nature neuroscience

Resource

https://doi.org/10.1038/s41593-023-01311-w

Single-cell transcriptomic landscape of the developing human spinal cord

Received: 8 December 2021

Accepted: 20 March 2023

Published online: 24 April 2023

Check for updates

Jimena Andersen[1,2,10,11], Nicholas Thom[1,2,11], Jennifer L. Shadrach[3], Xiaoyu Chen[1,2], Massimo Mario Onesto[1,2,4], Neal D. Amin[1,2], Se-Jin Yoon[1,2], Li Li[5], William J. Greenleaf[6,7], Fabian Müller[6,8], Anca M. Pașca[9], Julia A. Kaltschmidt[3] & Sergiu P. Pașca[1,2]

Understanding spinal cord assembly is essential to elucidate how motor behavior is controlled and how disorders arise. The human spinal cord is exquisitely organized, and this complex organization contributes to the diversity and intricacy of motor behavior and sensory processing. But how this complexity arises at the cellular level in the human spinal cord remains unknown. Here we transcriptomically profiled the midgestation human spinal cord with single-cell resolution and discovered remarkable heterogeneity across and within cell types. Glia displayed diversity related to positional identity along the dorso-ventral and rostro-caudal axes, while astrocytes with specialized transcriptional programs mapped into white and gray matter subtypes. Motor neurons clustered at this stage into groups suggestive of alpha and gamma neurons. We also integrated our data with multiple existing datasets of the developing human spinal cord spanning 22 weeks of gestation to investigate the cell diversity over time. Together with mapping of disease-related genes, this transcriptomic mapping of the developing human spinal cord opens new avenues for interrogating the cellular basis of motor control in humans and guides human stem cell-based models of disease.

The spinal cord plays a central role in integrating sensory and motor information to regulate movement. This is achieved by the coordinated function of motor neurons (MNs) and other neuronal and glial populations that are arranged in spatially distinct domains[1,2]. Damage or degeneration of the spinal cord can lead to devastating disorders, such as amyotrophic lateral sclerosis (ALS) or developmental disorders, such as spinal muscular atrophy or childhood leukodystrophies[3,4].

While cell diversity of the developing human neural tube and that of the mouse spinal cord has started to be unveiled[5,6], a comprehensive profiling of the developing human spinal cord at midgestation is needed. Access to data from human fetal samples is important as increasing evidence points at human-specific timing and characteristics in nervous system assembly[7–9]. For example, development of the cerebral cortex in humans takes place at a slower rate compared with rodents and other primates[7,10]. On the other hand, the human spinal cord has been proposed to develop and mature faster compared with rodents[11,12].

In this Resource, we used single-cell transcriptomics to generate a cell-type census and explore developmental landmarks in the fetal human spinal cord. We uncovered remarkable cellular diversity

[1]Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA, USA. [2]Stanford Brain Organogenesis, Wu Tsai Neurosciences Institute, Stanford, CA, USA. [3]Department of Neurosurgery, Stanford University, Stanford, CA, USA. [4]Neurosciences Graduate Program, Stanford University, Stanford, CA, USA. [5]Department of Human Genetics, Emory University, Atlanta, GA, USA. [6]Department of Genetics, Stanford University, Stanford, CA, USA. [7]Department of Applied Physics, Stanford University, Stanford, CA, USA. [8]Center for Bioinformatics, Saarland University, Saarbrücken, Germany. [9]Department of Pediatrics, Division of Neonatology, Stanford University, Stanford, CA, USA. [10]Present address: Department of Human Genetics, Emory University, Atlanta, GA, USA. [11]These authors contributed equally: Jimena Andersen, Nicholas Thom. e-mail: spasca@stanford.edu

along several axes and characterized the signature of putative alpha and gamma MNs. We anticipate that this single-cell transcriptomic landscape will be useful for understanding evolutionary innovations in the developing spinal cord and to improve organoid and assembloid models of disease.

Results

A transcriptomic profile of the developing human spinal cord

We used 10x Chromium to profile single cells and single nuclei from four samples at gestational week (GW) 17 and GW18 (Fig. 1a, Extended Data Fig. 1a,b and Supplementary Table 1). To increase the probability of capturing MNs, we also performed Thy1 immunopanning or NeuN sorting (Extended Data Fig. 1c). Following quality control, doublet removal and filtering (performed separately in cells and nuclei), we obtained transcriptomes for 112,554 cells and 34,884 nuclei (Extended Data Fig. 1b–i and Supplementary Tables 2 and 3). GW17 and GW18 samples were highly correlated and were analyzed together (Extended Data Fig. 1j–l). Single-cell and single-nucleus samples displayed some differences, as reported[13]. Specifically, we detected a higher number of molecules (nCount) and genes (nFeature) as well as a higher percentage of ribosomal and mitochondrial genes in cells, but a higher ratio of unspliced to spliced counts in nuclei (Extended Data Fig. 1i). Differential expression and Gene Ontology (GO) analysis showed that nuclei-enriched genes were related to synapse organization (*NRXN1* and *SHANK1*) and cell adhesion (*CDH18* and *PCDH9*), while genes enriched in single cells were related to housekeeping (*GAPDH*), cellular stress (*DNAJA1* and *HSPB1*) and immediate early gene response (*FOS* and *JUN*), differences possibly associated with dissociation artifacts (Extended Data Fig. 2a–c and Supplementary Table 4).

We used uniform manifold approximation and projection (UMAP) dimensionality reduction to visualize and cluster cells following the integration of single cells and single nuclei[14] (Fig. 1b and Supplementary Table 5). We found a group of astrocytes and cycling astrocytes as well as cells within the astroglia lineage (floor plate (FP), roof plate (RP) and midplate (MP)), cells in the oligodendrocyte lineage (oligodendrocyte progenitor cells (OPCs)/oligodendrocyte (oligo)) and a group of multipotent glial progenitor cells (mGPCs) linking the two lineages. We also identified a group of neurons that was enriched in the single-nucleus samples (Fig. 1c), and groups of vascular cells including endothelial cells (ECs), pericytes and vascular leptomeningeal cells (VLMCs), Schwann cells and immune cells including microglia and monocytes (MG/immune) (Fig. 1d,e). Cycling cells were distributed within the UMAP landscape and could be found as part of the OPC/oligo, VLMC, EC, MG/immune and Schwann cell groups (Extended Data Fig. 1m).

Pseudotime analysis in the OPC/oligo lineage

We first examined OPC/oligo cells (Fig. 2a). Subclustering identified cell types along the oligodendrocyte lineage (Fig. 2b, Extended Data Fig. 3a,b, Supplementary Fig. 1a and Supplementary Table 6): OPCs, cycling OPCs and OPC-like mGPCs, differentiation-committed oligo precursors (COP), newly formed oligos (NFOL), myelin-forming oligos (MFOL) and mature oligos (MOL).

To explore developmental progression in the OPC/oligo lineage, we computed RNA velocity[5,15]. Diffusion pseudotime analysis revealed putative origins of the trajectory in cycling cells (Extended Data Fig. 3c) and inferred a differentiation trajectory. We annotated each cell with a pseudotime value (Fig. 2c) and grouped variable genes along this pseudotime into ten clusters (Fig. 2d). GO analysis highlighted a sequence of cellular processes along the trajectory. Early in the pseudotime we found terms that included cell division, and terms associated with potassium transport and synapse assembly. Genes expressed in NFOL and MFOL were enriched for terms associated with axon guidance and migration, while MFOL and MOL were enriched for myelination-related genes (Extended Data Fig. 3d). For each major cell type, we identified marker genes exhibiting particularly high estimated velocities, some of which manifested in increased expression downstream in the trajectory (Fig. 2e,f and Supplementary Table 7). For instance, *GPC3* was characterized by high velocity in COPs and high expression in NFOLs followed by negative velocities indicating downregulation in MFOLs.

Diversity of astrocytes in the developing human spinal cord

In the rodent spinal cord, astrocytes are spatially and functionally heterogeneous[2,16–19]. To investigate this diversity in humans, we next subclustered astroglia (Fig. 3a–c, Extended Data Fig. 4a,b, Supplementary Fig. 1b and Supplementary Table 8). We found a group of glial progenitors corresponding to ventricular zone cells and dorsal and ventral mGPCs that expressed *SOX9*, but lower levels of the more mature astrocyte markers *AQP4* and *SPARCL1* (Extended Data Fig. 4c), cycling cells (Extended Data Fig. 4d,e) and a group of cells we called 'active' on the basis of their expression of activity-dependent genes such as *FOS*, *EGR1* and *ARC*. These cells were found almost exclusively in single-cell samples (Extended Data Fig. 4a) and their signature did not overlap with that of reactive astrocytes (Extended Data Fig. 4f) (ref. 20), suggesting a signature probably related to the dissociation method. In addition to these, we found nine clusters of astrocytes along two axes of diversity.

The first axis corresponded to dorso-ventral positioning, with some astroglia expressing dorsal-related genes (*PAX3* and *ZIC1*), and others expressing ventral-related genes (*EPHA5* and *SEMA3A*; Fig. 3d), in line with what has been described in the mouse spinal cord[16]. Within the ventral group, we also found a population of V1 astrocytes characterized by the expression of *RELN* and the absence of *SLIT1* (ref. 18) (Extended Data Fig. 4g–i). We validated the dorso-ventral division by immunohistochemistry with PAX3 and NKX6-1 (Extended Data Fig. 4j,k). In addition, we found that homeodomain and forkhead transcription factors that are generally associated with positional identity in the spinal cord were expressed in a subtype-specific pattern (for example, *PAX7*, *IRX2*, *DBX2*, *FOXP2*, *NKX6-1*, *NKX2-2* and *NKX6-2*) (Extended Data Fig. 4l). The second axis of diversity corresponded to white matter and gray matter astrocytes[9] (Extended Data Fig. 5a–c and Supplementary Tables 9–11). White matter or fibrous astrocytes were characterized by the expression of *CRYAB* and *ID3*, as well as higher expression of *GFAP*[9,21] (Fig. 3e and Extended Data Fig. 5a). Gray matter or protoplasmic astrocytes were characterized by the expression of the amino acid transporter *SLC7A10*, and the inward-rectifying potassium channel *KCNJ16* (*Kir5.1*; Fig. 3e), and higher expression of the excitatory amino acid transporter *SLC1A2* and glutamine synthetase *GLUL* (Extended Data Fig. 5a). We validated the presence of some these markers in the white and gray matter. While CRYAB, GLUL and SLC1A2 all colocalized with pan-astrocyte markers AQP4 and GFAP (Extended Data Fig. 5d,e), CRYAB was specifically expressed in the white matter in regions where myelinated SMI-312 axons were present (Fig. 3f and Extended Data Fig. 5f,g), and GLUL and SLC1A2 were specifically expressed in the gray matter (Fig. 3f). Quantification of the fluorescence intensity of each of these markers further validated this (Fig. 3g). Further examination of these axes of division allowed us to discover new astrocyte markers. For example, we found that the transmembrane glycoprotein CD38 was specifically expressed in ventral fibrous astrocytes (Fig. 3h and Extended Data Fig. 5h).

Astrocytes play crucial roles in the development and maintenance of neuronal function including neurotransmitter and potassium homeostasis, synapse formation and elimination, and blood–brain barrier function[22–25]. Therefore, we next plotted genes associated with astrocyte function and asked if astrocyte positional or anatomical diversity was linked to functional diversity (Fig. 3i and Extended Data Fig. 5i). We found that all types of astrocytes express Na^+–K^+–ATPases and potassium channels; however, different astrocyte subtypes specifically expressed different types (for example, fibrous: *ATP1A2*, *ATP1B2*, *ATP1B1* and *ATP1A1*; protoplasmic: higher *ATP1A2* and *ATP1B2*). We also found that different astrocyte subtypes express different ionotropic and metabotropic receptors (fibrous: *GRM7* and *GRM8*; protoplasmic:

https://doi.org/10.1038/s41593-023-01311-w

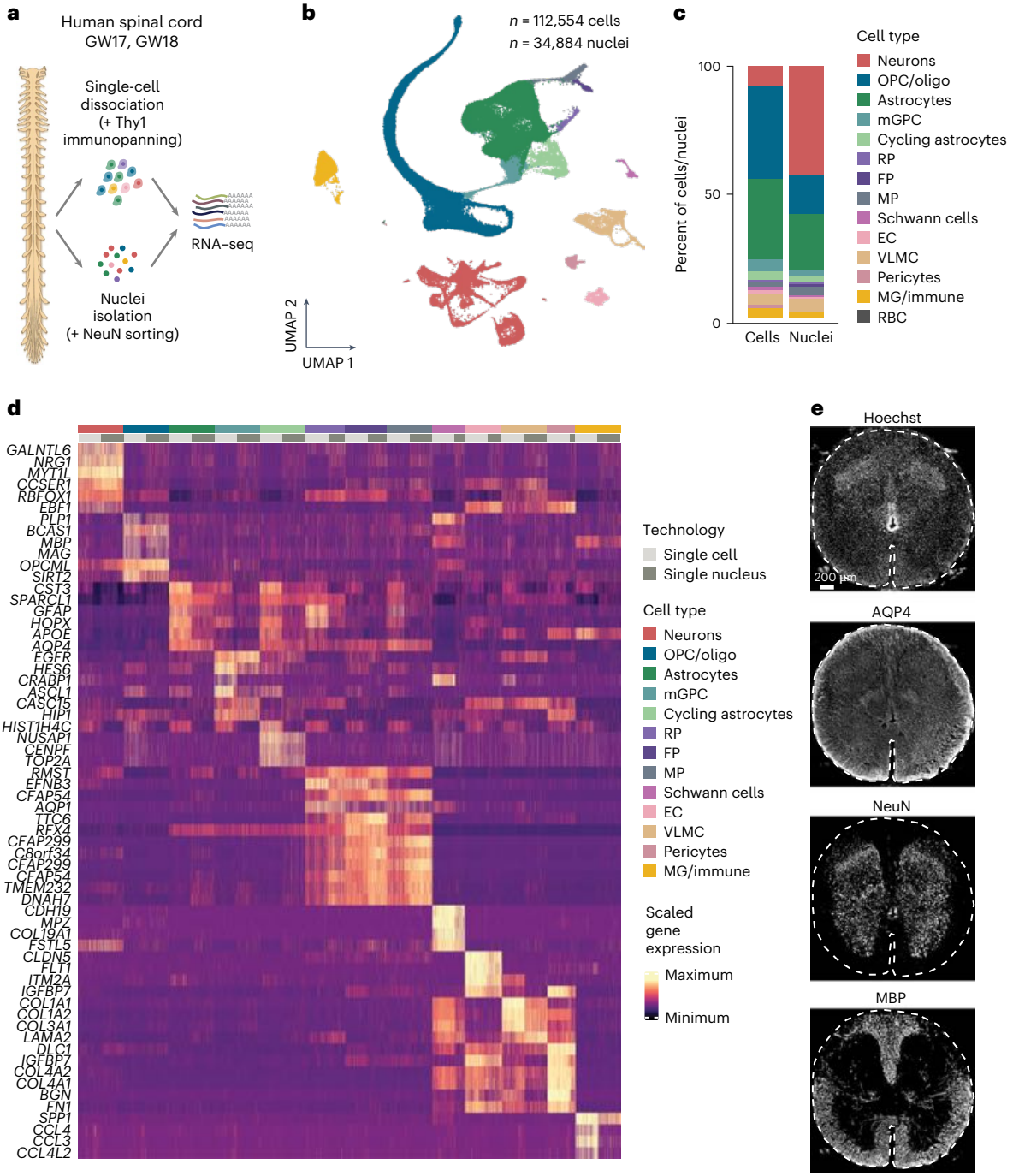


Fig. 1 | A single-cell atlas of the human spinal cord at midgestation. a, A schematic showing an overview of the samples and methods used for the study. **b**, UMAP representation of integrated single cells and single nuclei, colored by cell type. RBC, red blood cell. **c**, A bar plot showing the percent of cell types in single-cell and single-nucleus samples. **d**, A heat map showing normalized expression of the top four to six unique markers per cell type shown separated by isolation method (Technology), and downsampled to at most 500 cells per cell type and isolation method. **e**, Representative immunohistochemistry images of AQP4, NeuN (RBFOX3) and MBP in a coronal thoracic spinal cord cryosection at GW19. Hoechst shows nuclei. Immunohistochemistry was replicated four times in two independent samples. Scale bar, 200 µm (**e**).

