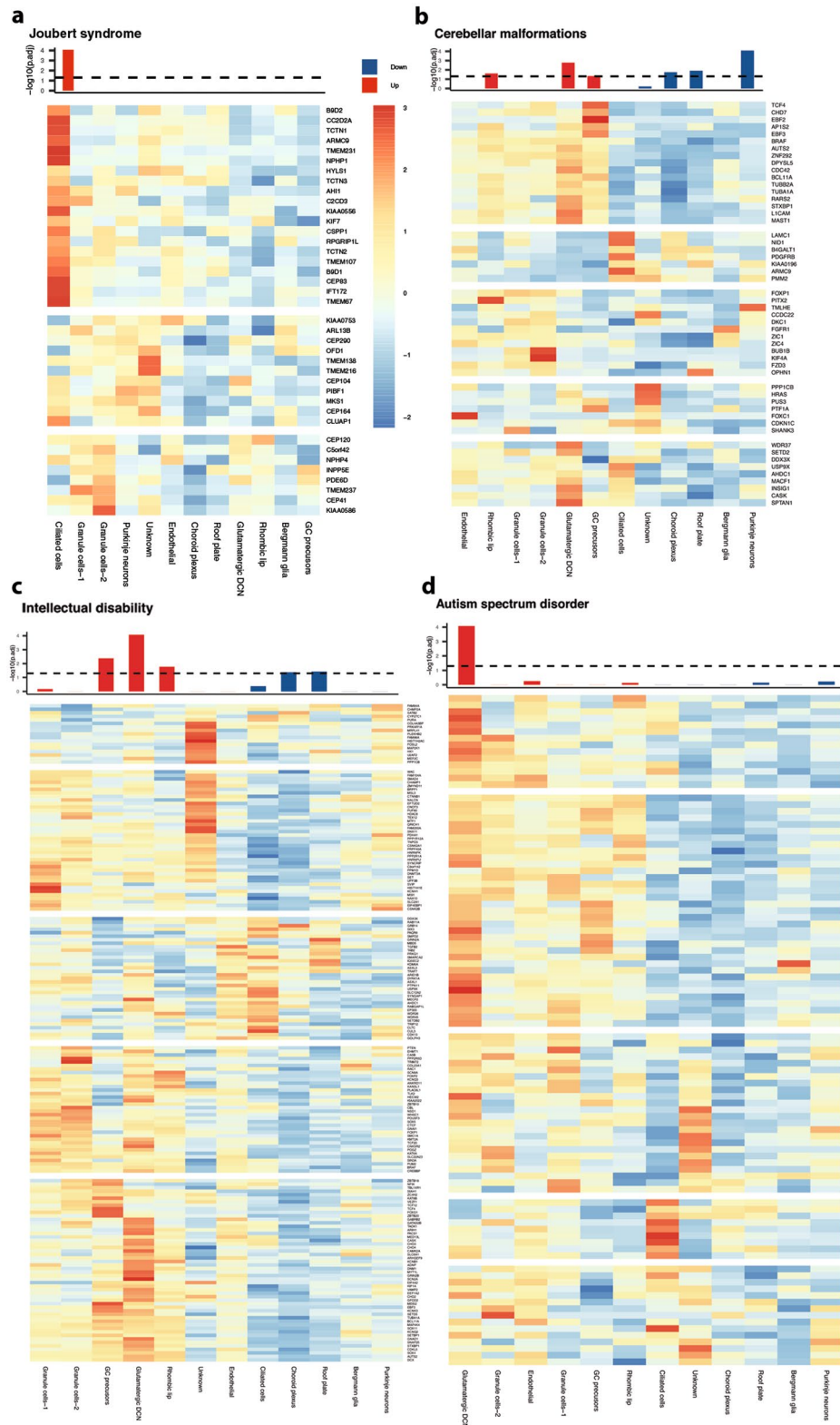


of brain-associated microglia<sup>46</sup>. Population 4 (GCs-1) showed differential regulation of ‘DNA packaging’ (FDR 9.27E-08), ‘nucleosome assembly’ (FDR 2.00E-06), ‘chromatin remodelling at centromere’ (FDR 1.25E-05) and ‘telomere maintenance via semi-conservative replication’ (FDR 1.78E-04). Population 6 (PNs) displayed enrichment for pathways ‘Intrinsic apoptotic signalling pathway in response to endoplasmic reticulum stress’ (FDR 5.68E-09), ‘response to starvation’ (FDR 5.03E-05) and ‘monocarboxylic acid biosynthetic process’ (FDR 1.71E-05). PNs express high levels of monocarboxylate transporter 2, an integral component of the lactate transport system required for energy metabolism and neuronal activity<sup>47</sup>. Overall, the relative effect of Matrigel encapsulation resulted in distinct responses in the various cerebellar populations, which could be related to their function in vivo. This was most notable in RL derivatives, where encapsulation seemed to strongly influence cell cycle dynamics, driving expansion of the GC population.

**Subclustering of Purkinje neuron clusters.** The majority of organoid-derived cerebellar cell types showed enrichment of canonical marker genes (Fig. 2g, Supplementary Figure S7), as well as resolution with their tissue counterparts when projected in mutual transcriptional space (Fig. 3, Supplementary Figure S9). PNs only clustered following deconvolution to their temporal components (PCA loadings are shown in Supplementary Figure S18, S19); however, we noticed a surprising lack of canonically-ascribed PN markers including *ALDOC*, *SKOR2*, *RORA* and *PCP2* localizing to the cluster we assigned PN identity. This population clustered diffusely within the human tissue PN cluster following integration (Supplementary Figure S20a shows UMAP projection with only human tissue PNs and organoid PNs projected), which was unsurprising given this cluster spanned ten developmental timepoints, presumably representing a rich biological complexity. The lack of obvious enrichment of biomarkers in our PN population may have been for a number of reasons including low cell numbers relative to tissue datasets (107 cells versus 25,724), high levels of ribosomal and stress-response genes, or intrinsic culture versus native tissue differences. One other possibility is that our clustering failed to resolve PNs at these parameters at a resolution that serviced delineation of the other cerebellar cell types because of differences in transcriptional complexity. To examine the contribution that our clustering parameters had on enrichment of PN markers, we subclustered these populations and examined overlap of enrichment of biomarkers in human and murine tissue data. We used Venice, a recent algorithm designed to overcome some of the challenges with mixed cell states/rare populations being represented by mean gene expression values<sup>48</sup>. BioTuring BBrowser<sup>49</sup> was used to identify markers in tissue-derived PNs (pooled time points) and organoid-derived PNs, identifying 10.2% shared marker expression (977 genes logFC > 0.001, Supplementary Table S33, Supplementary Figure S20b). The ‘Unknown’ cluster, which previously clustered alongside human tissue PNs (Supplementary Figure S9), also showed diffuse clustering and overlap with more than one human tissue PN region (Supplementary Figure S20c). This cluster shared a high degree of marker overlap with PN population (63.6%), with 8.1% of biomarkers shared between the two organoid-derived populations and the tissue PN group (852 genes, Supplementary Figure S20d and Supplementary Table S34). Lastly, we compared PN markers from human tissue, murine tissue and both the PN and Unknown clusters, identifying 236 genes in common (2.2%, Supplementary Figure S20e and Supplementary Table S35). 21 of these were represented in the Ma’ayan PN gene set<sup>50</sup>, including *DCX* (previously observed in PNs<sup>51,52</sup>), the Reelin receptor *VLDLR*, known to be involved with patterning of PN topography<sup>53</sup>, *DLGAP4* (known to be expressed by PNs<sup>54,55</sup>), *CRMP1* (required for PN migration<sup>56</sup>) as well as *RAB3C* and *TNFRSF21*. *NOVA2*, which regulates neuronal migration in PNs<sup>57</sup> was also detected. *SOX4*, enriched here, was recently demonstrated to be expressed in PNs<sup>16</sup>. *CALB1* and *PCP4* were detected as biomarkers in all four groups adding further evidence for this cluster as PNs.