GRM3 and *GRM5*). Moreover, while most astrocytes are highly permeable to Ca^{2+}, dorsal protoplasmic astrocytes are not, since they specifically express the AMPA receptor subunit *GluA2* (refs. 26,27). Finally, we found that protoplasmic astrocytes have a higher expression of neurotransmitter transport genes (for example, *SLC1A2*, *SLC7A10* and *SLC6A9*), suggesting a more prominent role in synaptically released neurotransmitter clearance. In contrast, we found that fibrous astrocytes expressed genes involved in the regulation of neuronal transmission at the node of Ranvier. For example, they highly express the membrane proteins ezrin (*EZR*) and radixin (*RDX*), which mediate the motility of peripheral astrocyte processes[28]. We also found that both secreted extracellular matrix-binding proteins tenascin-C (*TNC*) and tenascin-R (*TNR*) were highly expressed in fibrous astrocytes. TNC and TNR bind and cluster sodium channels at the nodes of Ranvier[29]. On the basis of these data and on their physical position in the white matter, we next wondered if fibrous astrocytes might be interacting with oligos

https://doi.org/10.1038/s41593-023-01311-w

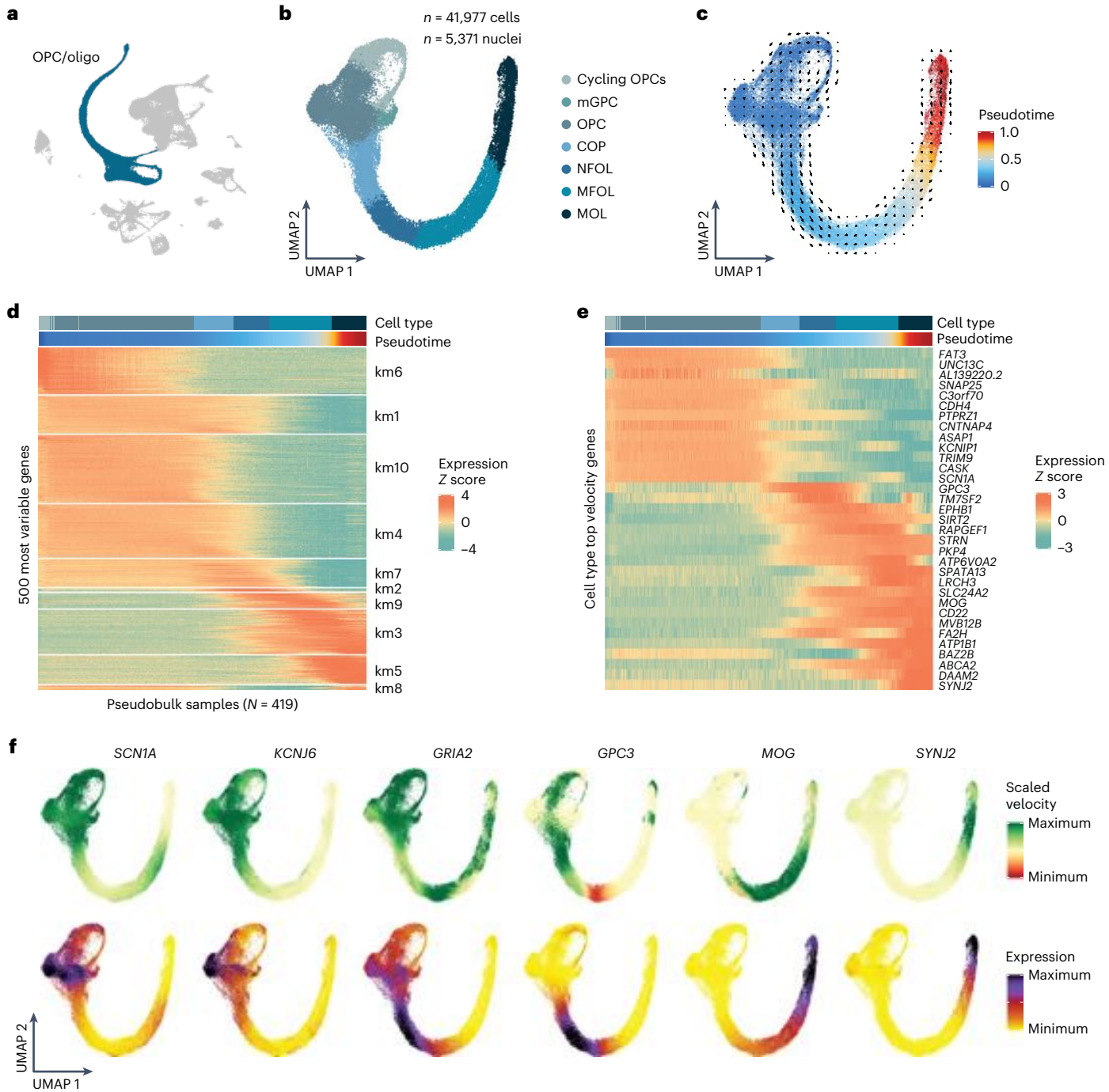


Fig. 2 | OPC/Oligo lineage in the spinal cord. a, Highlight of the OPC/oligo lineage in the main UMAP, showing the cells selected for subclustering. **b**, UMAP of the OPC/oligo subcluster colored by cell type. **c**, A UMAP plot showing the OPC/oligo subcluster colored by the diffusion pseudotime using single-cell velocity as computed by scVelo. The arrow overlay shows aggregate projected velocities. **d**, A heat map showing scaled expression of the 500 most variable genes across 419 pseudobulk samples aggregated along pseudotime bins. Genes were clustered into ten groups using *k* means. **e**, A heat map showing scaled expression of cell type dynamic genes. For each cell type the five highest-ranking genes according to velocity are shown. **f**, UMAPs showing examples of dynamic genes with cell-type-specific velocity along the trajectory showing computed velocity (top) and normalized expression (bottom).

to play other roles in the spinal cord. We performed a cell–cell interaction analysis using the tool Network Analysis Toolkit for Multicellular Interactions (NATMI) and found that ventral fibrous astrocytes and MOL showed the highest specificity of interactions (Extended Data Fig. 5j,k and Supplementary Table 12), which suggested that crosstalk between these two cell types might play roles in trophic support, cell adhesion and guidance, and blood–brain barrier homeostasis.

Midline glia and mGPCs in the developing human spinal cord

We next explored the diversity of progenitors. First, we focused on cells lining the ventricular zone (VZ), which included RP, FP and MP cells (Fig. 4a,b and Supplementary Table 13). RP cells expressed *GDF7*, *GDF10*, *ZIC1* and *MSX1* (Fig. 4c and Extended Data Fig. 6a), and FP cells expressed *SHH* and *FOXA2* (Fig. 4c), as well as PAX7 as previously described[6,30,31] (Extended Data Fig. 6b). We validated ZIC1, FOXA2 and PAX7 expression

https://doi.org/10.1038/s41593-023-01311-w

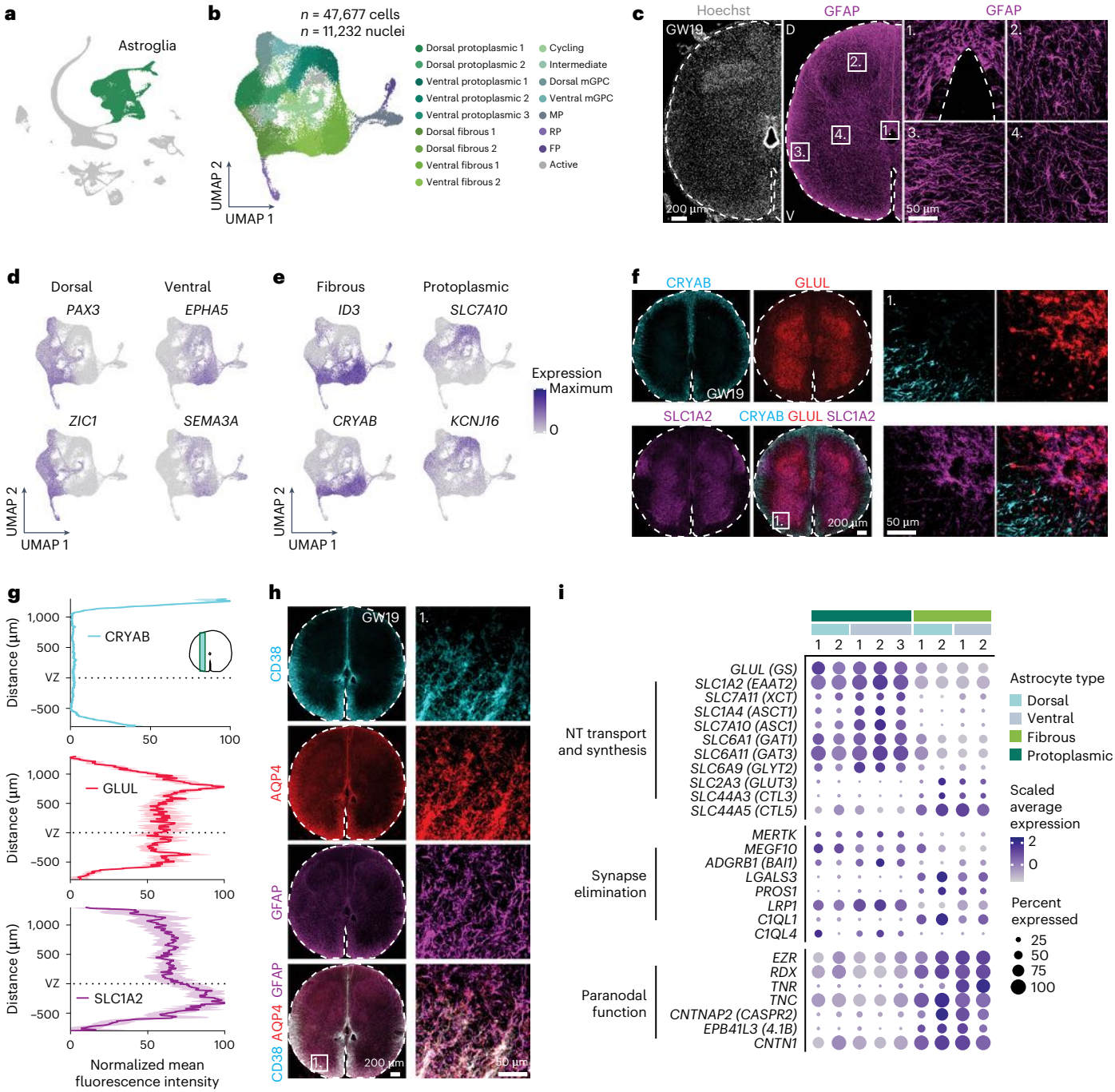


Fig. 3 | Astrocytes in the spinal cord show diversity along several axes. **a**, Highlight of astroglia in the main UMAP, showing the cells selected for subclustering. **b**, UMAP of the astroglia subcluster colored by type. **c**, A representative immunohistochemistry image for the astroglia marker GFAP in a coronal upper thoracic spinal cord cryosection at GW19. Insets (numbers 1–4) show different types of astroglia in the spinal cord. Immunohistochemistry was replicated eight times in two independent samples. D, dorsal; V, ventral. **d**, UMAP plots showing gene expression of markers associated with dorsal and ventral identities in the astroglia subcluster. **e**, UMAP plots showing gene expression of markers associated with fibrous and protoplasmic astrocyte subtypes. **f**, Representative immunohistochemistry images showing the fibrous astrocyte marker CRYAB and the protoplasmic astrocyte markers GLUL and SLC1A2. **g**, Quantification of the normalized mean fluorescence intensity for CRYAB, GLUL and SLC1A2 in the spinal cord at GW19 ($n = 5$, 4 and 4 sections for CRYAB, GLUL and SLC1A2, respectively). **h**, A representative immunohistochemistry image showing the ventral fibrous marker CD38 with the pan-astrocyte markers AQP4 and GFAP in a coronal spinal cord cryosection at GW19. Inset (number 1) shown on the right. The immunohistochemistry combination was replicated once in one independent sample, and independently 9, 10 and 16 times, respectively, in two independent samples. **i**, A dot plot showing the expression of selected genes associated with astrocyte functions in fibrous and protoplasmic subtypes. NT, neurotransmitter. The size of the dots represents the percent of cells expressing each gene, while the color depicts the scaled average expression per subtype. Scale bars, 50 µm (insets in **c** and **h**) and 200 µm (in **c**, **f** and **h**).

in cryosections (Extended Data Fig. 6c,d). We found FOXA2$^+$ FP cells dissociating from the VZ wall (Extended Data Fig. 6d), a process that in mice takes place starting at embryonic day (E)16.5 (ref. 32). We also noticed that, although cells in the VZ showed distinct domains of expression along the dorso-ventral axis (Extended Data Fig. 6e–h), this organization was different to that present at earlier stages[6,33]. For example, we

https://doi.org/10.1038/s41593-023-01311-w

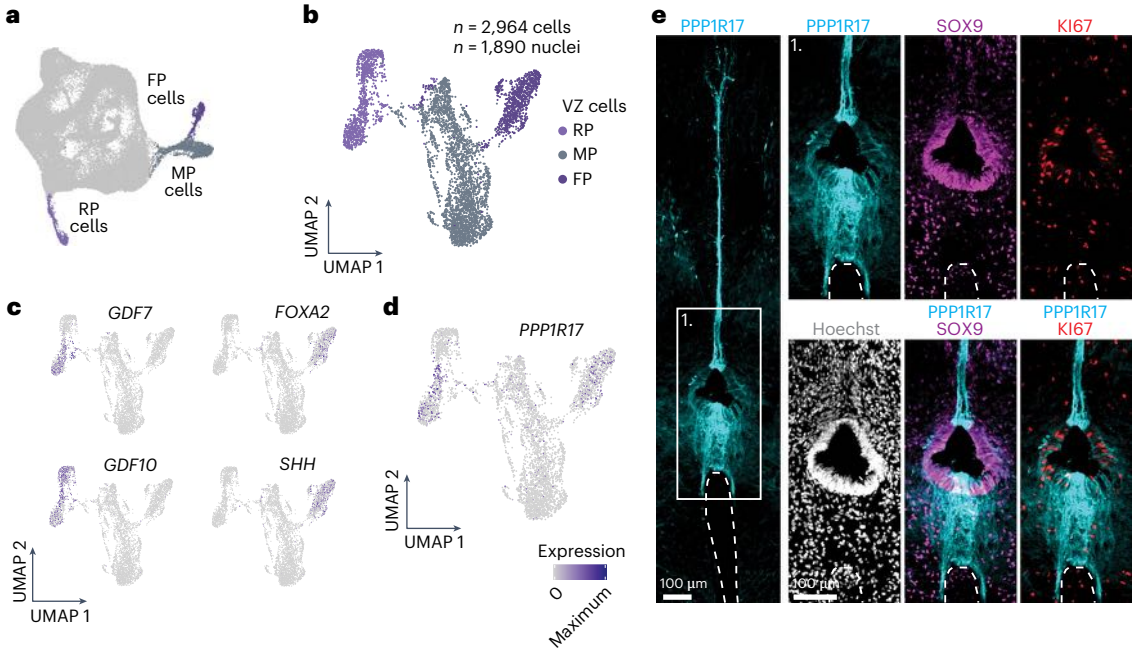


Fig. 4 | Midline glia in the developing human spinal cord. a, Highlight of ventricular zone (VZ) cells in the astroglia subcluster, highlighting the cells selected for subclustering. **b**, UMAP of the VZ cell subcluster, colored by cell type. **c**, UMAP plots showing gene expression of markers specific to RP (*GDF7* and *GDF10*) and FP (*FOXA2* and *SHH*) cells. **d**, UMAP plot showing expression of *PPP1R17* in RP and FP cells. **e**, Representative immunohistochemistry images showing PPP1R17, SOX9 and the cell cycle marker KI67 in VZ cells in a coronal upper thoracic spinal cord cryosection. Hoechst shows nuclei. Immunohistochemistry replicated four times in two independent samples. Scale bars, 100 µm (**e**).

found that the p3, pMN and p2 marker NKX6-1 was not restricted to ventral domains, but was instead expressed along the entire MP and FP (Extended Data Fig. 6g,h), similar to the expanded expression of NKX6-2 observed at human embryonic stages[6]. Moreover, the pMN marker OLIG2 was not expressed in the VZ at this stage (Extended Data Fig. 6i,j), although oligodendrogenesis is still ongoing[34].

Closer inspection of the FP and RP cells revealed expression of the phosphatase inhibitor *PPP1R17* (Fig. 4d), a marker for intermediate progenitor cells in the cerebral cortex[35,36] with human-specific gene regulation[37]. We performed immunohistochemistry and were surprised to find *PPP1R17* expressed in a group of FP and RP cells and along the dorsal and ventral midline (Fig. 4e). Their position and radial-like morphology extending from the VZ to the pia is suggestive of a midline glia-like or Nestin$^+$ radial glia identity and reminiscent of roof MSX1$^+$ putative radial quiescent cells[31,38]. Midline glia in the *Drosophila* embryo express highly conserved guidance molecules and play important roles in directing axonal crossing[39]. Analysis of FP and RP *PPP1R17*$^+$ cells indicated that they express ligands and receptors important for midline crossing[40], such as Netrin (*NTN1*), SLITs, semaphorins and ephrins (for example, *EFNB3*; Extended Data Fig. 6k and Supplementary Tables 14 and 15). Moreover, immunohistochemistry with the axonal marker SMI-312 and the sensory neuron marker substance P showed axons crossing both the dorsal and ventral midlines at GW19 (Extended Data Fig. 6l,m). We found minimal expression of Ppp1r17 in mouse spinal cord ventricular zone cells (Extended Data Fig. 6n–t). Future experiments should determine if PPP1R17 plays human-specific roles in the spinal cord.

In the midgestation human cerebral cortex, we identified a group of multipotent EGFR$^+$/OLIG2$^+$/ASCL1$^+$ glial progenitors (mGPCs) that may generate both astrocytes and oligos[36]. Similarly, previous studies have identified Olig2-expressing astrocytes generated from the MN and oligo-generating pMN domain in the developing rodent spinal cord[17,41,42], Ascl1-expressing glial progenitors in the spinal cord giving rise to oligos and dorsally restricted astrocytes[43], and dual lineage Olig2-expressing precursor cells in the adult subventricular zone in the brain[44] and injured spinal cord[45]. Here we found a similar population expressing *EGFR*, *OLIG2* and *ASCL1* (Extended Data Fig. 7a,b) that formed a bridge between the astroglia and OPC/oligo clusters, suggestive of multipotency; half of these cells expressed astroglia markers (*SOX9* and *AQP4*), while the other half expressed OPC/oligo markers (*SOX10* and *PDGFRA*) (Extended Data Fig. 7c). Immunohistochemistry for EGFR and OLIG2 identified cells positioned mainly in the intermediate gray matter adjacent to the VZ (Extended Data Fig. 7d,e, inset 1) expressing either the astroglia marker SOX9 or the oligo lineage marker NKX2.2, consistent with mGPCs.

Early MN identity specification in humans

Next, we investigated cell diversity in spinal cord neurons (Fig. 5a). We subclustered all neurons and found groups with different neurotransmitter identities (Fig. 5b–d, Extended Data Fig. 8a–c, Supplementary Fig. 2a and Supplementary Table 16), including glutamatergic (Glut: *SLC17A6*), GABAergic and glycinergic (GABA and GABA/Glyc.: *GAD1*, *GAD2*, *SLC32A1* and *SLC6A5*) and cholinergic (*SLC5A7*, *SLC18A3* and *CHAT*). We validated the presence of these neuronal types by immunohistochemistry (Fig. 5e). Next, we used label transfer to compare our cell types with a harmonized atlas of cell types in the mouse spinal cord[46] (Fig. 5f and Extended Data Fig. 8d; for label transfer of individual datasets, see Extended Data Fig. 8g–m; Supplementary Table 17)[47–51]. This revealed the presence of different types of excitatory, inhibitory and cholinergic neurons, as well as dorsal, ventral and intermediate types. Comparison of the top genes per cluster revealed similarities (for example, SSTR2 and RORB; Extended Data Fig. 8e,f and Supplementary Table 18). However, out of the top 50 markers per cluster from Russ et al.[46] and the top 50 markers from our dataset, only 22% of genes were shared. Using all markers from our clusters with an adjusted *P* value of 0.01 revealed a ~55% overlap with these same top 50 mouse markers. This is consistent with transcriptomic studies comparing mice and

human in other parts of the nervous system[21,52], although here the different ages of the samples between human and mouse could account for some of the differences.

We next focused on cholinergic neuron diversity. Of the two cholinergic clusters we identified, one expressed visceral MN markers (*NOS1*) and the other expressed markers of skeletal MNs (*RET* and *GFRA1*; Fig. 5g, Extended Data Fig. 9a and Supplementary Table 19). Subclustering revealed that skeletal MNs were composed of two subgroups (Fig. 5h). Using label transfer of MNs from adult mouse spinal cord[53], we found that GW17–18 visceral MNs mapped onto mouse adult visceral MNs, while skeletal MNs were divided into α and γ subgroups (Fig. 5i–k, Extended Data Fig. 9b–e, Supplementary Fig. 2b,c and Supplementary Tables 20 and 21). Comparison of human α and γ MNs with adult mouse counterparts revealed some similarities (Extended Data Fig. 9f,g and Supplementary Tables 22 and 23). For example, αMNs expressed *CHODL*, *SV2B* and *VIPR2* in both datasets, while γMNs expressed *ESRRB* and *WLS*. In addition, we found differences among the two datasets that, in part, might highlight the different ages sampled. For example, we found that, unlike adult mouse αMNs, human GW17–18 αMNs expressed several type II cadherins (*CDH7*, *CDH9*, *CDH10* and *CDH20*) that might regulate motor pool segregation[54,55]. Other differences included the expression of *UNC5B*, *TXK* and *DGKK* in human αMNs, and the expression of *NDNF*, *CLIC5* and *GFRA2* in human γMNs. To validate the expression of some of these markers in human MNs, we performed RNAScope for the α marker *CHODL* and the novel γ marker *NDNF* at GW18 and GW19, both of which were specific to MNs at this stage (Extended Data Fig. 9h,i). We found that *CHODL* and *NDNF* were expressed in *CHAT*$^+$ cells in the ventral horn, but did not colocalize (Fig. 5l and Extended Data Fig. 9j), indicating their presence in different types of MNs. Interestingly, we did not see this separation between *Chodl* and *Ndnf* in mouse spinal cord sections at P0 or P25 (Extended Data Fig. 9k,l). At both ages *Chodl* was expressed in a proportion of MNs, but *Ndnf* was either not expressed or was expressed at low levels in *Chodl* neurons, indicating that *NDNF* may be enriched in developing γMNs in human.

Mapping disease-related genes

We next mapped the expression of genes associated with disease in our spinal cord dataset. First, we focused on genes associated with myelin-related disorders, including leukodystrophies and degenerative white matter disorders (Fig. 6a and Supplementary Table 24). We found a subgroup of genes expressed specifically in myelinating oligos or Schwann cells (*PLP1*, *FA2H* and *GJC2*) and a subgroup expressed in different types of astrocytes (*GJA1*, *GFAP* and *EDNRB*) or in microglia (*CSF1R* and *TREM2*), suggesting that interactions between different cell types may contribute to pathogenesis[56]. Next, we focused on genes linked to the Charcot–Marie–Tooth neuropathy (CMT, Fig. 6b). We found a group of genes expressed in the OPC/oligo lineage or Schwann cells (*MPZ* and *PMP22*), astrocytes (*FGD4*) and microglia (*EGR2*), as well as genes that appeared to be ubiquitously expressed (*KIF1B* and *SBF2*). Interestingly, some genes associated with myelin-related disorders and CMT were also high-velocity marker genes in our pseudotime analysis (for example, *FA2H* in NFOL, *HSPB1* in MFOL and *GNB4* in COP). Lastly, we mapped the expression of genes associated with ALS, which is characterized by the selective loss of MNs (Fig. 6c). We found that, although some genes appeared to be highly expressed in neurons (*MAPT* and *UNC13A*) or specifically expressed in MNs (*NEFH* and *PRPH*), the majority of them were ubiquitously expressed (*SOD1*), or present in non-neuronal cell types such as OPC/oligo (*MOBP*) or VLMC, pericytes and microglia (*SQSTM1* and *GRN*), highlighting the role of non-cell autonomous toxicity[57]. We have generated an interactive online browser to further explore the dataset, which can be found at http://devspinalcord.su.domains.

Developing human spinal cord dataset integration

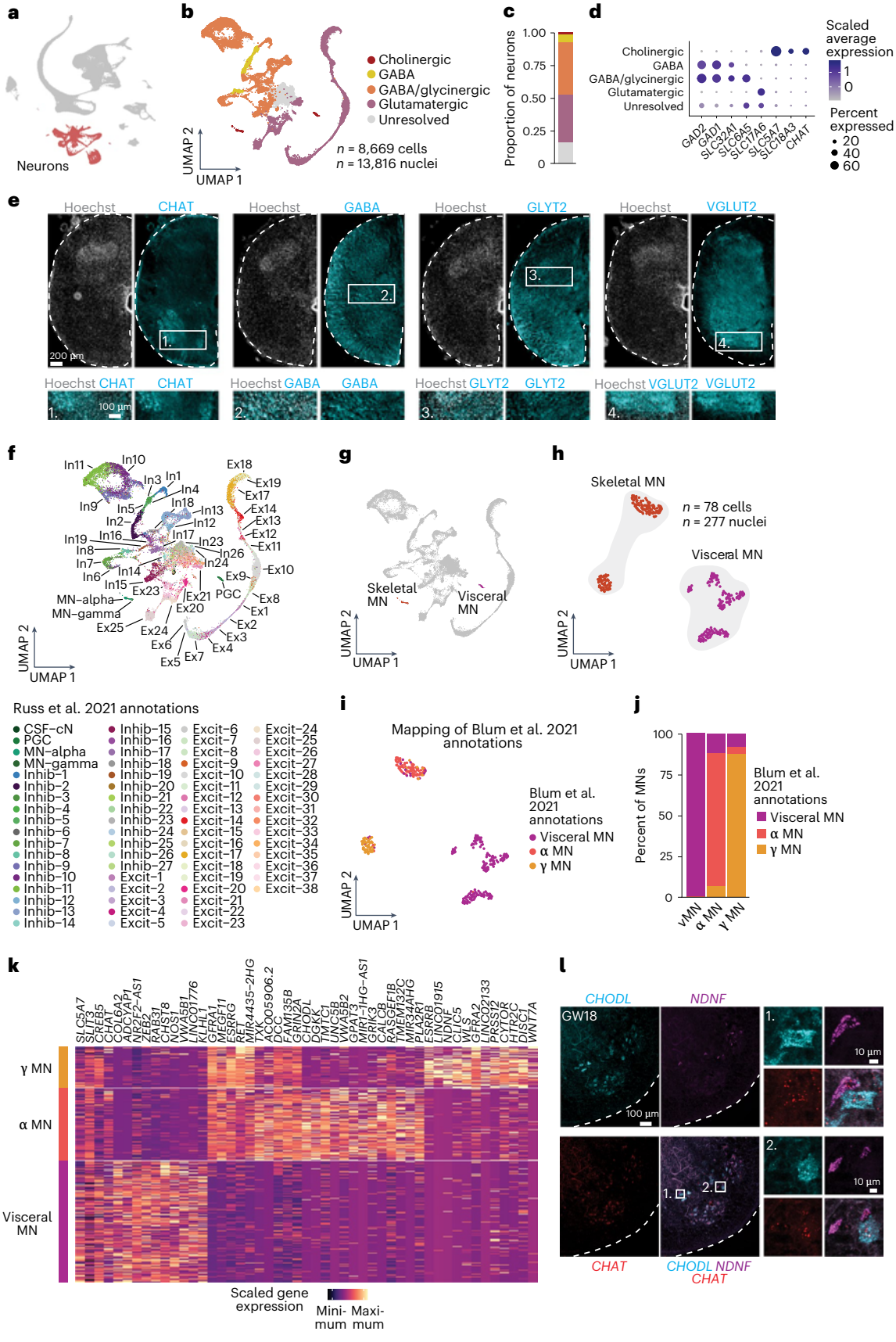
To place our dataset into a broader developmental context, we integrated our dataset with two other developing human spinal cord datasets[6,58]. The integrated dataset includes 35 samples spanning 22 weeks of gestation with a total of 950,215 cells and nuclei after quality control (QC) and filtering (Fig. 7a,b; for details of integration, see Methods; briefly, downsampling was performed to obtain variable features and label the main cell types based on our cell type annotations using nearest neighbors, Supplementary Figs. 3–6). Using these annotations, we then quantified the percentage of each cell type over time. In the OPC/oligo subcluster we found, not surprisingly, that clusters further along the inferred trajectory in pseudotime (for example, myelinating oligos), were enriched at later timepoints (Fig. 7c). In a similar way, within the astroglia subcluster, we observed a reduction in the percentage of progenitor clusters (for example, cycling, FP, RP and MP) over developmental time (Fig. 7d). In addition, we found that protoplasmic astrocytes represented a higher proportion of astrocytes earlier in development, while fibrous astrocytes appeared to be more abundant at late mid-fetal stages (Fig. 7d), suggesting that protoplasmic astrocytes may acquire their identity before fibrous astrocytes. In fact, immunohistochemistry at the early fetal 1 stage (GW11), when astrocytes and oligos start to populate the ventral spinal cord (Fig. 7e), showed the presence of the protoplasmic marker GLUL in the ventral spinal cord, and absence of the fibrous marker CRYAB at the same stage (Fig. 7f). Further validation also confirmed the presence of PPP1R17$^+$ midline glia and EGFR$^+$ mGPCs expressing either NKX2-2 or SOX9 at GW11 (Extended Data Fig. 10a–f).