One possibility regarding underrepresentation of canonical biomarkers is our clustering resolution, which delineated all other cell types. Our PN and Unknown clusters could thus in fact represent two PN subtypes, separation of which would occlude conventional biomarker identification. To explore this, we grouped these cells together, also separating cells on the periphery of the BG/PN cluster, noting 180 genes including the appearance of several markers previously observed in PNs, including *NRG1*<sup>58</sup>, *FOXPI*<sup>59</sup> and *MEF2C*<sup>60</sup>, and 12 of which were in common with the Ma’ayan PN gene-set list (Supplementary Figure S20f., Supplementary Table S36).

**Relevance to cerebellar disease.** The cerebellum is linked to a growing number of neurodevelopmental disorders<sup>61</sup>. However, the underlying genetic and molecular mechanisms remain incompletely understood. To investigate whether cerebellar organoids might serve as model systems to identify vulnerable cell populations and better understand underlying disease mechanisms in these developmental conditions, we examined transcriptomic data from individual cell clusters to identify enrichment for disease genes. We first examined genes causing Joubert syndrome<sup>16</sup>, a recessive neurodevelopmental ciliopathy, and identified significant positive enrichment in the ciliated cell population (Fig. 5a). We next examined the enrichment of genes that are associated with structural cerebellar malformations (cerebellar hypoplasia (CH) and Dandy-Walker syndrome (DWS))<sup>62,63</sup>. Genes associated with these disorders showed significant positive enrichment in RL, GC precursors and glutamatergic DCN, as well as negative enrichment in choroid plexus, roof plate, and PN populations (Fig. 5b), pointing towards potentially vulnerable cell populations in these disorders. Genes associated with intellectual disability (ID)<sup>64–66</sup> showed prominent positive representation in RL, GC precursors, and glutamatergic DCN. Negative enrichment was associated with choroid plexus and roof plate clusters (Fig. 5c). Genes associated with autism spectrum disorder (ASD)<sup>64,67–71</sup> were enriched in glutamatergic DCN (Fig. 5d). We did not identify any significant enrichment for genes associated with the autosomal recessive and dominant ataxias<sup>72,73</sup> that represent neurodegenerative disorders of the cerebellum (Supplementary Figure S21). This is consistent with the identified developmental stage of the human cerebellar organoids at day 90, at which most ataxia genes are not yet expressed. Together, these findings underscore the utility of cerebellar organoids for studies of human neurodevelopmental disorders.



**Figure 5.** Enrichment of disease genes in discrete organoids cell populations. Heatmaps showing the average scaled expression across each of human organoid-derived cell type of risk genes for (a) Joubert Syndrome, (b) cerebellar malformation, (c) intellectual disability and (d) autism spectrum disorder. Enrichment p-values (-Log10 P value) for each cell type based on GSEA (each cell type vs all others) are shown in the top bar plots. The dashed line is the Bonferroni significance threshold ( $P < 0.05$ ). Bonferroni correction was calculated by dividing nominal p-value by total number of cell types and gene sets ( $12 \times 6$ ). The colour of each bar in the barplot shows upregulation (red) or downregulation of gene set in particular cell type vs all other cell types based on positive or negative NES values from GSEA results (Supplementary Table S37). Heatmaps were generated using Pheatmap (v1.0.12).

## Discussion

In this study, we have employed a robust and reproducible protocol for the generation and culture of three-dimensional cerebellar organoids from hiPSCs that does not rely on co-culture with mouse progenitors. By performing scRNA-seq of individually barcoded organoids, we demonstrate that this protocol reliably generates transcriptionally discrete populations encompassing major cerebellar neuronal cell types including RL, GC progenitors, GCs, PNs, roof plate, choroid plexus, endothelial cells, ciliated cells, Bergmann glia, and glutamatergic DCN. The identified cell populations largely correspond to cells present in the early developing human and mouse cerebellum. Moreover, we show that growth and lineage commitment of the organoids can be influenced by treatment with Matrigel. Finally, we identify cell populations in the hiPSC-derived cerebellar organoids that are enriched for disease genes associated with cerebellar developmental disorders.

Cell autonomous processes, as well as the local microenvironment drive specification and development of the cerebellum. A popular method to promote sustained growth of cortical organoids has been the addition of Matrigel to the culture setting, promoting differentiation through exposure to molecules present at the basement membrane<sup>74</sup>. The ECM-enriched basement membrane of the cerebellum is located underneath the pial meningeal layer, where it supplies anchor points for the endfeet of radial processes of VZ precursors, as well as acting as a barrier to migrating neurons<sup>75,76</sup>. These processes provide a scaffold, facilitating both cortical and cerebellar neuronal migration and correct layer formation. Evidence suggests that the basement membrane influences early patterning events at the MHB. In agreement with these observations, our data show that migratory and proliferative activity is significantly altered in Matrigel-embedded cerebellar organoids. However, Matrigel encapsulation also decreased the reproducibility of generating cerebellar cell types, with embedded organoids showing a greater degree of variability in cell type composition. Interestingly, this heterogeneity comprised a bias towards RL derivatives, including glutamatergic DCN and GCs. This is intriguing given that previous cerebellar differentiation protocols have been reliant on co-culture with isolated murine RL progenitors or GCs to provide necessary factors for growth and maturation of the VZ component<sup>9,10,12</sup>. Our population-level differential expression analysis suggests that GCs were especially sensitive to the addition of Matrigel, showing a greater degree of responsiveness as marked by increased numbers of significantly altered genes, particularly of those involved in cell cycle-related pathways. An alternative hypothesis involves embedding driving differential oxygen gradients and hypoxia within the organoids, a mechanism which has previously been shown to drive proliferation of GCPs in the cerebellum<sup>77,78</sup>. In this regard, hypoxic pockets during the differentiation may actually be advantageous.