We next wondered whether the developmental dynamics we observed were similar along the rostro-caudal axis. We used data from samples that were divided into cervical, thoracic and lumbar regions (Extended Data Fig. 1d–f and Supplementary Figs. 4 and 5). We found no clear region-specific distribution in the OPC/oligo cluster. In contrast, in the astroglia cluster, we found that thoracic regions had a lower proportion of protoplasmic astrocytes compared with cervical and lumbar regions. This was also the case at all the timepoints we analyzed (Fig. 7g,h and Extended Data Fig. 10g). On closer inspection, we found that the ventral proto 2 cluster of astrocytes was depleted in thoracic regions (thoracic cells represented 7.4% in the ventral proto cluster 2, and 37.9% on average in the rest of the clusters; Extended Data Fig. 10h–j).

Fig. 5 | Neuron and MN diversity in the spinal cord. a, Highlight of neurons in the main UMAP, highlighting the cells selected for subclustering. **b**, UMAP of the neuron subcluster, colored by neurotransmitter identity. **c**, A bar plot showing the proportion of cells within the neuron subcluster in each neurotransmitter identity group. **d**. Bubble plot showing expression of neurotransmitter identity-related genes as average scaled expression per group. *SLC32A1*, *SLC6A5*, *SLC17A6*, *SLC5A7* and *SLC18A3* are also known as *VGAT*, *GLYT2*, *VGLUT2*, *CHT1* and *VACHT*, respectively. **e**, Representative immunohistochemistry images showing neurons with different neurotransmitter identities (CHAT shows cholinergic neurons, GLYT2 shows glycinergic neurons, GABA shows GABAergic neurons and VGLUT2, also known as SLC17A6, shows glutamatergic neurons) in coronal spinal cord cryosections at GW19. Insets (numbers 1–4) shown on the bottom. Immunohistochemistry was replicated nine, three, three and four times in two independent samples. **f**, Label transfer showing neuronal annotations from Russ et al.[46] in the neuron subcluster. For more details, including common markers for each group, see Supplementary Table 18. **g**, Highlight of skeletal and visceral MNs in the neuron subcluster. **h**, MN subcluster colored by identity. **i**, MN subcluster showing the predicted identities of cells on the basis of Blum et al.[53] label transfer. **j**, Bar plot showing the percent of predicted identities in the MN subcluster (this study) on the basis of annotations from Blum et al.[53]. **k**, A heat map of scaled gene expression across cells in the MN subcluster. The parameter min.diff.pct in 'FindMarkers' was set to 0.3 to select features specific to each group. **l**, Representative in situ hybridization of *CHAT*, *CHODL* and *NDNF* in coronal lumbar spinal cord cryosections at GW18. Insets show *CHAT*$^+$ MNs that express either the alpha marker *CHODL* or the gamma marker *NDNF*. Replicated in $n = 2$ samples. Scale bars, 10 µm (insets in **l**), 100 µm (**l** and insets in **e**) and 200 µm (**e**).

https://doi.org/10.1038/s41593-023-01311-w

Finally, we used this dataset to investigate the timing of MN diversification. We subclustered cholinergic cells from the integrated neuron dataset and identified the α and γ MN clusters by creating modules of expression on the basis of the markers we previously identified, as well as mouse MN markers from Blum et al.[53] and Russ et al.[46] (Extended Data Fig. 10k,l and Supplementary Fig. 7a). We then quantified the



https://doi.org/10.1038/s41593-023-01311-w

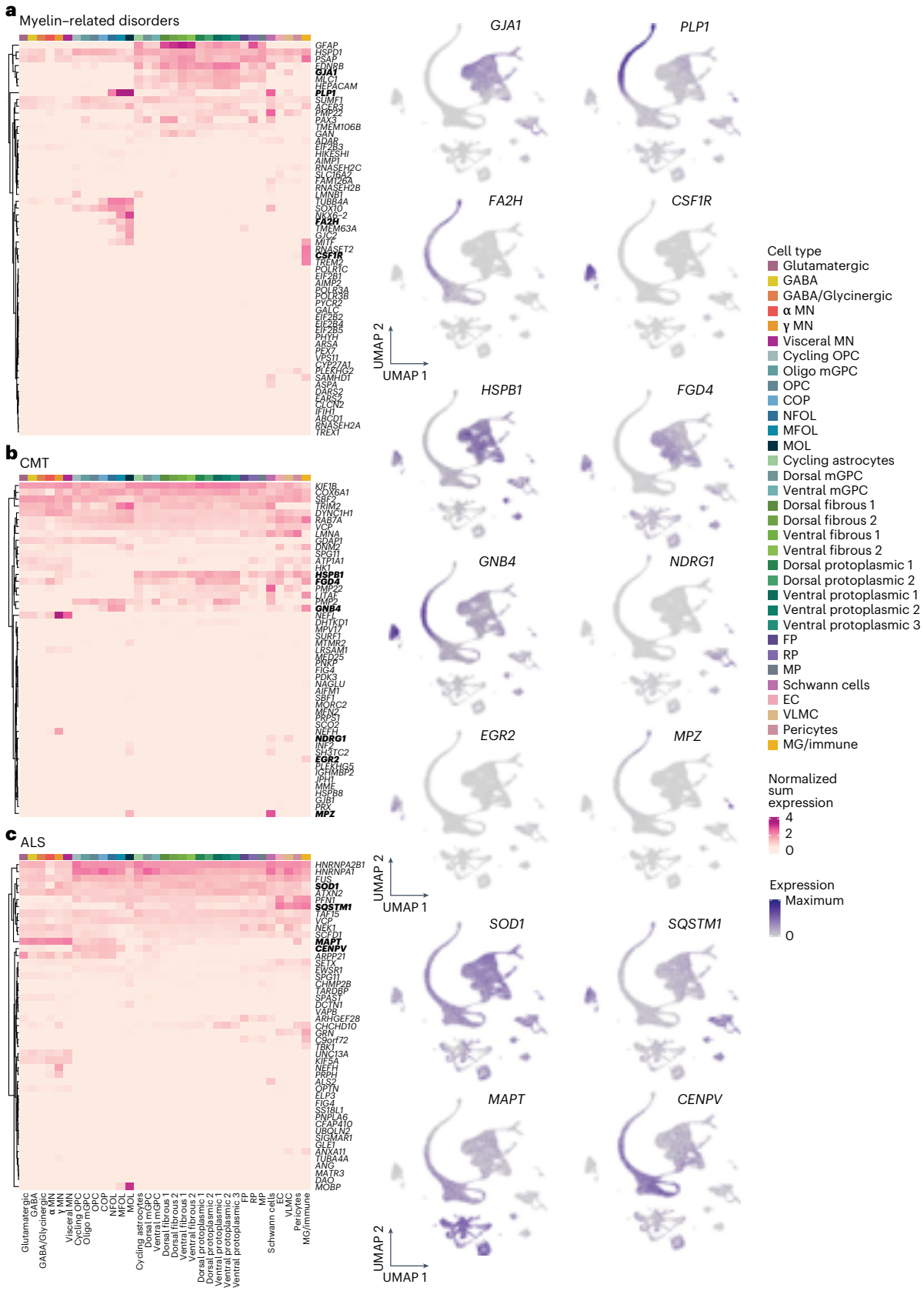
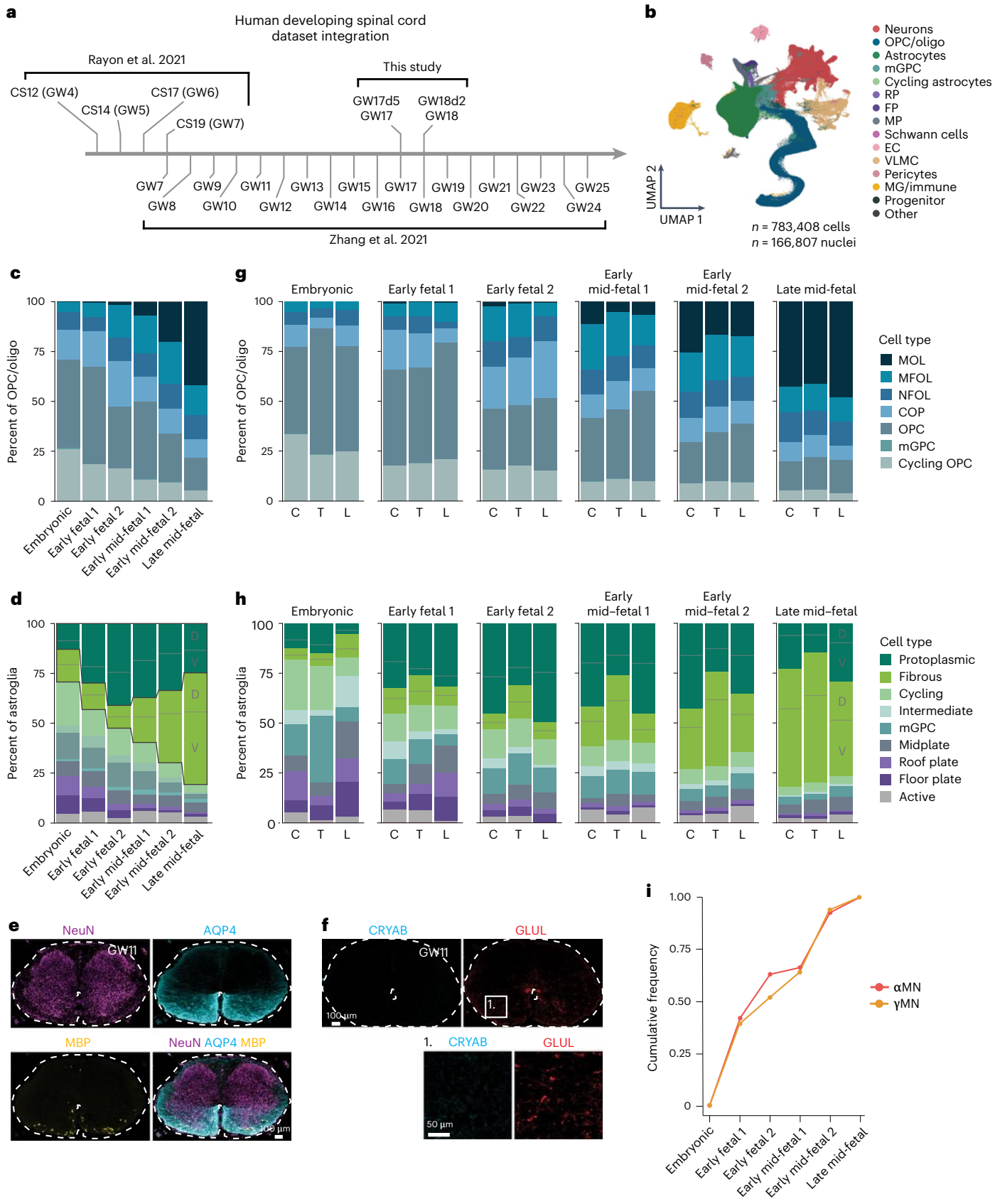


Fig. 6 | Mapping of disease gene sets on a human developing spinal cord atlas. **a**, A heat map (left) showing the sum of gene counts normalized to total counts per cell cluster for genes associated with myelin-related disorders and UMAP plots (right) showing gene expression of selected genes (bold). For references, see Methods. **b**, A heat map (left) showing the sum of gene counts normalized to total counts per cell cluster for genes associated with Charcot-Marie-Tooth (CMT) and UMAP plots (right) showing expression of selected genes (bold). For references, see Methods. **c**, A heat map (left) showing the sum of gene counts normalized to total counts per cell cluster for genes associated with Amyotrophic Lateral Sclerosis (ALS) and UMAP plots (right) showing expression of selected genes (bold). For references, see Methods.

https://doi.org/10.1038/s41593-023-01311-w



relative contribution of ages to these clusters and found that distinct α and γ signatures are present as early as GW8–10 (Fig. 6i, Extended Data Fig. 10m and Supplementary Table 25). MN diversification into extrafusal- and intrafusal-innervating αMNs and γMNs is only apparent at late stages of rodent development, with the earliest marker of γMNs detectable at E17.5 (ref. 59), so we were surprised to see this separation at early fetal stages in human. To determine whether the absence of early markers for αMNs and γMNs hindered identification of these

Fig. 7 | Integration of developing human spinal cord datasets. a, A schematic showing all the human developing spinal cord samples included in the dataset integration. GW, gestational week; d, day. **b**, UMAP representation of integrated datasets, colored by cell type. **c**, Bar plot showing the percent of cells within the OPC/oligo lineage over gestational time. **d**, A bar plot showing the percent of cells within the astroglia subcluster over gestational time. Dorsal (D) and ventral (V) subtypes are shown. **e**, Representative immunohistochemistry images showing the neuronal marker NeuN, the astrocyte marker AQP4 and the myelinating oligo marker MBP in GW11 spinal cord cryosections. Immunohistochemistry was replicated twice in one independent sample. **f**, Representative immunohistochemistry images showing the fibrous and protoplasmic astrocyte markers CRYAB and GLUL in GW11 spinal cord cryosections. Immunohistochemistry was replicated twice in one independent sample. **g**, Bar plots showing the percent of cells within the OPC/oligo lineage over gestational time and separated in cervical (C), thoracic (T) and lumbar (L) regions of the spinal cord. **h**, Bar plots showing the percent of cells within the astroglia subcluster over gestational time and separated in cervical (C), thoracic (T) and lumbar (L) regions. Dorsal (D) and ventral (V) subtypes are shown. **i**, A curve plot showing the cumulative frequency of alpha (α) and gamma (γ) MNs over time in the integrated dataset. Scale bars, 50 μm (insets in **f**) and 100 μm (**e** and **f**).

populations at earlier timepoints in mouse, we used α, γ and visceral MN markers identified in this study and performed label transfer onto a mouse dataset at earlier stages (E9.5–E13.5, ref. 60), but we did not see clear separations between these three MN groups at early developmental stages (Extended Data Fig. 10n–q and Supplementary Fig. 7b,c), and we did not see this separation when utilizing MNs from the embryonic dataset in Rayon et al.[6] in a similar way (Extended Data Fig. 10r–u). Altogether, these data suggest that α and γ MN diversification takes place at relatively early fetal stages of human development.

Discussion

Single-cell transcriptomics technologies have enabled the systematic profiling of cell diversity across the central nervous system at an unprecedented rate[61]. To generate a comprehensive census of cell types of the midgestation human spinal cord, we performed single-cell and single-nucleus RNA sequencing (RNA-seq). This allowed us to describe glial and neuronal diversity at this stage, and point at an early fetal diversification of α and γ MNs.