While we showed that differences in Matrigel encapsulated organoids were subtle, encapsulation of cerebellar-patterned organoids may provide a viable alternative to costly, time-restrictive rodent RL isolations that hamper downstream analyses of differentiated human cells. However, this method should be carefully considered given batch-to-batch variability associated with Matrigel and the extra degree of handling this introduces to the culture workflow.

An exciting potential for organoid-based protocols is the fact that neighbouring regions can influence the development of adjacently located tissues. For example, both roof plate epithelium and choroid plexus epithelium are required for full development of the RL in vivo<sup>79,80</sup>. It has been suggested that diffusible factors such as retinoic acid influence development of the cerebellum, following secretion from the choroid plexus<sup>81</sup>. We show that the organoids differentiated in the absence of murine feeder layers contain choroid plexus, roof plate, ciliated cells, RL precursors, glutamatergic DCN, GCs, Bergmann glia and PNs. This suggests that organoids transit through stages akin to physiological embryonic cerebellar development, which is supported by the pseudo-temporal ordering. Future studies utilizing spatial sequencing or high-resolution imaging will give insight whether these cell types function in a manner akin to their in vivo counterparts. The presence of choroid plexus cells and endothelial cells within the organoids is particularly interesting as it raises the possibility that cerebellar organoids may possess an intrinsic angiogenic capability as well as cells of non-ectodermal origin.

In this study we did not employ SDF1 and FGF19 treatment, which encourages cerebellar plate neuroepithelium formation but is not necessary to derive mature neurons<sup>10,14</sup>. Future single-cell studies will be well-suited to provide molecular insights into how these growth factors promote more efficient recapitulation of the primitive cerebellar germinal zones.

Our integration with existing single-cell data from the developing mouse and human cerebellum suggests that overall, human organoids at day 90 represent mid-to-late embryonic cerebellar development. This is also reflected by the absence of some well-known, more mature markers for distinct cell populations. For example, organoid-derived PNs, which clustered most strongly with 20 PCW human tissue counterparts, failed to robustly express mature markers including *ALDOC*, *RORA*, and *PCP2*. This is consistent with the notion that PN maturation occurs during the late-gestational stage of development and is reliant on the establishment of the internal granule layer (IGL)<sup>16</sup>. PN maturation occurs in late gestation and is reliant on GCPs completing migration to the IGL from the EGL. In further support of this, we did not detect appreciable levels of *SHH* expression, and Calbindin-expressing neuronal cells present in D90 cerebellar organoids did not resemble the characteristic teardrop shaped soma with a single bifurcating axon, characteristic of mature PNs. Findings in the literature suggest that PNs and GABAergic DCN arise from the VZ prior to dependence on SHH signalling<sup>82</sup>, consistent with the presence of immature PNs in the organoids that are still undergoing maturation, potentially in suboptimal culture settings. Moreover, organoid PNs exhibited increased expression of stress-response genes, which has previously shown to occlude identification and specification<sup>83</sup>. Another limitation of this study has been the limited number of identified PNs, which limited our ability to effectively survey this cluster. We have looked in greater detail at expression of genes from a number of datasets which also transcriptionally identified PNs, expanding our list of potential markers. Higher resolution studies are required to definitively place these cells on a maturation axis. Further comparison of the organoid-derived PNs to murine and human PN transcriptomes may yield some clues about supportive trophic settings that could be exploited experimentally in the future to promote survival

and maturation. While we did not observe strong *SHH* expression, it is worth noting that transventricular SHH delivery, from outside of the cerebellum has been demonstrated to drive early VZ development<sup>84</sup>.

One limitation in comparing cell types across species and studies is incongruence of cell type labels, which are ultimately ascribed subjectively by the end-user. This, as well as technical variability may underlie the fact that not all cell types consistently clustered with their tissue counterparts. However, some cell types did not resolve with their expected counterparts, which could be down to culture-induced differences. For example, the organoid RL cluster clustered most closely with brain stem and excitatory CN. While excitatory CN arise from the RL, it is reasonable that a residual gene expression signature explains this clustering. We did not detect brainstem marker *HOXB3* and can only assume some active transcriptional programs in the brain stem resemble those in the cultured GCP/RL cells. Deconvolution of the temporal component enabled and improved cluster identification, depending on the cell type, presumably due to the transcriptional complexity/stability over time of the population. Recent advances<sup>85,86</sup> to adopt standardized repositories, QC metrics and nomenclature, including and transcriptome-based taxonomic classification systems for neurons will be necessary in future work.

Our study shows that cerebellar organoid technology enables the generation of human cell types that are normally inaccessible, and provide insights into windows of development that are normally restricted. Enrichment of disease genes in distinct cell populations of the human cerebellar organoids highlights the exciting potential of this model system to provide insights into the pathogenesis of cerebellar disorders and as a valuable platform for drug screening and personalised medicine.

## Methods

**Ethics statement.** The human iPSC cell line used in this study was derived from a healthy 67-year-old female donor (AHO17-3) following signed informed consent and with approval from the National Health Service, Health Research Authority, NRES Committee South Central—Berkshire, UK as previously described<sup>95</sup>. All experiments were performed in accordance with relevant guidelines and regulations.

**hiPSC growth and maintenance.** hiPSCs from a healthy 67-year-old female donor (AHO17-3) were used as previously described<sup>11</sup> between passage 25–35. Samples were tested and free of Mycoplasma using the MycoAlert kit (Lonza, Switzerland). hiPSCs were maintained on Matrigel-coated plates in a humidified CO<sub>2</sub> incubator with mTeSR1 media changed daily (STEMCELL Technologies, UK). hiPSC were passaged upon reaching approximately 80% confluence using EDTA at 37 °C for three to five minutes and plated in mTeSR1 with 10 μM Y-27562 (Tocris, UK). Media without Y-27562 was supplemented the day following.

**Cerebellar organoid differentiation.** Figure 1a provides an overview of the differentiation protocol. Briefly, embryoid bodies (EBs) were made by dissociating sub-confluent hiPSCs with TrypLE before allowing to aggregate in ultra-low attachment 96 well-plates (PrimeSurface MS-9096 V, Sumitomo Bakelite, Japan) in growth factor-free chemically-defined medium in the presence of 7 μg/ml Insulin (Sigma-Aldrich), 50 μM Y-27562 (Sigma-Aldrich) and 10 μM SB431542 (Tocris)<sup>9</sup>. Approximately 1.1E04 cells were deposited into each well. Two days later, media was supplemented with 50 ng/ml FGF2 (R&D Systems). On day seven, one-third of the media was removed and supplemented with fresh media without growth factors. On day 14 full volume media replacement was performed. On day 21 of differentiation, individual EBs were embedded in undiluted Matrigel or transferred into differentiation media as free-floating aggregates (four EBs per well of a 24-well low-attachment plate). EBs were harvested on day 21 for cryosectioning/qPCR to validate acquisition of MHB fate. Medium was changed on day 28 and 35 before assessing expression of rhombic lip/ventricular zone markers. EBs were monitored for growth by photographing every two days and inspected for visual signs of differentiation including the presence of polarized neuroepithelia (rosettes). On day 35, individual EBs were transferred to transwell membranes (Millicell Cell Culture inserts, PTFE, 0.4 μm, Merck Millipore, USA) and allowed to further develop at the air–liquid interface. Organoids were harvested at multiple timepoints throughout for validation of successful cell-fate acquisition.

**Immunostaining.** EBs and organoids were harvested for cryosectioning at various timepoints by fixing in 4% paraformaldehyde for 15 min at room temperature, followed by three washes of PBS. Samples were cryoprotected overnight in 20% sucrose/PBS before embedding in optimal cutting temperature compound and sectioning on a cryotome at 10 μm. Cryosections were blocked in 2% milk power/PBS for one hour at RT following permeabilization with 0.3% Triton-X100 for 10 min at RT. Sections were incubated with primary antibody diluted in blocking buffer with 0.3% Triton-X100 overnight at 4 °C followed by application of relevant fluorescent secondary-conjugated antibody (Alexa Fluor, Invitrogen). An identical preparation without the primary antibody was used as a negative control. Hoechst was applied for 5 min at RT to visualize nuclei. Immunostained sections were mounted in ProLong Gold Antifade (VECTASHIELD). Antibodies are listed in Supplementary Table S29.

**Quantitative real-time PCR.** RNA was isolated from EBs/organoids on days 0, 21 and 35 using the RNeasy mini kit (Qiagen, Germany) according to manufacturer's instructions including an on-column DNase digestion step. cDNA was generated using the Superscript III First-strand synthesis kit (Invitrogen, USA). Quantitative real-time PCR was performed on an Applied Biosystems StepOne plus qPCR machine using Fast SYBR Green Mix according to the manufacturer's specifications. No template and minus reverse transcriptase controls were used to test for specificity. Primers are listed in Supplementary Table S30. Relative fold-change of mRNA abundance was calculated using the  $\Delta\Delta Cq$  method. Samples were calibrated relative to *ACTB*. Results are shown from between 5 and 6 independent differentiation experiments.  $\Delta Cq$  values were tested for significance with  $P < 0.05$  accepted as significant, Mann–Whitney test.

**Imaging and quantification.** Imaging was performed on mounted cryosections/whole-organoids using a Zeiss upright fluorescent microscope. CellProfiler was used to segment and identify immunopositive nuclei in fluorescently-labelled cryosection images. This was divided by the total number of nuclei as ascertained by visualisation with Hoechst. Organoid growth was assessed by measuring diameter using ImageJ and assessed for significance through two-way ANOVA. Formation of polarized neuroepithelial tissue was quantified in blinded samples and subject to statistical testing using one-way ANOVA.

**Statistical analysis.** Statistical tests were performed with the PRISM software package (v8) or R. qPCR data is shown from multiple independent differentiation experiments,  $\Delta Cq$  values tested for significance using Mann–Whitney test (two-tailed). Figure 1b; Five-six biological replicates (differentiation experiments) and three technical replicates (PCR replicates). Figure 1d; Eight-nine biological replicates (differentiation experiments), three technical replicates (PCR replicates). Figure 1e; Growth rate of organoids was assessed by two-way ANOVA (Dunnet's multiple comparison test); Eight biological replicates (differentiation experiments), three-five technical replicates (individual organoids).  $F$  (DFn, DFD); (7, 54) = 2.466 (Interaction), (7, 54) = 17.69 (Row factor), (1, 54) = 7.110. Figure 1f; Three biological replicates (differentiation experiments) and three technical replicates (PCR replicates). Chi-square testing was used to assess lineage-specific bias in SC data at both utilized clustering resolutions.  $P$ -value < 0.05 accepted as significant for all statistical testing. Weighted Kolmogorov–Smirnov statistics and gene-set based permutations were used to derive nominal  $P$ -values for gene-set enrichment analysis for disease risk analysis.

**Preparation of organoids for single cell sequencing.** Following 90 days of differentiation, three organoids embedded in Matrigel (at day 21) and three unembedded organoids were manually dissected from transwell membranes and enzymatically digested with Accumax dissociation reagent (Sigma-Aldrich) with manual trituration. Organoids were resuspended in 0.01% BSA in PBS on ice and labelled with Hash-tagged Oligonucleotide (HTO)-conjugated antibodies for B2M/ATP1B3 at RT for one hour (HTO sequences and antibodies are listed in Supplementary Table S31). Samples were processed according to the standard 10× Chromium 3' workflow and pooled before sequencing over two lanes (Illumina HiSeq 4000).

**Processing of single cell data.** A total of 2,526 cells from two hashed pools were sequenced at an average of 53,751 reads per cell, with a median detection rate of 2,617 genes per cell. Hashed pools were demultiplexed using Seurat<sup>87–89</sup> HTODemux with default options and 65% of the cells (1653) were classified as singlets. The SCTransform function in Seurat (v 3.2.1) was used to normalize UMI depth across cells using negative binomial regression followed by additional regression for mitochondrial proportion across all cells. Each hash pool was analysed separately and then also integrated and projected in UMAP space using canonical correlation analysis (CCA) and MNN (Mutual nearest neighbours). The ClusTree package and manual curation was used to visualize and determine optimal clustering parameters. Populations were identified through unsupervised clustering and projection of known cell-type specific markers using the FindConservedMarkers using default parameters<sup>90–92</sup>. Organoid composition was projected by plotting the fraction of each population per organoid after integration/CCA + MNN projection. Expression of cluster-specific markers was validated through comparison to sagittal ISH sections of the developing murine brain (Allen Brain Atlas). Differential expression testing was performed using the FindConservedMarkers command using a Wilcoxon Rank Sum Test using identity classes comprising the two hash pools (Matrigel-embedded samples vs control). Heatmaps were generated using Seurat (v3.2.1) and pheatmap (v1.0.12). Mouse vs human comparisons were made by performing Integration and Label Transfer<sup>93</sup> and incorporating high-confidence/one-to-one human to mouse orthologues from Ensembl. Pseudo-time reconstruction was performed using Monocle. A list of identities pertaining to cluster identity from murine dataset is listed in Supplementary Table S32.