A common theme that emerged through our analysis was the existence of a transcriptional positional identity code governing cell diversity. Our data suggest that the acquisition of positional identity is a common fundamental principle guiding patterning in the spinal cord beyond neuronal specification. In the astroglia lineage, positional differences appeared to be linked to functional differences on the basis of gene expression; however, further work is needed to demonstrate this. Sampling at earlier timepoints could also determine the relationship between some of the cell types we described. For example, although analysis of the integrated dataset suggested that protoplasmic astrocytes acquired their identity before fibrous ones, how these astrocytes are generated during development or if they are derived from a common progenitor remains unknown. Similarly, the dual lineage potential of mGPCs is reminiscent of ependymal cells in the adult murine spinal cord that, upon injury, are activated to generate both oligos and astrocytes[45], yet whether mGPCs are generated from VZ/ependymal cells during development is unclear. Zhang et al.[58] explored a population of EGFR$^+$ glial progenitors that probably corresponds to our mGPC population. They found that EGFR-expressing cells acquire glial characteristics over time and transform through intermediate stages into astrocytes and oligos. In addition, they found that, similar to what we observe, a large proportion of EGFR$^+$ cells in GW17–23 samples concentrated within the dorsal horn area, and they speculate these may play a role in the positioning of dorsal neuron localization. Future studies should explore this.

The integration of published human developing spinal cord datasets spanning 22 weeks of gestation[6,58] allowed us to validate our findings and put them into a broader developmental context. In addition, this dataset represents a comprehensive resource for exploring the biology of the human spinal cord over time. A remaining challenge will be to comprehensively compare developmental landmarks and cellular diversity across species to identify unique features and uncover the molecular machinery underlying maturation at a different developmental pace. For example, Rayon et al.[6] compared equivalent staged mouse and human embryonic spinal cord and found several differences, including expression of unique markers in the human FP, a higher proportion of human pMN and p3 progenitors early on, and a slower rate of dorsal neurogenesis in humans. Our analysis suggests that diversification of MNs into α and γ takes place in early fetal development, while the earliest evidence of this diversification in mouse has been described at E17.5 (ref. 59). Evidence for early muscle innervation by MNs in human compared with rodents may explain, in part, this discrepancy[12].

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41593-023-01311-w.

References

1. Jessell, T. M. Neuronal specification in the spinal cord: inductive signals and transcriptional codes. *Nat. Rev. Genet.* **1**, 20–29 (2000).
2. Haim, L. B. & Rowitch, D. H. Functional diversity of astrocytes in neural circuit regulation. *Nat. Rev. Neurosci.* **18**, 31–41 (2017).
3. Monani, U. R. Spinal muscular atrophy: a deficiency in a ubiquitous protein; a motor neuron-specific disease. *Neuron* **48**, 885–896 (2005).
4. Swash, M., Leader, M., Brown, A. & Swettenham, K. W. Focal loss of anterior horn cells in the cervical cord in motor neuron disease. *Brain* **109**, 939–952 (1986).
5. Bergen, V., Lange, M., Peidli, S., Wolf, F. A. & Theis, F. J. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat. Biotechnol.* **38**, 1408–1414 (2020).
6. Rayon, T., Maizels, R. J., Barrington, C. & Briscoe, J. Single-cell transcriptome profiling of the human developing spinal cord reveals a conserved genetic programme with human-specific features. *Development* **148**, dev199711 (2021).
7. Sousa, A. M. M., Meyer, K. A., Santpere, G., Gulden, F. O. & Sestan, N. Evolution of the human nervous system function, structure, and development. *Cell* **170**, 226–247 (2017).
8. Gu, Z. et al. Control of species-dependent cortico-motoneuronal connections underlying manual dexterity. *Science* **357**, 400–404 (2017).
9. Oberheim, N. A. et al. Uniquely hominid features of adult human astrocytes. *J. Neurosci.* **29**, 3276–3287 (2009).
10. Somel, M. et al. Transcriptional neoteny in the human brain. *Proc. Natl Acad. Sci. USA* **106**, 5743–5748 (2009).
11. Ohmura, Y. & Kuniyoshi, Y. A translational model to determine rodent's age from human foetal age. *Sci. Rep.* **7**, 17248 (2017).
12. Tadros, M. A., Lim, R., Hughes, D. I., Brichta, A. M. & Callister, R. J. Electrical maturation of spinal neurons in the human fetus: comparison of ventral and dorsal horn. *J. Neurophysiol.* **114**, 2661–2671 (2015).

13. Bakken, T. E. et al. Single-nucleus and single-cell transcriptomes compared in matched cortical cell types. *PLoS ONE* **13**, e0209648 (2018).
14. Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902.e21 (2019).
15. La Manno, G. et al. RNA velocity of single cells. *Nature* **560**, 494–498 (2018).
16. Molofsky, A. V. et al. Astrocyte-encoded positional cues maintain sensorimotor circuit integrity. *Nature* **509**, 189–194 (2014).
17. Tsai, H.-H. et al. Regional astrocyte allocation regulates CNS synaptogenesis and repair. *Science* **337**, 358–362 (2012).
18. Hochstim, C., Deneen, B., Lukaszewicz, A., Zhou, Q. & Anderson, D. J. Identification of positionally distinct astrocyte subtypes whose identities are specified by a homeodomain code. *Cell* **133**, 510–522 (2008).
19. Khakh, B. S. & Deneen, B. The emerging nature of astrocyte diversity. *Annu. Rev. Neurosci.* **42**, 187–207 (2019).
20. Liddelow, S. A. et al. Neurotoxic reactive astrocytes are induced by activated microglia. *Nature* **541**, 481–487 (2017).
21. Hodge, R. D. et al. Conserved cell types with divergent features in human versus mouse cortex. *Nature* **573**, 61–68 (2019).
22. Allen, N. J. Astrocyte regulation of synaptic behavior. *Annu. Rev. Cell Dev. Biol.* **30**, 439–463 (2014).
23. Allen, N. J. & Barres, B. A. Glia—more than just brain glue. *Nature* **457**, 675–677 (2009).
24. Allen, N. J. & Eroglu, C. Cell biology of astrocyte–synapse interactions. *Neuron* **96**, 697–708 (2017).
25. Chung, W.-S. et al. Astrocytes mediate synapse elimination through MEGF10 and MERTK pathways. *Nature* **504**, 394–400 (2013).
26. Man, H.-Y. GluA2-lacking, calcium-permeable AMPA receptors—inducers of plasticity? *Curr. Opin. Neurobiol.* **21**, 291–298 (2011).
27. Droste, D. et al. Ca2+-permeable AMPA receptors in mouse olfactory bulb astrocytes. *Sci. Rep.* **7**, 44817 (2017).
28. Lavialle, M. et al. Structural plasticity of perisynaptic astrocyte processes involves ezrin and metabotropic glutamate receptors. *Proc. Natl Acad. Sci. USA* **108**, 12915–12919 (2011).
29. Srinivasan, J., Schachner, M. & Catterall, W. A. Interaction of voltage-gated sodium channels with the extracellular matrix molecules tenascin-C and tenascin-R. *Proc. Natl Acad. Sci. USA* **95**, 15753–15757 (1998).
30. Betters, E., Liu, Y., Kjaeldgaard, A., Sundström, E. & García-Castro, M. I. Analysis of early human neural crest development. *Dev. Biol.* **344**, 578–592 (2010).
31. Ghazale, H. et al. RNA profiling of the human and mouse spinal cord stem cell niches reveals an embryonic-like regionalization with MSX1+ roof-plate-derived cells. *Stem Cell Rep.* **12**, 1159–1177 (2019).
32. Cañizares, M. A. et al. Multiple steps characterise ventricular layer attrition to form the ependymal cell lining of the adult mouse spinal cord central canal. *J. Anat.* **236**, 334–350 (2020).
33. Sagner, A. & Briscoe, J. Establishing neuronal diversity in the spinal cord: a time and a place. *Development* **146**, dev182154 (2019).
34. Weidenheim, K. M., Epshteyn, I., Rashbaum, W. K. & Lyman, W. D. Patterns of glial development in the human foetal spinal cord during the late first and second trimester. *J. Neurocytol.* **23**, 343–353 (1994).
35. Pollen, A. A. et al. Molecular identity of human outer radial glia during cortical development. *Cell* **163**, 55–67 (2015).
36. Trevino, A. E. et al. Chromatin and gene-regulatory dynamics of the developing human cerebral cortex at single-cell resolution. *Cell* **184**, 5053–5069.e23 (2021).
37. Girskis, K. M. et al. Rewiring of human neurodevelopmental gene regulatory programs by human accelerated regions. *Neuron* **109**, 3239–3251.e7 (2021).
38. Shinozuka, T. & Takada, S. Morphological and functional changes of roof plate cells in spinal cord development. *J. Dev. Biol.* **9**, 30 (2021).
39. Lemke, G. Glial control of neuronal development. *Annu. Rev. Neurosci.* **24**, 87–105 (2001).
40. Chédotal, A. Roles of axon guidance molecules in neuronal wiring in the developing spinal cord. *Nat. Rev. Neurosci.* **20**, 380–396 (2019).
41. Masahira, N. et al. Olig2-positive progenitors in the embryonic spinal cord give rise not only to motoneurons and oligodendrocytes, but also to a subset of astrocytes and ependymal cells. *Dev. Biol.* **293**, 358–369 (2006).
42. Wang, H. et al. Region-specific distribution of Olig2-expressing astrocytes in adult mouse brain and spinal cord. *Mol. Brain* **14**, 36 (2021).
43. Vue, T. Y., Kim, E. J., Parras, C. M., Guillemot, F. & Johnson, J. E. Ascl1 controls the number and distribution of astrocytes and oligodendrocytes in the gray matter and white matter of the spinal cord. *Development* **141**, 3721–3731 (2014).
44. Marshall, C. A. G., Novitch, B. G. & Goldman, J. E. Olig2 directs astrocyte and oligodendrocyte formation in postnatal subventricular zone cells. *J. Neurosci.* **25**, 7289–7298 (2005).
45. Llorens-Bobadilla, E. et al. A latent lineage potential in resident neural stem cells enables spinal cord repair. *Science* **370**, eabb8795 (2020).
46. Russ, D. E. et al. A harmonized atlas of mouse spinal cord cell types and their spatial organization. *Nat. Commun.* **12**, 5722 (2021).
47. Rosenberg, A. B. et al. Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science* **360**, 176–182 (2018).
48. Sathyamurthy, A. et al. Massively parallel single nucleus transcriptional profiling defines spinal cord neurons and their activity during behavior. *Cell Rep.* **22**, 2216–2225 (2018).
49. Osseward, P. J. et al. Conserved genetic signatures parcellate cardinal spinal neuron classes into local and projection subsets. *Science* **372**, 385–393 (2021).
50. Zeisel, A. et al. Molecular architecture of the mouse nervous system. *Cell* **174**, 999–1014.e22 (2018).
51. Häring, M. et al. Neuronal atlas of the dorsal horn defines its architecture and links sensory input to transcriptional cell types. *Nat. Neurosci.* **21**, 869–880 (2018).
52. Fang, R. et al. Conservation and divergence of cortical cell organization in human and mouse revealed by MERFISH. *Science* **377**, 56–62 (2022).
53. Blum, J. A. et al. Single-cell transcriptomic analysis of the adult mouse spinal cord reveals molecular diversity of autonomic and skeletal motor neurons. *Nat. Neurosci.* **24**, 572–583 (2021).
54. Price, S. R., Garcia, N. V. D. M., Ranscht, B. & Jessell, T. M. Regulation of motor neuron pool sorting by differential expression of type ii cadherins. *Cell* **109**, 205–216 (2002).
55. Dewitz, C., Duan, X. & Zampieri, N. Organization of motor pools depends on the combined function of N-cadherin and type II cadherins. *Development* **146**, dev180422 (2019).
56. van der Knaap, M. S. & Bugiani, M. Leukodystrophies: a proposed classification system based on pathological changes and pathogenetic mechanisms. *Acta Neuropathol.* **134**, 351–382 (2017).
57. Ilieva, H., Polymenidou, M. & Cleveland, D. W. Non-cell autonomous toxicity in neurodegenerative disorders: ALS and beyond. *J. Cell Biol.* **187**, 761–772 (2009).

58. Zhang, Q. et al. Single-cell analysis reveals dynamic changes of neural cells in developing human spinal cord. *EMBO Rep.* **22**, e52728 (2021).
59. Ashrafi, S. et al. Wnt7A identifies embryonic γ-motor neurons and reveals early postnatal dependence of γ-motor neurons on a muscle spindle-derived signal. *J. Neurosci.* **32**, 8725–8731 (2012).
60. Delile, J. et al. Single cell transcriptomics reveals spatial and temporal dynamics of gene expression in the developing mouse spinal cord. *Development* **146**, dev173807 (2019).
61. Winnubst, J. & Arber, S. A census of cell types in the brain's motor cortex. *Nature* **598**, 33–34 (2021).

Methods

Human tissue

De-identified spinal cord samples were obtained at Stanford University School of Medicine from elective pregnancy terminations under a protocol approved by the Research Compliance Office at Stanford University. Samples were delivered on ice and processed for single-cell analyses or immunocytochemistry within 3 h of the procedure.

Sample collection and single-cell data generation

Single-cell dissociation and Thy1 immunopanning was performed as previously described[36,62,63]. Briefly, spinal cords were dissected out of the vertebral column, chopped and incubated in 30 U ml^{-1} of papain enzyme solution (Worthington, LS03126) for 45 min at 37 °C. After digestion, samples were washed with a protease inhibitor solution and gently triturated using progressively smaller pipette tips to achieve a single-cell suspension. Cells were resuspended in 0.02% bovine serum albumin/phosphate-buffered saline (BSA/PBS) and passed through a 70 µm flowmi filter and either continued to single-cell sample preparation or Thy1 immunopanning to enrich for neurons. For immunopanning, the single-cell suspension was added to a plastic Petri dish precoated with an anti-Thy1 antibody (CD90, BD Biosciences, 550402) and incubated for 10–30 min at room temperature (RT). Bound cells were incubated in an Accutase solution (Innovative Cell Technologies, AT104) at 37 °C for 3–5 min, and then gently washed off, spun down at 200*g* for 5 min and resuspended in cold 0.02% BSA/PBS.

Nuclei isolation was performed as described in ref. 64 with some modifications. Briefly, dissected spinal cords were disrupted using the detergent–mechanical cell lysis method using a 2 ml glass tissue grinder (Sigma-Aldrich/KIMBLE, D8938). Crude nuclei were then filtered using a 40 µm filter and centrifuged at 320*g* for 10 min at 4 °C before performing a sucrose density gradient to separate them from cellular debris. After a centrifugation step (320*g*, 20 min at 4 °C), samples were resuspended in 0.04% BSA/PBS supplemented with 0.2 U $µl^{-1}$ RNAse inhibitor (Ambion 40U $µl^{-1}$, AM2682) and passed through a 40 µm flowmi filter. NeuN sorting was performed as described in ref. 65. Briefly, the cell suspension was mixed with 1.2 µl of mouse anti-NeuN antibody (Millipore, MAB377) and 1 µl of anti-mouse Alexa Fluor 488 (Thermo Fisher Scientific, A21202) in PBS with 0.5% BSA and 10% normal donkey serum for 45 min on ice. The sorting was performed on a BD fluorescence-activated cell sorting (FACS) Aria II Cell Sorter at the Stanford Shared FACS Facility. A 100 µm nozzle and Purity mode were used during the sorting. Nuclei suspension stained with secondary antibody only was used as a control for setting up the gate. Sorted nuclei were collected in 0.04% BSA/PBS supplemented with 0.2 U $µl^{-1}$ RNAse inhibitor.