**Comparison with BrainSpan RNAseq data.** Normalised gene-level expression data in RPKM for the BrainSpan RNA-seq dataset generated from post-mortem human brain were downloaded (<https://www.brainspan.org/static/download.html>) and only protein coding genes were retained. The human organoid SCT data was pseudobulked (Average expression) according to control and Matrigel groups and only protein-coding genes retained. A ComBat batch correction for study was performed on the log2 transformed joint BrainSpan and pseudobulked organoid protein-coding gene \* sample matrix. The organoid samples were then projected onto the BrainSpan maturity PCA axis (formed by using the ComBat corrected BrainSpan data).

**Pathway analysis.** GO Term Biological process over-representation tests across all cell populations was performed using the *CompareCluster* function from the Bioconductor package ClusterProfiler<sup>94</sup>. For each cell type population, differentially-expressed genes between treatments with a  $p$ -value < 0.05 were included in the analysis and the background universe consisted of 17,125 genes that were expressed in a minimum of 3 cells with a UMI count > = 1.

**Cerebellar disease enrichment.** We used Gene Set Enrichment Analysis (GSEA) to test enrichments of each of the six disease risk gene sets in one cell type vs all others. We calculated the average expression of each of the integrated organoid cell types and then performed GSEA with GSEA GUI app for each cell type vs all for all the 6 gene sets with the following parameters : enrichment statistic : weighted, gene ranking metric : Diff\_of\_Classes, Gene list sorting mode : real, Gene list ordering mode: descending, number of permutations: 10,000, permutation\_type: gene labels, Max GeneSet size : 500, Minimum GeneSet size : 15, Normalisation\_mode: mean

and div. Nominal p-values from GSEA were corrected for multiple testing using the Bonferroni method calculated over all cell types and gene lists used [ $P < 0.05/(12 \times 6)$ ]. A heatmap of the average z-scores of the diseases risk gene sets across each of the cell types was plotted and the gene and cell types were clustered with the *ward.D2* method using the pheatmap R library.

### Data availability

Data is available on GEO (GSE150153).

### Code availability

Code is available on Github repository <https://github.com/SamN1985/General-code> and [https://github.com/deevdevil88/10x\\_Human\\_Cerebellum\\_Organoid](https://github.com/deevdevil88/10x_Human_Cerebellum_Organoid).

Received: 9 November 2020; Accepted: 26 May 2021

Published online: 21 June 2021

### References

1. Reeber, S. L., Otis, T. S. & Sillitoe, R. V. New roles for the cerebellum in health and disease. *Front. Syst. Neurosci.* **7**, 83. <https://doi.org/10.3389/fnsys.2013.00083> (2013).
2. Schmammann, J. D., Guell, X., Stoodley, C. J. & Halko, M. A. The theory and neuroscience of cerebellar cognition. *Annu. Rev. Neurosci.* **42**, 337–364. <https://doi.org/10.1146/annurev-neuro-070918-050258> (2019).
3. Sathyanesan, A. *et al.* Emerging connections between cerebellar development, behaviour and complex brain disorders. *Nat. Rev. Neurosci.* **20**, 298–313. <https://doi.org/10.1038/s41583-019-0152-2> (2019).
4. Stoodley, C. J. & Limperopoulos, C. Structure-function relationships in the developing cerebellum: Evidence from early-life cerebellar injury and neurodevelopmental disorders. *Semin. Fetal. Neonatal. Med.* **21**, 356–364. <https://doi.org/10.1016/j.siny.2016.04.010> (2016).
5. Herculano-Houzel, S. & Lent, R. Isotropic fractionator: A simple, rapid method for the quantification of total cell and neuron numbers in the brain. *J. Neurosci.* **25**, 2518–2521. <https://doi.org/10.1523/JNEUROSCI.4526-04.2005> (2005).
6. Butts, T., Green, M. J. & Wingate, R. J. Development of the cerebellum: Simple steps to make a “little brain”. *Development* **141**, 4031–4041. <https://doi.org/10.1242/dev.106559> (2014).
7. van Essen, M. J., Nayler, S., Becker, E. B. E. & Jacob, J. Deconstructing cerebellar development cell by cell. *PLoS Genet.* **16**, e1008630. <https://doi.org/10.1371/journal.pgen.1008630> (2020).
8. Wechsler-Reya, R. J. & Scott, M. P. Control of neuronal precursor proliferation in the cerebellum by Sonic Hedgehog. *Neuron* **22**, 103–114. [https://doi.org/10.1016/s0896-6273\(00\)80682-0](https://doi.org/10.1016/s0896-6273(00)80682-0) (1999).
9. Muguruma, K., Nishiyama, A., Kawakami, H., Hashimoto, K. & Sasai, Y. Self-organization of polarized cerebellar tissue in 3D culture of human pluripotent stem cells. *Cell Rep.* **10**, 537–550. <https://doi.org/10.1016/j.celrep.2014.12.051> (2015).
10. Ishida, Y. *et al.* Vulnerability of Purkinje Cells Generated from Spinocerebellar Ataxia Type 6 Patient-Derived iPSCs. *Cell Rep.* **17**, 1482–1490. <https://doi.org/10.1016/j.celrep.2016.10.026> (2016).
11. Watson, L. M., Wong, M. M. K., Vowles, J., Cowley, S. A. & Becker, E. B. E. A simplified method for generating purkinje cells from human-induced pluripotent stem cells. *Cerebellum* <https://doi.org/10.1007/s12311-017-0913-2> (2018).
12. Sundberg, M. *et al.* Purkinje cells derived from TSC patients display hypoexcitability and synaptic deficits associated with reduced FMRP levels and reversed by rapamycin. *Mol. Psychiatry* <https://doi.org/10.1038/s41380-018-0018-4> (2018).
13. Buchholz, D. E. *et al.* Novel genetic features of human and mouse Purkinje cell differentiation defined by comparative transcriptomics. *Proc. Natl. Acad. Sci. USA* **117**, 15085–15095. <https://doi.org/10.1073/pnas.2000102117> (2020).
14. Silva, T. P. *et al.* Maturation of human pluripotent stem cell-derived cerebellar neurons in the absence of co-culture. *Front. Bioeng. Biotechnol.* **8**, 70. <https://doi.org/10.3389/fbioe.2020.00070> (2020).
15. Carter, R. A. *et al.* A Single-Cell Transcriptional Atlas of the Developing Murine Cerebellum. *Curr. Biol.* **28**, 2910–2920 e2912. <https://doi.org/10.1016/j.cub.2018.07.062> (2018).
16. Aldinger, K. A. *et al.* Spatial and single-cell transcriptional landscape of human cerebellar development. *bioRxiv*. <https://doi.org/10.1101/2020.06.30.174391> (2020).
17. Randles, M. J. *et al.* Basement membrane ligands initiate distinct signalling networks to direct cell shape. *Matrix Biol.* **90**, 61–78. <https://doi.org/10.1016/j.matbio.2020.02.005> (2020).
18. Lancaster, M. A. *et al.* Guided self-organization and cortical plate formation in human brain organoids. *Nat. Biotechnol.* **35**, 659–666. <https://doi.org/10.1038/nbt.3906> (2017).
19. Gibbs, H. C., Chang-Gonzalez, A., Hwang, W., Yeh, A. T. & Lekven, A. C. Midbrain-hindbrain boundary morphogenesis: At the intersection of Wnt and Fgf signaling. *Front. Neuroanat.* **11**, 64. <https://doi.org/10.3389/fnana.2017.00064> (2017).
20. Gutzman, J. H., Graeden, E. G., Lowery, L. A., Holley, H. S. & Sive, H. Formation of the zebrafish midbrain-hindbrain boundary constriction requires laminin-dependent basal constriction. *Mech. Dev.* **125**, 974–983. <https://doi.org/10.1016/j.mod.2008.07.004> (2008).
21. Giandomenico, S. L. *et al.* Cerebral organoids at the air-liquid interface generate diverse nerve tracts with functional output. *Nat. Neurosci.* **22**, 669–679. <https://doi.org/10.1038/s41593-019-0350-2> (2019).
22. Wizeman, J. W., Guo, Q., Wilton, E. M. & Li, J. Y. Specification of diverse cell types during early neurogenesis of the mouse cerebellum. *Elife* <https://doi.org/10.7554/eLife.42388> (2019).
23. Grimmer, M. R. & Weiss, W. A. BMPs oppose Math1 in cerebellar development and in medulloblastoma. *Genes Dev.* **22**, 693–699. <https://doi.org/10.1101/gad.1657808> (2008).
24. Gupta, I. *et al.* Single-cell isoform RNA sequencing characterizes isoforms in thousands of cerebellar cells. *Nat. Biotechnol.* <https://doi.org/10.1038/nbt.4259> (2018).
25. Rahimi-Balaei, M., Bergen, H., Kong, J. & Marzban, H. Neuronal Migration During Development of the Cerebellum. *Front. Cell Neurosci.* **12**, 484. <https://doi.org/10.3389/fncel.2018.00484> (2018).
26. Li, X. *et al.* Expression of prosaposin and its receptors in the rat cerebellum after kainic acid injection. *IBRO Re.* <https://doi.org/10.1016/j.ibror.2017.02.002> (2017).
27. Thul, P. J. *et al.* A subcellular map of the human proteome. *Science* **356**, eaal3321. <https://doi.org/10.1126/science.aal3321> (2017).
28. Williams, S. E. *et al.* Aspm sustains postnatal cerebellar neurogenesis and medulloblastoma growth in mice. *Development* **142**, 3921–3932. <https://doi.org/10.1242/dev.124271> (2015).
29. Leightner, A. C. *et al.* The Meckel syndrome protein meckelin (TMEM67) is a key regulator of cilia function but is not required for tissue planar polarity. *Hum. Mol. Genet.* **22**, 2024–2040. <https://doi.org/10.1093/hmg/ddt054> (2013).
30. Stauber, M. *et al.* Identification of FOXJ1 effectors during ciliogenesis in the foetal respiratory epithelium and embryonic left-right organiser of the mouse. *Dev. Biol.* **423**, 170–188. <https://doi.org/10.1016/j.ydbio.2016.11.019> (2017).

31. Eura, N. *et al.* Brainstem organoids from human pluripotent stem cells. *Front. Neurosci.* <https://doi.org/10.3389/fnins.2020.00538> (2020).
32. Pellegrini, L. *et al.* Human CNS barrier-forming organoids with cerebrospinal fluid production. *Science* **369**, eaaz5626. <https://doi.org/10.1126/science.aaz5626> (2020).
33. Wang, L. & Liu, Y. Signaling pathways in cerebellar granule cells development. *Am. J. Stem. Cells* **8**, 1–6 (2019).
34. Legué, E. *et al.* Differential timing of granule cell production during cerebellum development underlies generation of the foliation pattern. *Neural Dev.* **11**, 17. <https://doi.org/10.1186/s13064-016-0072-z> (2016).
35. Haldipur, P. *et al.* Spatiotemporal expansion of primary progenitor zones in the developing human cerebellum. *Science* **366**, 454–460. <https://doi.org/10.1126/science.aax7526> (2019).
36. Vilas-Boas, F. & Henrique, D. HES6-1 and HES6-2 function through different mechanisms during neuronal differentiation. *PLoS ONE* **5**, e15459. <https://doi.org/10.1371/journal.pone.0015459> (2010).
37. Fiedler, L. R. & Eble, J. A. Decorin regulates endothelial cell-matrix interactions during angiogenesis. *Cell Adh. Migr.* **3**, 3–6. <https://doi.org/10.4161/cam.3.1.7275> (2009).
38. Tech, K. *et al.* Pyruvate kinase inhibits proliferation during postnatal cerebellar neurogenesis and suppresses medulloblastoma formation. *Cancer Res.* **77**, 3217–3230. <https://doi.org/10.1158/0008-5472.CAN-16-3304> (2017).
39. Watanabe, M. *et al.* Contribution of thyrotropin-releasing hormone to cerebellar long-term depression and motor learning. *Front. Cell Neurosci.* **12**, 490. <https://doi.org/10.3389/fncel.2018.00490> (2018).
40. Wrigley, S., Arafa, D. & Tropea, D. Insulin-like growth factor 1: At the crossroads of brain development and aging. *Front. Cell Neurosci.* **11**, 14. <https://doi.org/10.3389/fncel.2017.00014> (2017).
41. Park, C. R. *et al.* Spatiotemporal expression and functional implication of CXCL14 in the developing mice cerebellum. *Mol. Cells* **34**, 289–293. <https://doi.org/10.1007/s10059-012-0116-0> (2012).
42. Mosquera, N., Rodriguez-Trillo, A., Mera-Varela, A., Gonzalez, A. & Conde, C. Uncovering Cellular retinoic acid-binding protein 2 as a potential target for rheumatoid arthritis synovial hyperplasia. *Sci. Rep.* **8**, 8731. <https://doi.org/10.1038/s41598-018-26027-x> (2018).
43. Duncan, M. K., Bordas, L., Diccio-Bloom, E. & Chada, K. K. Expression of the helix-loop-helix genes Id-1 and NSCL-1 during cerebellar development. *Dev. Dyn.* **208**, 107–114. [https://doi.org/10.1002/\(SICI\)1097-0177\(199701\)208:1%3c107::AID-AJA10%3e3.0.CO;2-3](https://doi.org/10.1002/(SICI)1097-0177(199701)208:1%3c107::AID-AJA10%3e3.0.CO;2-3) (1997).
44. Kant, S., Stopa, E. G., Johanson, C. E., Baird, A. & Silverberg, G. D. Choroid plexus genes for CSF production and brain homeostasis are altered in Alzheimer's disease. *Fluids Barriers CNS* **15**, 34. <https://doi.org/10.1186/s12987-018-0120-7> (2018).
45. Mizee, M. R. *et al.* Retinoic acid induces blood-brain barrier development. *J. Neurosci.* **33**, 1660–1671. <https://doi.org/10.1523/JNEUROSCI.1338-12.2013> (2013).
46. Grubman, A. *et al.* A CX3CR1 Reporter hESC line facilitates integrative analysis of in-vitro-derived microglia and improved microglia identity upon neuron-glia co-culture. *Stem Cell Rep.* **14**, 1018–1032. <https://doi.org/10.1016/j.stemcr.2020.04.007> (2020).
47. Bergersen, L. *et al.* A novel postsynaptic density protein: The monocarboxylate transporter MCT2 is co-localized with delta-glutamate receptors in postsynaptic densities of parallel fiber-Purkinje cell synapses. *Exp. Brain Res.* **136**, 523–534. <https://doi.org/10.1007/s002210000600> (2001).
48. Vuong, H., Truong, T., Phan, T. & Pham, S. Venice: A new algorithm for finding marker genes in single-cell transcriptomic data. *bioRxiv*. <https://doi.org/10.1101/2020.11.16.384479> (2020).
49. Le, T. *et al.* BBrowser: Making single-cell data easily accessible. *bioRxiv*. <https://doi.org/10.1101/2020.12.11.414136> (2020).
50. Rouillard, A. D. *et al.* The harmonizome: A collection of processed datasets gathered to serve and mine knowledge about genes and proteins. *Database (Oxford)* <https://doi.org/10.1093/database/baw100> (2016).