Cell and nucleus suspensions were then loaded onto a Chromium Single cell 3′ chip, version 3 (10x Genomics) and processed according to instructions with a target of 10,000 cells, 8–10,000 nuclei and 4,000 NeuN-sorted nuclei.

Analysis and processing of single-cell RNA and single-nucleus RNA-sequencing data

Single-cell and single-nucleus libraries were sequenced by Admera Health on a Novaseq S4 (Illumina) using 150 ×2 chemistry. Fastq files were obtained from Admera Health and subsequently aligned to the human reference genome GRCh38 (2020-A) using Cell Ranger (version 5.0.0, 10x Genomics) on the Sherlock Stanford Computing Cluster. We used the 'include introns' option for both single-cell and single-nucleus samples and all other parameters were kept as default. Spliced and unspliced counts were computed using velocyto (version 0.17.17 (ref. 15)) from the BAM files produced by Cell Ranger with default parameters and the same reference used for alignment.

Analysis of the counts-by-gene matrix created by Cell Ranger was performed using 'Seurat' (version 3.9.9.9024, beta of version 4 (ref. 14)) in R (version 3.6.1). QC and filtering were performed on single cells and single nuclei separately. These were first filtered to remove cells or nuclei with fewer than 1,000 features or more than 30% mitochondrial reads. Subsequently, through repeated rounds of clustering and QC, we removed doublets and low-quality cells. For each round of clustering and QC, known cell-type-specific marker genes were used to annotate cells. We used *SOX10*, *OLIG2*, *PDGFRA*, *MOG* and *MBP* to identify cells in the OPC/oligo lineage; *AQP4*, *GFAP*, *FGFR3* and *SOX9* to identify astroglia; *STMN2* and *MYT1L* to identify neurons; *AIF1* for microglia; *TOP2A* for cycling cells; *PECAM1* for ECs; *NGFR* for Schwann cells; and *COL1A1* for VLMC. Clusters containing a majority of cells that expressed well-known cell-type-specific markers for two or more cell types (for example, clusters expressing both *MYT1L* and *SOX10* or *AQP4*, *FGFR3* and *AIF1*) were assumed to be doublets and removed. We also removed clusters that were low quality, which was determined by a majority of cells with low nCount and high mitochondrial percentage. From the 138,159 single cells identified by Cell Ranger, 25,605 were identified as doublets or low-quality cells and removed. Similarly, 9,247 nuclei were removed out of the 44,131 initially identified by Cell Ranger. We note that NeuN-sorted nuclei had a low 'fraction of reads in cells' on the basis of Cell Ranger, suggesting possible ambient RNA contamination.

Remaining single cells were then normalized, scaled and integrated to correct for batch effects using the Integrate pipeline from Seurat, with 3,000 integration features used across the two batches of cells we collected. Single nuclei were also normalized, scaled and integrated with 2,000 integration features. Principal component analysis was performed with the first 75 and first 50 principal components selected for single cells and single nuclei, respectively. These principal components were then used in the 'FindNeighbors' and 'FindClusters' functions to determine cell groupings and to generate a two-dimensional UMAP projection via 'FindUMAP' for each dataset. These functions were run with the parameters recommended in Seurat's tutorial for large datasets ('FindNeighbors' nn.eps of 0.5) and were parallelized using four threads using the R package 'future.'

After QC and filtering we then integrated the two datasets together to generate our final UMAP. To aid the integration, we removed mitochondrial and ribosomal genes from the count matrices before integration. We used 2,000 features for canonical-correlation analysis (CCA) integration and 60 dimensions along with parameters in the pipeline described above for 'FindNeighbors'. This produced a total of 53 clusters, which were then combined and manually annotated on the basis of the cell-type-specific markers described above. The astrocyte and oligo subclusters were generated by rerunning 'FindVariableFeatures', 'ScaleData', 'RunPCA', 'FindNeighbors', 'FindClusters' and 'RunUMAP' on the integrated assay of the combined dataset. The mGPC cluster was split between the astroglia and OPC/oligo subclusters: Seurat cluster 7 was included with the astroglia and Seurat cluster 46 with the OPC/oligo. For neurons, we performed a further round of doublet removal and then reintegrated the data due to the prevalence of Thy1 immunopanned cells and NeuN-sorted nuclei, integrating across collection type instead of sequencing run. Within the astroglia and neuron subclusters, we further subclustered cycling astroglia (10 dimensions and 1,000 variable features), VZ cells (15 dimensions and 1,000 variable features), and MNs (8 dimensions and 750 features) respectively. In the MN subcluster we removed a cluster of cells ($n = 15$) that expressed DRG markers. Annotation of subgroups or cell types within subclusters was determined by gene expression of known markers or label transfer with published datasets. In some cases, several Seurat clusters were combined to form a subgroup, and the identity of these are presented in the supplementary tables.

For heat maps showing cluster expression, we pseudobulked each cluster by summing up gene expression before normalization via Seurat's 'NormalizeData'. Cell cycle scores were generated in Seurat using the 'CellCycleScoring' function and the default 's.genes' and 'g2.genes' objects loaded in Seurat. For MA plots, differential expression was computed with 'FindMarkers' with the option 'logfc.threshold' set to 0.

Myelin-related disorder genes were selected from refs. 56,66–70. CMT-associated genes were gathered from the Human Phenotype Ontology website (https://hpo.jax.org/app/). ALS-associated genes were selected from the following human exome sequencing studies: refs. 71–77.

Differential expression and GO analysis

Differential expression was performed in Seurat with either 'FindAllMarkers' or 'FindMarkers' using a Wilcoxon sum rank test and default parameters, except that 'min.pct' was changed to 0.25 unless otherwise specified. All tests were done using the Wilcoxon rank sum test. Finally, top markers were selected from those that had adjusted *P* values <0.01, then ranked by their log fold change (FC).

GO analysis for differential expression between single cells and single nuclei was completed using 'fgsea' (version 1.19.2) with terms supplied via the GSEA mSigDB, specifically those related to Biological Processes. To find terms that were enriched, an overexpression analysis was carried out with a hypergeometric test using the 'fora' function. The gene universe was taken to be all genes present in the RNA counts matrix, and the gene list was taken to be at most the top 250 genes by logFC with adjusted *P* values <0.01 based on results from differential expression. Calculated adjusted *P* values were then used in subsequent analyses.

Label transfer with reference datasets

Data for label transfer analysis performed in this study were obtained from the Gene Expression Omnibus (GEO) using the following accession numbers: GSE110823 (ref. 47), GSE103892 (ref. 48), GSE103840 (ref. 51), GSE171607 (ref. 49), GSE161621 (ref. 53) and GSE171892 (ref. 6). Data for Zeisel et al.[50] were obtained from mousebrain.org. Data for Delile et al.[60] were obtained from E-MTAB-7320. RNA count data were obtained from the authors from Russ et al.[46]. The provided annotations and counts by gene matrices were then processed in the same manner as our data before comparison. For mouse datasets, genes were converted from mouse MGI or ENSEMBL to HGNC symbols via a list of mouse to human homologs[78]. In the case of the Rayon et al.[6], dataset, integration, normalization and clustering were performed on the full dataset (2,000 variable features, 2,000 integration features, 40 dimensions, 0.3 resolution, 30% mitochondrial gene cutoff and 1,000 nCount cutoff). MNs for this dataset were then identified on the basis of the expression of *ISL1*, *MNX1* and *CHAT* and subclustered (1,000 variable features, 10 dimensions and 0.3 resolution). Label transfer was then carried out with all reference datasets using a CCA projection of the dataset on the first 25 dimensions using variable features from our dataset from the integrated assay between the neuron or MN subcluster of our dataset and the RNA assay of the reference dataset.

Cell–cell interaction analysis

We used NATMI to investigate ligand–receptor interactions between cell types. Counts were taken from the RNA assay of an object, then normalized using 'NormalizeData' with default parameters and transformed into counts per million/transcripts per million (CPM/TPM) as demonstrated by the NATMI instructions. The command line tool was then used with Python (3.6.1) to both extract and visualize the edges of the ligand–receptor interaction network. We used 32 threads, a specificity cutoff of 0.05 and a detection cutoff of 0.1. This provided a list of edges with information on their expression, specificity and detection. For visualization, we selected edges that were between astrocyte and OPC/oligo subtypes but not between either astrocytes and astrocytes or OPC/oligo and OPC/oligo.

RNA velocity and pseudotime

We computed RNA velocity using custom R scripts interfacing with the 'scVelo' toolkit (version 0.2.3) (ref. 5) via the 'reticulate' R–Python interface. For this, we exported the Velocyto-derived spliced and unspliced counts along with Seurat-derived PC and UMAP representations of single cells as 'AnnData' objects. We filtered the dataset using the scVelo function 'pp.filter_and_normalize' (parameters: min_shared_counts = 10 and n_top_genes = 2,000) and computed moments using 'pp.moments' (n_pcs = 30, n_neighbors = 30). We then used 'tl.velocity' with mode = 'dynamical' to compute cell velocities and 'tt.velocity_graph' to compute a velocity graph. Potential root and end point cells for the trajectory were computed using 'tt.terminal_states'. 'tt.velocity_pseudotime' was applied to compute cell pseudotime scores. We then reimported the scVelo-derived gene and cell annotations (gene velocities and other scVelo-inferred model parameters, as well as cell pseudotime, velocity projections, root and end point probabilities) into the metadata of the R-based Seurat objects. Cell-type-specific velocity genes were obtained using scVelo's 'rank_velocity_genes' function.

To facilitate aggregate analysis along pseudotime, we obtained pseudobulk samples by sorting cells on the basis of their pseudotime scores and merging bins of 100 cells. For merging gene expression levels, counts for all cells assigned to a pseudobulk sample were summed and the data were renormalized using counts per million normalization.

Dataset integration

Data for integration[6,58] were obtained from GEO. Filtering of cells/nuclei was performed using the same metrics as described above (30% MT percent and 1,000 minimum nFeature). Due to the large size of the combined dataset, we first randomly selected 300,000 cells from the Zhang et al. dataset to combine with nuclei from the Zhang dataset, our own dataset and the Rayon et al. dataset. We then ran both the seurat functions NormalizeData and FindVariableFeatures to find the 3,000 most variable features. Next, we created a matrix using all cells but only the variable features, and used this for principal component analysis and dimensionality reduction, and for visualizing and clustering the data using UMAP. We then used the top markers per cluster (of the 3,000 variable features) to annotate all cells. We note that one cluster (cluster 43) in this dataset probably represents Schwann cells, but was labeled as 'Other' since it appeared to have a mixed signature, and potentially included a large number of doublets. Finally, we separately subclustered all the cell types (astroglia, OPC/oligo, VZ cells and neurons), which allowed us to use all features in a reduced number of cells for further analysis. We performed one round of doublet removal in each subcluster. In each case, as described above, clusters containing a majority of cells that expressed well-known cell-type-specific markers for two or more cell types (for example, clusters expressing both *MYT1L* and *SOX10* or *AQP4*, *FGFR3* and *AIF1*) were assumed to be doublets and removed. For each subcluster, we then used a nearest neighbor method (get.knnx from the FNN package) to annotate the cells using our annotations as a reference. As part of the QC for each subcluster, we also generated module scores for each sample using the GO Oxidative Stress gene list. To compute this score we used the AddModuleScore function in Seurat with the default parameters, where the score is normalized to 100 random genes. We also used the 'AddModuleScore' function in Seurat to calculate module scores for alpha and gamma MN clusters in this dataset using the genes shown in Fig. 5k or using a maximum of 50 genes with a *P* value <0.01 and $\log_2$FC >1 or <−1 for Blum et al. and Russ et al.

We note that GW22–24 samples from Zhang et al., probably due to difficulties with dissociation of myelinated tissue, had only a small percent of neurons. Therefore, the neuronal subcluster consists mainly of samples from GW4–21.

Age bins for analysis of samples over time corresponded to the following ages: embryonic (GW4–9), early fetal 1 (GW10–11), early fetal 2 (GW12–14), early mid-fetal 1 (GW15–17), early mid-fetal 2 (GW18–19) and late mid-fetal (GW20–25).

Immunohistochemistry

Immunohistochemistry was performed as previously described[36,62,63]. In brief, GW18 and GW19 spinal cords were dissected out of the vertebral column and fixed for 2–3 h at 4 °C with 4% paraformaldehyde (PFA,

Electron Microscopy Sciences), washed with PBS and transferred to a 30% sucrose solution for 48–72 h. Once the samples had sunk in this solution, they were embedded in an optimal cutting temperature (OCT)/30% sucrose solution (1:1) and snap frozen in dry ice. Cryosections were obtained using a Leica cryostat set at 20–30 µm and mounted on Superfrost Plus Micro slides (VWR, 48311-703). On the day of staining, sections were then blocked and permeabilized for 1 h at RT in blocking solution containing 10% normal donkey serum and 0.3% Triton-X in PBS and then incubated with primary antibodies at 4 °C overnight. The primary antibodies used are presented in Supplementary Table 26. Primary antibodies were washed using PBS, and sections were incubated with Alexa Fluor secondary antibodies (1:1,000, Life Technologies) for 1 h at RT. Hoechst 33258 (Thermo Fisher Scientific, H3569) was used to visualize nuclei. Sections were mounted with glass coverslips using Aqua Poly-mount (Polysciences, 18606-5). Images were taken using a Leica SP8 confocal microscope or a Keyance fluorescence microscope and processed using ImageJ (Fiji). All immunohistochemistry validations were performed in GW18 and GW19 samples.

Quantification of immunofluorescence images were performed using ImageJ (Fiji). Regions of interest (ROIs) were selected from each hemicord by defining a 300 × 2,100 µm region that extended from the ventral to dorsal white matter. ROIs were processed by using the subtract background plugin with a rolling ball radius of 50. A Gaussian blur with sigma radius of 10 was applied to each ROI. Following processing, plot profiles were generated within ImageJ to quantify average pixel intensity of markers along the ventral–dorsal axis.