51. Capes-Davis, A., Tolhurst, O., Dunn, J. M. & Jeffrey, P. L. Expression of doublecortin (DCX) and doublecortin-like kinase (DCLK) within the developing chick brain. *Dev. Dyn.* **232**, 457–467. <https://doi.org/10.1002/dvdy.20240> (2005).
52. Gleeson, J. G., Lin, P. T., Flanagan, L. A. & Walsh, C. A. Doublecortin Is a Microtubule-Associated Protein and Is Expressed Widely by Migrating Neurons. *Neuron* **23**, 257–271. [https://doi.org/10.1016/S0896-6273\(00\)80778-3](https://doi.org/10.1016/S0896-6273(00)80778-3) (1999).
53. Larouche, M., Beffert, U., Herz, J. & Hawkes, R. The Reelin receptors Apoer2 and Vldlr coordinate the patterning of Purkinje cell topography in the developing mouse cerebellum. *PLoS ONE* **3**, e1653. <https://doi.org/10.1371/journal.pone.0001653> (2008).
54. Minocherhomji, S. *et al.* Epigenetic remodelling and dysregulation of DLGAP4 is linked with early-onset cerebellar ataxia. *Hum. Mol. Genet.* **23**, 6163–6176. <https://doi.org/10.1093/hmg/ddu337> (2014).
55. Rasmussen, A. H., Rasmussen, H. B. & Silahatoglu, A. The DLGAP family: Neuronal expression, function and role in brain disorders. *Mol. Brain* **10**, 43. <https://doi.org/10.1186/s13041-017-0324-9> (2017).
56. Yamazaki, Y. *et al.* Phosphorylation of CRMP2 is required for migration and positioning of Purkinje cells: Redundant roles of CRMP1 and CRMP4. *Brain Res.* **1736**, 146762. <https://doi.org/10.1016/j.brainres.2020.146762> (2020).
57. Yano, M., Hayakawa-Yano, Y., Mele, A. & Darnell, R. B. Nova2 regulates neuronal migration through an RNA switch in disabled-1 signaling. *Neuron* **66**, 848–858. <https://doi.org/10.1016/j.neuron.2010.05.007> (2010).
58. Law, A. J., Shannon Weickert, C., Hyde, T. M., Kleinman, J. E. & Harrison, P. J. Neuregulin-1 (NRG-1) mRNA and protein in the adult human brain. *Neuroscience* **127**, 125–136. <https://doi.org/10.1016/j.neuroscience.2004.04.026> (2004).
59. Estruch, S. B. *et al.* Proteomic analysis of FOXP proteins reveals interactions between cortical transcription factors associated with neurodevelopmental disorders. *Hum. Mol. Genet.* **27**, 1212–1227. <https://doi.org/10.1093/hmg/ddy035> (2018).
60. Kamath, S. P. & Chen, A. I. Myocyte enhancer factor 2c regulates dendritic complexity and connectivity of cerebellar purkinje cells. *Mol. Neurobiol.* **56**, 4102–4119. <https://doi.org/10.1007/s12035-018-1363-7> (2019).
61. Stoodley, C. J. The cerebellum and neurodevelopmental disorders. *Cerebellum* **15**, 34–37. <https://doi.org/10.1007/s12311-015-0715-3> (2016).
62. Aldinger, K. A. *et al.* Redefining the etiologic landscape of cerebellar malformations. *Am. J. Hum. Genet.* **105**, 606–615. <https://doi.org/10.1016/j.ajhg.2019.07.019> (2019).
63. Aldinger, K. A. & Doherty, D. The genetics of cerebellar malformations. *Semin. Fetal. Neonatal. Med.* **21**, 321–332. <https://doi.org/10.1016/j.siny.2016.04.008> (2016).
64. Willsey, A. J. *et al.* The psychiatric cell map initiative: A convergent systems biological approach to illuminating key molecular pathways in neuropsychiatric disorders. *Cell* **174**, 505–520. <https://doi.org/10.1016/j.cell.2018.06.016> (2018).
65. Prevalence and architecture of de novo mutations in developmental disorders. *Nature* **542**, 433–438. <https://doi.org/10.1038/nature21062> (2017).
66. Rauch, A. *et al.* Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: An exome sequencing study. *Lancet* **380**, 1674–1682. [https://doi.org/10.1016/s0140-6736\(12\)61480-9](https://doi.org/10.1016/s0140-6736(12)61480-9) (2012).
67. Feliciano, P. *et al.* Exome sequencing of 457 autism families recruited online provides evidence for autism risk genes. *NPJ Genomic Med.* **4**, 19. <https://doi.org/10.1038/s41525-019-0093-8> (2019).
68. Yuen, R. K. *et al.* Whole genome sequencing resource identifies 18 new candidate genes for autism spectrum disorder. *Nat. Neurosci.* **20**, 602–611. <https://doi.org/10.1038/nn.4524> (2017).
69. Ruzzo, E. K. *et al.* Inherited and de novo genetic risk for autism impacts shared networks. *Cell* **178**, 850–866.e826. <https://doi.org/10.1016/j.cell.2019.07.015> (2019).
70. Sanders, S. J. *et al.* Insights into autism spectrum disorder genomic architecture and biology from 71 risk loci. *Neuron* **87**, 1215–1233. <https://doi.org/10.1016/j.neuron.2015.09.016> (2015).

71. Yuen, R. K. *et al.* Whole-genome sequencing of quartet families with autism spectrum disorder. *Nat Med* **21**, 185–191. <https://doi.org/10.1038/nm.3792> (2015).
72. Beaudin, M., Klein, C. J., Rouleau, G. A. & Dupré, N. Systematic review of autosomal recessive ataxias and proposal for a classification. *Cerebellum Ataxias* **4**, 3. <https://doi.org/10.1186/s40673-017-0061-y> (2017).
73. McLoughlin, H. S., Moore, L. R. & Paulson, H. L. Pathogenesis of SCA3 and implications for other polyglutamine diseases. *Neurobiol. Dis.* **134**, 104635. <https://doi.org/10.1016/j.nbd.2019.104635> (2020).
74. Lancaster, M. A. *et al.* Publisher Correction: Guided self-organization and cortical plate formation in human brain organoids. *Nat. Biotechnol.* **36**, 1016. <https://doi.org/10.1038/nbt1018-1016a> (2018).
75. Halfter, W. & Yip, J. An organizing function of basement membranes in the developing nervous system. *Mech. Dev.* **133**, 1–10. <https://doi.org/10.1016/j.mod.2014.07.003> (2014).
76. Nguyen, H. *et al.* Glial scaffold required for cerebellar granule cell migration is dependent on dystroglycan function as a receptor for basement membrane proteins. *Acta Neuropathol. Commun.* **1**, 58. <https://doi.org/10.1186/2051-5960-1-58> (2013).
77. Eyrich, N. W., Potts, C. R., Robinson, M. H., Maximov, V. & Kenney, A. M. Reactive oxygen species signaling promotes hypoxia-inducible factor 1 $\alpha$  stabilization in sonic hedgehog-driven cerebellar progenitor cell proliferation. *Mol. Cell. Biol.* **39**, e00268-e1218. <https://doi.org/10.1128/mcb.00268-18> (2019).
78. Kullmann, J. A. *et al.* Oxygen tension and the VHL-Hif1 $\alpha$  pathway determine onset of neuronal polarization and cerebellar germinal zone exit. *Neuron* **106**, 607–623.e605. <https://doi.org/10.1016/j.neuron.2020.02.025> (2020).
79. Broom, E. R., Gilthorpe, J. D., Butts, T., Campo-Paysaa, F. & Wingate, R. J. The roof plate boundary is a bi-directional organiser of dorsal neural tube and choroid plexus development. *Development* **139**, 4261–4270. <https://doi.org/10.1242/dev.082255> (2012).
80. Chizhikov, V. V. *et al.* The roof plate regulates cerebellar cell-type specification and proliferation. *Development* **133**, 2793–2804. <https://doi.org/10.1242/dev.02441> (2006).
81. Krizhanovsky, V. & Ben-Arie, N. A novel role for the choroid plexus in BMP-mediated inhibition of differentiation of cerebellar neural progenitors. *Mech. Dev.* **123**, 67–75. <https://doi.org/10.1016/j.mod.2005.09.005> (2006).
82. Huang, X. *et al.* Transventricular delivery of Sonic hedgehog is essential to cerebellar ventricular zone development. *Proc. Natl. Acad. Sci. USA* **107**, 8422–8427. <https://doi.org/10.1073/pnas.0911838107> (2010).
83. Bhaduri, A. *et al.* Cell stress in cortical organoids impairs molecular subtype specification. *Nature* **578**, 142–148. <https://doi.org/10.1038/s41586-020-1962-0> (2020).
84. Huang, X. *et al.* Transventricular delivery of Sonic hedgehog is essential to cerebellar ventricular zone development. *Proc. Natl. Acad. Sci.* **107**, 8422–8427. <https://doi.org/10.1073/pnas.0911838107> (2010).
85. Yuste, R. *et al.* A community-based transcriptomics classification and nomenclature of neocortical cell types. *Nat. Neurosci.* <https://doi.org/10.1038/s41593-020-0685-8> (2020).
86. Regev, A. *et al.* The human cell atlas. *Elife* <https://doi.org/10.7554/eLife.27041> (2017).
87. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *bioRxiv*. <https://doi.org/10.1101/2020.10.12.335331> (2020).
88. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420. <https://doi.org/10.1038/nbt.4096> (2018).
89. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502. <https://doi.org/10.1038/nbt.3192> (2015).
90. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550. <https://doi.org/10.1186/s13059-014-0550-8> (2014).
91. Trapnell, C. *et al.* The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386. <https://doi.org/10.1038/nbt.2859> (2014).
92. McDavid, A. *et al.* Data exploration, quality control and testing in single-cell qPCR-based gene expression experiments. *Bioinformatics* **29**, 461–467. <https://doi.org/10.1093/bioinformatics/bts714> (2013).
93. Stuart, T. *et al.* Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902 e1821. <https://doi.org/10.1016/j.cell.2019.05.031> (2019).
94. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287. <https://doi.org/10.1089/omi.2011.0118> (2012).
95. Handel, A. E. *et al.* Assessing similarity to primary tissue and cortical layer identity in induced pluripotent stem cell-derived cortical neurons through single-cell transcriptomics. *Hum. Mol. Genet.* **25**, 989–1000. <https://doi.org/10.1093/hmg/ddv637> (2016).
96. Zappia, L., & Oshlack, A. Clustering trees: a visualization for evaluating clusterings at multiple resolutions. *GigaScience*. **7**(7). <https://doi.org/10.1093/gigascience/giy083> (2018).

## Acknowledgements

We thank D. Sims, M. Attar, H. Slawinski, and S. Revale for technical assistance and M. Cooper, L. Harris for their comments on the manuscript. S.N. is the recipient of a fellowship from the Oxford Nuffield Medical Trust. This work was supported by BrAshA-T, Action for A-T, the Rosetrees Trust, and the Wellcome Trust via core funding to the Wellcome Centre for Human Genetics (award 203141/Z/16/Z).

## Author contributions

S.N. and E.B. conceived the project. S.N. performed differentiation experiments, qPCR, immunocytochemistry, sample preparation, scRNAseq analysis and data visualisation. D.A. assisted with scRNAseq analysis, interpretation and data visualisation. F.C. assisted with scRNAseq analysis and interpretation. R.B. assisted with scRNAseq analysis and interpretation. S.N. and E.B. wrote the manuscript with feedback from all authors.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-91846-4>.

**Correspondence** and requests for materials should be addressed to S.N. or E.B.E.B.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.





**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021