In situ hybridization (RNAScope)

GW18 ($n = 1$) and GW19 ($n = 1$) samples were processed as described above. Spinal cords from P0 ($n = 4$) and P25 ($n = 3$) mice were isolated by hydraulic extrusion. Tissues were then fixed in 4% PFA either for 2 h or overnight and cryopreserved in 30% sucrose. Lumbar spinal cords were embedded in optimal cutting temperature compound (OCT) and 20 µm transverse sections were cut on a Leica CM3050 S Cryostat. For RNAScope, frozen cryosections were hydrated in PBS and pretreatment was performed as follows: slides were baked at 60 °C for 45 min, post-fixed in 4% PFA at RT for 1 h and dehydrated in 50%, 70%, 100% and 100% ethanol (5 min each). Sections were then incubated in hydrogen peroxide for 10 min, and antigen retrieval was performed in a vegetable steamer for 4 min. Finally, slides were baked a final time for 45 min. Sections were processed with the RNAScope Multiplex Fluorescent V2 Kit (ACD Biosciences) following the manufacturer's guidelines. The following probes were used: Hs-CHAT-C2 (450671-C2), Hs-CHODL (506601), Hs-NDNF-C3 (495251-C3), Mm-Chat-C2 (408731-C2), Mm-Chodl (450211) and Mm-Ndnd-C3 (447471-C3). All imaging was done on a Leica SP8 confocal; maximum projections were generated using LASX software and further processed using ImageJ (Fiji).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Data generated for the analyses presented in this work are available under GEO accession number GSE188516. A website associated with the manuscript, including an interactive data browser, is available at https://devspinalcord.su.domains/.

Code availability

All code used for data analysis is available as part of the packages mentioned above.

References

62. Sloan, S. A., Andersen, J., Paşca, A. M., Birey, F. & Paşca, S. P. Generation and assembly of human brain region-specific three-dimensional cultures. *Nat. Protoc.* **13**, 2062–2085 (2018).
63. Trevino, A. E. et al. Chromatin accessibility dynamics in a model of human forebrain development. *Science* **367**, eaay1645 (2020).
64. Matson, K. J. E. et al. Isolation of adult spinal cord nuclei for massively parallel single-nucleus RNA sequencing. *J. Vis. Exp.* https://doi.org/10.3791/58413 (2018).
65. Matevossian, A. & Akbarian, S. Neuronal nuclei isolation from human postmortem brain tissue. *J. Vis. Exp.* https://doi.org/10.3791/914 (2008).
66. Charzewska, A. et al. Hypomyelinating leukodystrophies—a molecular insight into the white matter pathology. *Clin. Genet* **90**, 293–304 (2016).
67. Abrams, C. K. & Scherer, S. S. Gap junctions in inherited human disorders of the central nervous system. *Biochim. Biophys. Acta* **1818**, 2030–2047 (2012).
68. Brenner, M. et al. Mutations in GFAP, encoding glial fibrillary acidic protein, are associated with Alexander disease. *Nat. Genet.* **27**, 117–120 (2001).
69. Zardadi, S. et al. Four mutations in MITF, SOX10 and PAX3 genes were identified as genetic causes of waardenburg syndrome in four unrelated Iranian patients: case report. *BMC Pediatr.* **21**, 70 (2021).
70. Depienne, C. et al. Brain white matter oedema due to ClC-2 chloride channel deficiency: an observational analytical study. *Lancet Neurol.* **12**, 659–668 (2013).
71. Kenna, K. P. et al. NEK1 variants confer susceptibility to amyotrophic lateral sclerosis. *Nat. Genet.* **48**, 1037–1042 (2016).
72. Cirulli, E. T. et al. Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways. *Science* **347**, 1436–1441 (2015).
73. Farhan, S. M. K. et al. Exome sequencing in amyotrophic lateral sclerosis implicates a novel gene, DNAJC7, encoding a heat-shock protein. *Nat. Neurosci.* **22**, 1966–1974 (2019).
74. Smith, B. N. et al. Exome-wide rare variant analysis identifies TUBA4A mutations associated with familial ALS. *Neuron* **84**, 324–331 (2014).
75. Smith, B. N. et al. Mutations in the vesicular trafficking protein annexin A11 are associated with amyotrophic lateral sclerosis. *Sci. Transl. Med.* **9**, eaad9157 (2017).
76. Nicolas, A. et al. Genome-wide analyses identify KIF5A as a novel ALS gene. *Neuron* **97**, 1268–1283.e6 (2018).
77. van Rheenen, W. et al. Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis. *Nat. Genet.* **48**, 1043–1048 (2016).
78. Skene, N. G. et al. Genetic identification of brain cell types underlying schizophrenia. *Nat. Genet.* **50**, 825–833 (2018).

Acknowledgements

We thank members of S. P. Paşca, A. M. Paşca and W. G. Greenleaf laboratories for support, discussion and advice, especially S. Kanton, Y. Miura, L. Li, A. Trevino, K. W. Kelley and F. Birey. This work was supported by the S. Coates and V. Coates Foundation (S.P.P.), the Stanford Brain Organogenesis Program and the Big Idea Grant in the Stanford Wu Tsai Neurosciences Institute (S.P.P.), Bio-X (S.P.P.), the Kwan Fund (S.P.P), the Senkut Research Funds (S.P.P.), the Chan Zuckerberg Initiative Ben Barres Investigator Award (S.P.P), the Stanford University Department of Neurosurgery (J.A.K.) and the Stanford Wu Tsai Neurosciences Institute (J.A.K.). S.P.P. is a New York Stem Cell Foundation Robertson Stem Cell investigator and a Chan Zuckerberg Biohub investigator. W.J.G. is a Chan Zuckerberg Biohub investigator and acknowledges grants 2017-174468 and 2018-182817 from the Chan Zuckerberg Initiative. Fellowship support was provided by the Idun Berry Postdoctoral Fellowship (J.A.). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

J.A., N.T. and S.P.P. conceived the project and designed experiments. N.T. performed data analysis with guidance from F.M. and W.J.G. J.A. guided the biological interpretation of the analysis and performed immunohistochemistry validations. F.M. performed the RNA velocity analysis. J.L.S and J.A.K. performed and interpreted RNAScope validations. A.M.P., N.D.A., L.L. and J.A. processed samples. X.C. performed the nuclei NeuN-sort experiment. M.M.O performed immunohistochemistry validations and quantification. S.-J.Y. performed Thy1 immunopanning. J.A. performed Chromium 10x capture. J.A., N.T. and S.P.P. wrote the manuscript with input from all authors. J.A. and S.P.P. supervised the work.

Competing interests

W.J.G. was a consultant for 10x Genomics. The remaining authors declare no competing interests.

Additional information

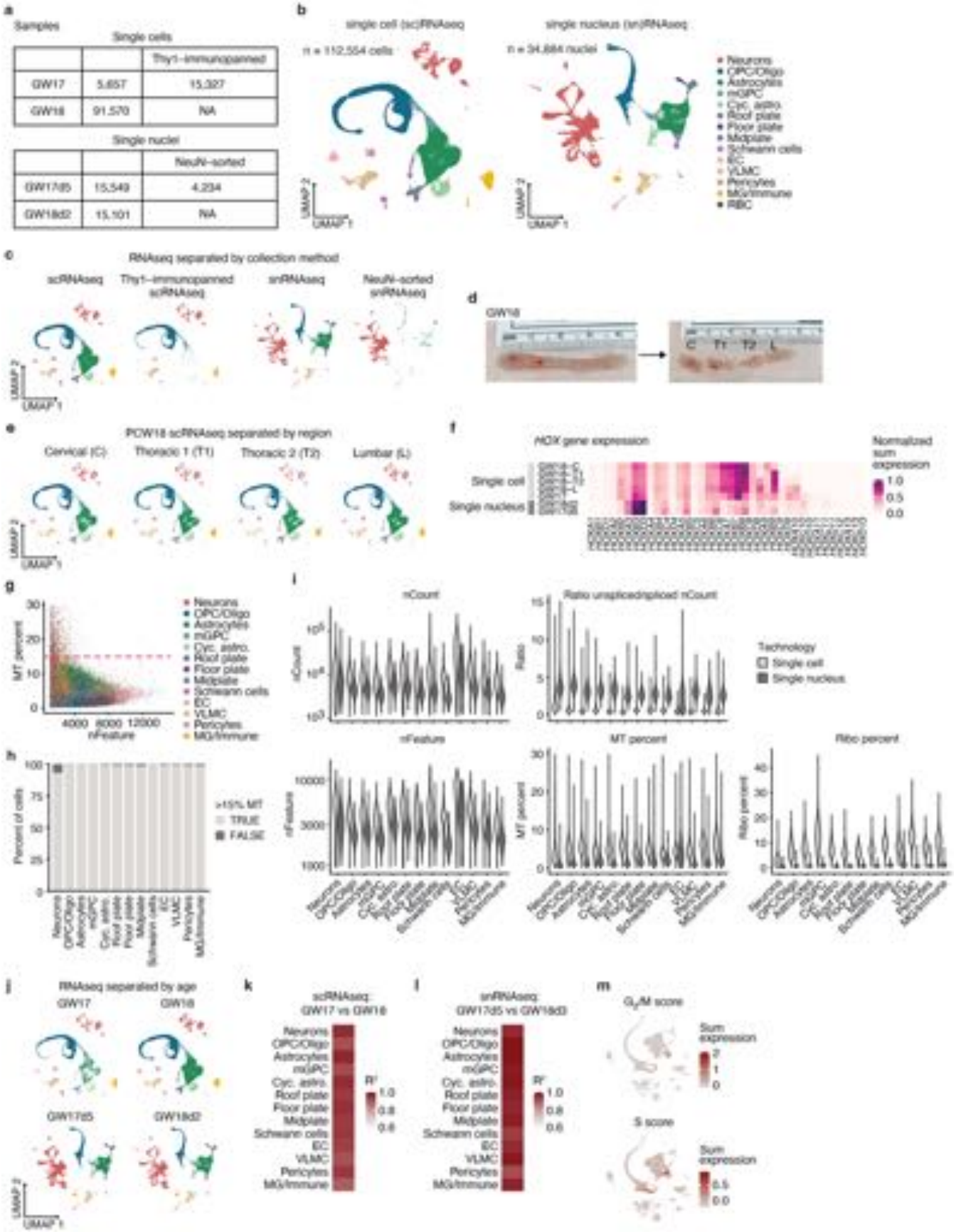
Extended data is available for this paper at https://doi.org/10.1038/s41593-023-01311-w.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41593-023-01311-w.

Correspondence and requests for materials should be addressed to Sergiu P. Pașca.

Peer review information *Nature Neuroscience* thanks the anonymous reviewers for their contribution to the peer review of this work.

https://doi.org/10.1038/s41593-023-01311-w

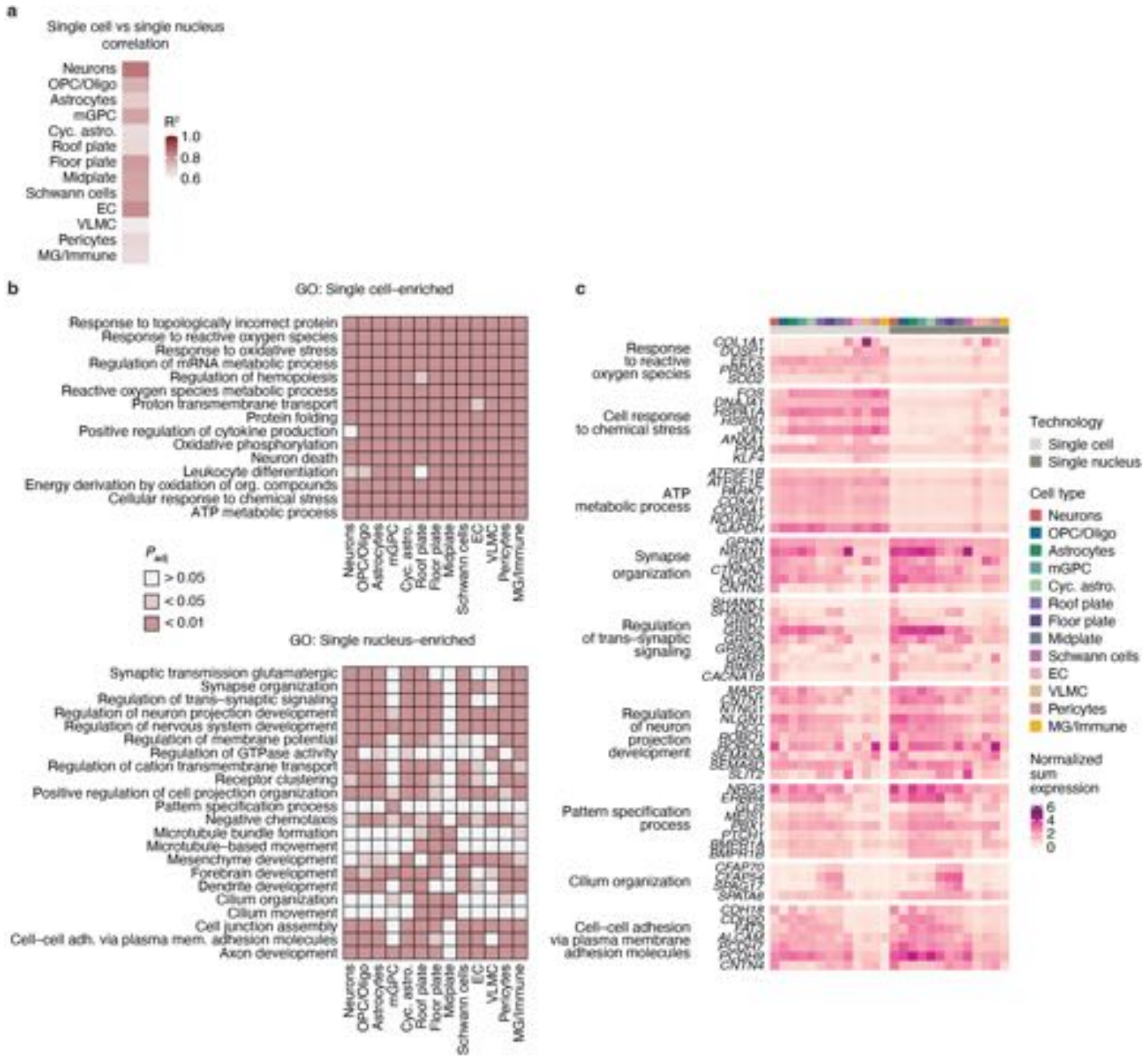


Extended Data Fig. 1 | See next page for caption.

https://doi.org/10.1038/s41593-023-01311-w

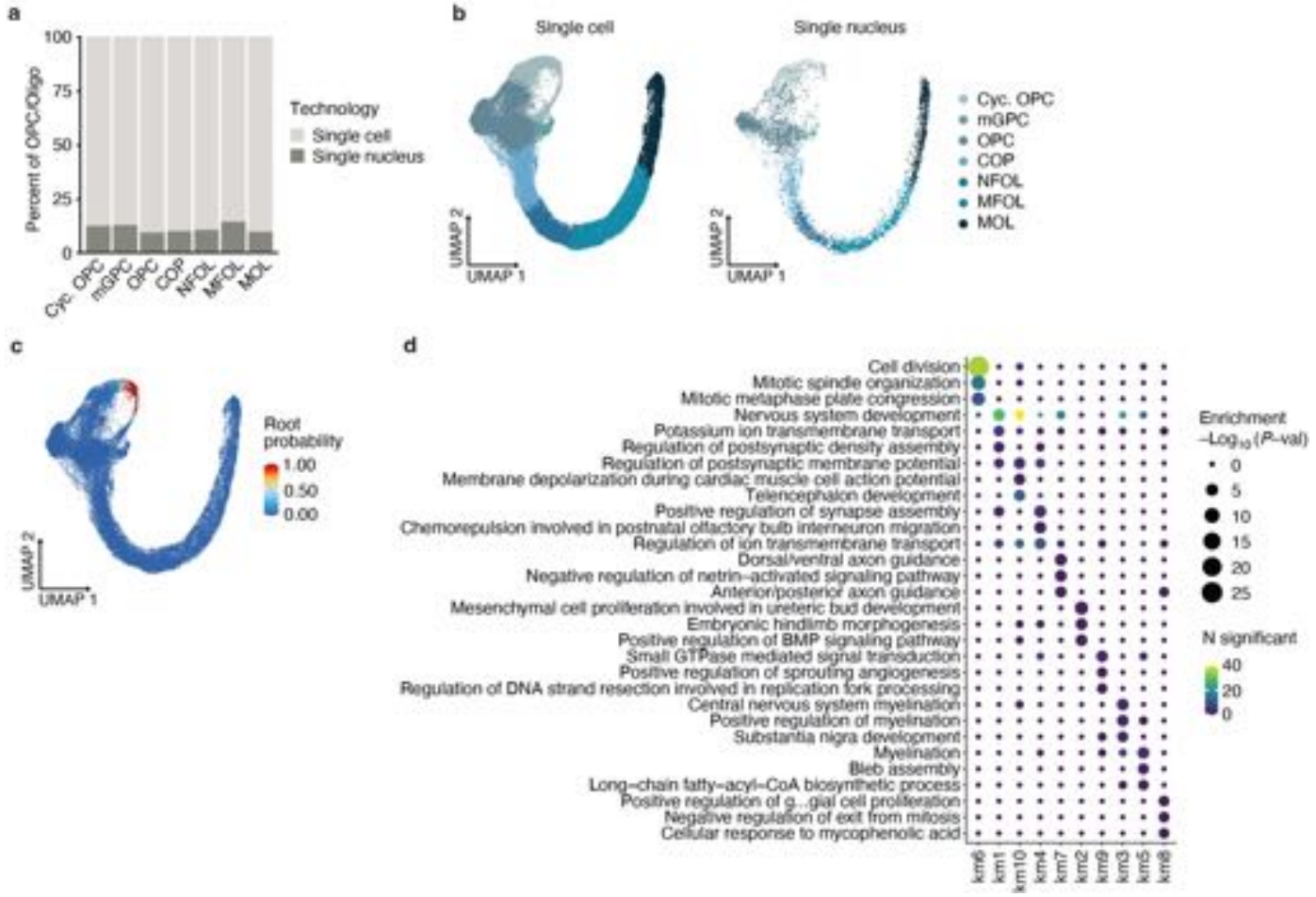
Extended Data Fig. 1 | Sample overview and quality control metrics. a. Table showing the number of single cells and single nuclei per sample and collection method. Numbers correspond to total numbers of cells or nuclei recovered after quality control (QC) and filtering. **b**. UMAPs showing single cells (left) and single nuclei (right) before integration, colored by cell type. **c**. Single cell and single nucleus UMAPs split by collection method. **d**. Picture of GW18 spinal cord sample before and after dividing it into rostro-caudal regions: cervical (C), thoracic (T1, T2) and lumbar (L). **e**. Single cell UMAPs showing GW18 cells split by region. **f**. Heatmap showing *HOX* gene expression as the sum of gene counts normalized to total sample counts in all samples. **g**. Scatter plot showing mitochondrial (MT) percent versus nFeature with points colored by cell type. The red dash line shows a less conservative hypothetical cutoff. **h**. Bar plot showing the proportion of cells/nuclei with MT percent above or below the hypothetical 15% cutoff per cell type. **i**. Violin plots showing total counts (nCount), ratio of spliced to unspliced counts, total number of features (nFeature), percent of mitochondrial (MT) counts, and percent of ribosomal (ribo) counts separated by cell type in both single cell and nucleus samples. **j**. Single cell and single nucleus UMAPs split by sample age. **k**. Heatmap showing the correlation of the normalized average gene expression between sample ages in single cell samples separated by cell type represented as R^2 values. **l**. Heatmap showing the correlation of the normalized average gene expression between ages in single nuclei samples separated by cell type represented as R^2 values. **m**. G_2/M and S cell cycle scores in the integrated single cell and single nuclei UMAP.

https://doi.org/10.1038/s41593-023-01311-w



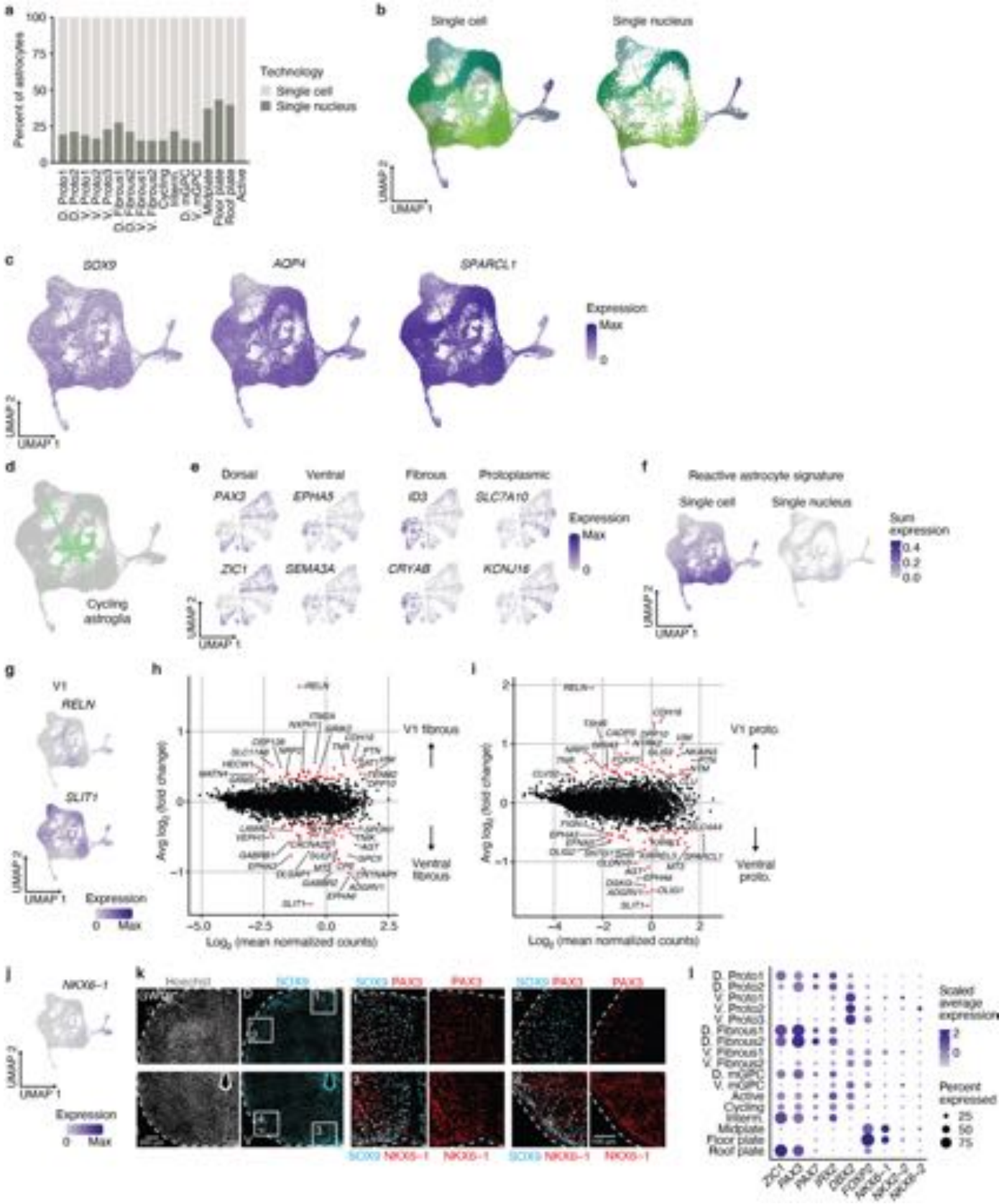
Extended Data Fig. 2 | Comparison of single cell and single nucleus transcriptional data. a. Heatmap showing the correlation of the normalized average gene expression between single cell and single nucleus samples separated by cell type represented as R^2 values. **b.** Heatmaps showing the top 5 gene ontology (GO) terms by cell type of genes enriched in either single cell samples (top) or single nucleus samples (bottom) per cell type, colored by significance level. Values shown were calculated using an over-representation analysis with a hypergeometric test. Exact p-values are shown in Supplementary Table 4. **c.** Heatmap showing gene expression of selected genes within GO terms that are differentially expressed in single cell and single nucleus samples as the sum of gene counts normalized to total counts per cell type.

https://doi.org/10.1038/s41593-023-01311-w



Extended Data Fig. 3 | Quality control and cell types in the OPC/Oligo lineage. **a**. Bar plot showing the percent of single cells and single nuclei in the OPC/Oligo subcluster. **b**. UMAP of OPC/Oligo subcluster split to show single cell and single nucleus samples separately. **c**. Root probability in UMAP space as computed by scVelo. **d**. Bubble plot showing Gene Ontology (GO) enrichment analysis of genes represented in the 10 clusters of velocity variable genes in Fig. 2d. Enrichment was computed using the 'TopGO' R package with a Fisher's Exact test.

https://doi.org/10.1038/s41593-023-01311-w

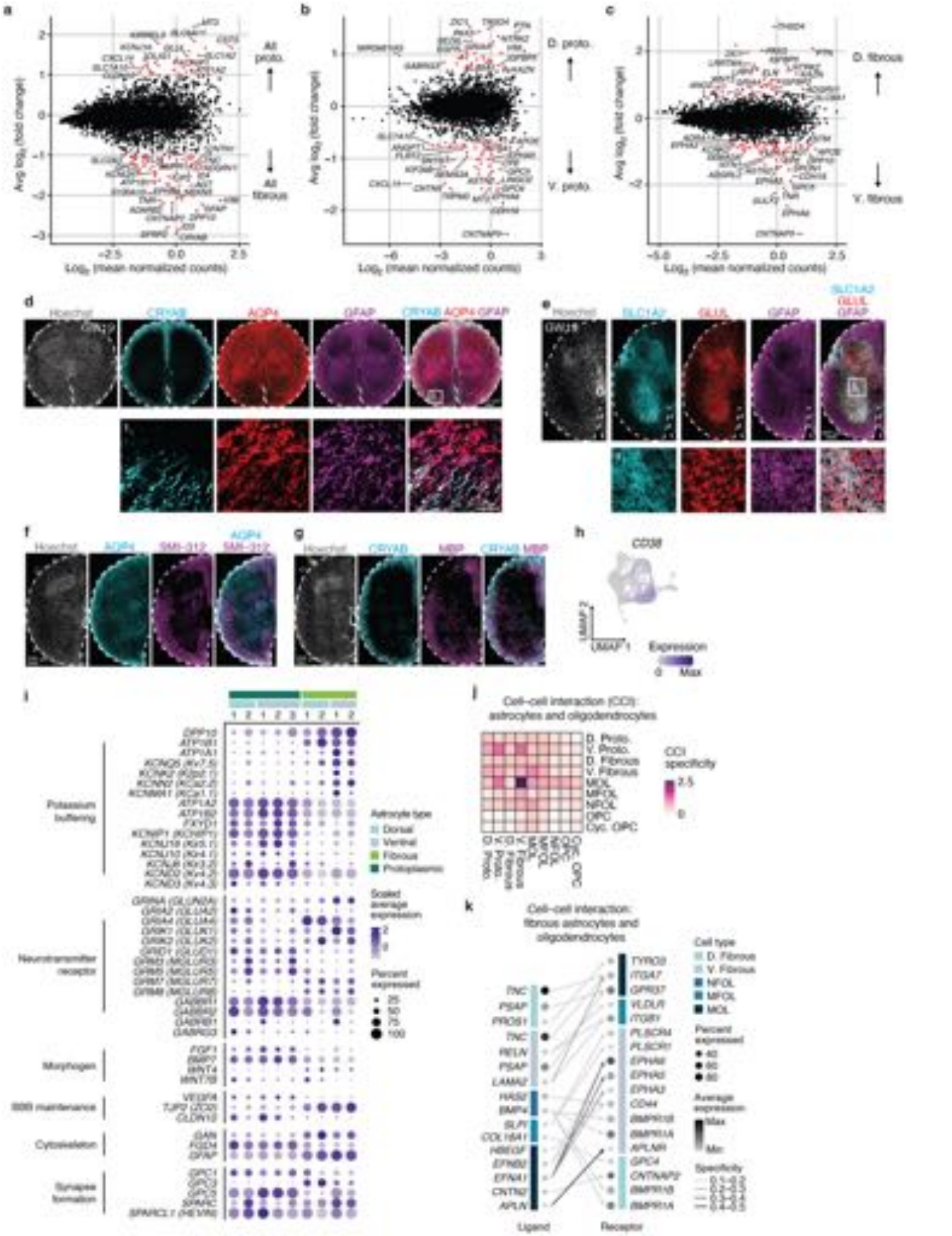


Extended Data Fig. 4 | See next page for caption.

https://doi.org/10.1038/s41593-023-01311-w

Extended Data Fig. 4 | Cell types and quality control in the astroglia subcluster. a. Bar plot showing the percent of single cells and single nuclei in the astroglia subcluster. **b**. UMAP of astroglia subcluster split to show single cell and single nucleus samples separately. **c**. UMAP plots showing expression of astroglia genes. **d**. Highlight of cycling astroglia, highlighting the cells selected for subclustering. **e**. UMAP plots showing gene expression of markers associated with dorsal and ventral, and fibrous and protoplasmic identities within the cycling astroglia subcluster. **f**. Reactive astrocyte signatures shown as sum expression of reactive astrocyte genes from[20] in single cell and single nucleus samples separately. **g**. UMAP plots showing gene expression of V1 astrocyte-associated markers. **h**. MA plot showing differential expression between V1 fibrous astrocytes and the rest of fibrous astrocytes. V1 fibrous astrocytes were selected based on their expression of *RELN* and absence of *SLIT1* (*RELN* counts > 1, *SLIT1* counts = 0). Red dots indicate genes in the top one percent of differentially expressed genes by log fold change. **i**. MA plot showing differential expression between V1 protoplasmic astrocytes and the rest of protoplasmic astrocytes. V1 protoplasmic astrocytes were selected based on their expression of *RELN* and absence of *SLIT1* (*RELN* counts > 1, *SLIT1* counts = 0). Red dots indicate genes in the top one percent of differentially expressed genes by log fold change. **j**. UMAP plot showing expression of *NKX6-1* in the astroglia subcluster. **k**. Representative immunohistochemistry images of SOX9 astrocytes expressing either the dorsal marker PAX3 or the ventral marker NKX6-1 in a coronal GW19 spinal cord cryosection. Top and bottom images correspond to different cryosections. Immunohistochemistry replicated 6 and 4 times in 2 independent samples. **l**. Dot plot showing the expression of transcription factors associated with patterning in the spinal cord. The size of the dots represents the percent of cells expressing each gene while the color depicts the scaled average expression per cell type. Scale bars: 100 µm (insets in **k**), 200 µm (**k**).

https://doi.org/10.1038/s41593-023-01311-w

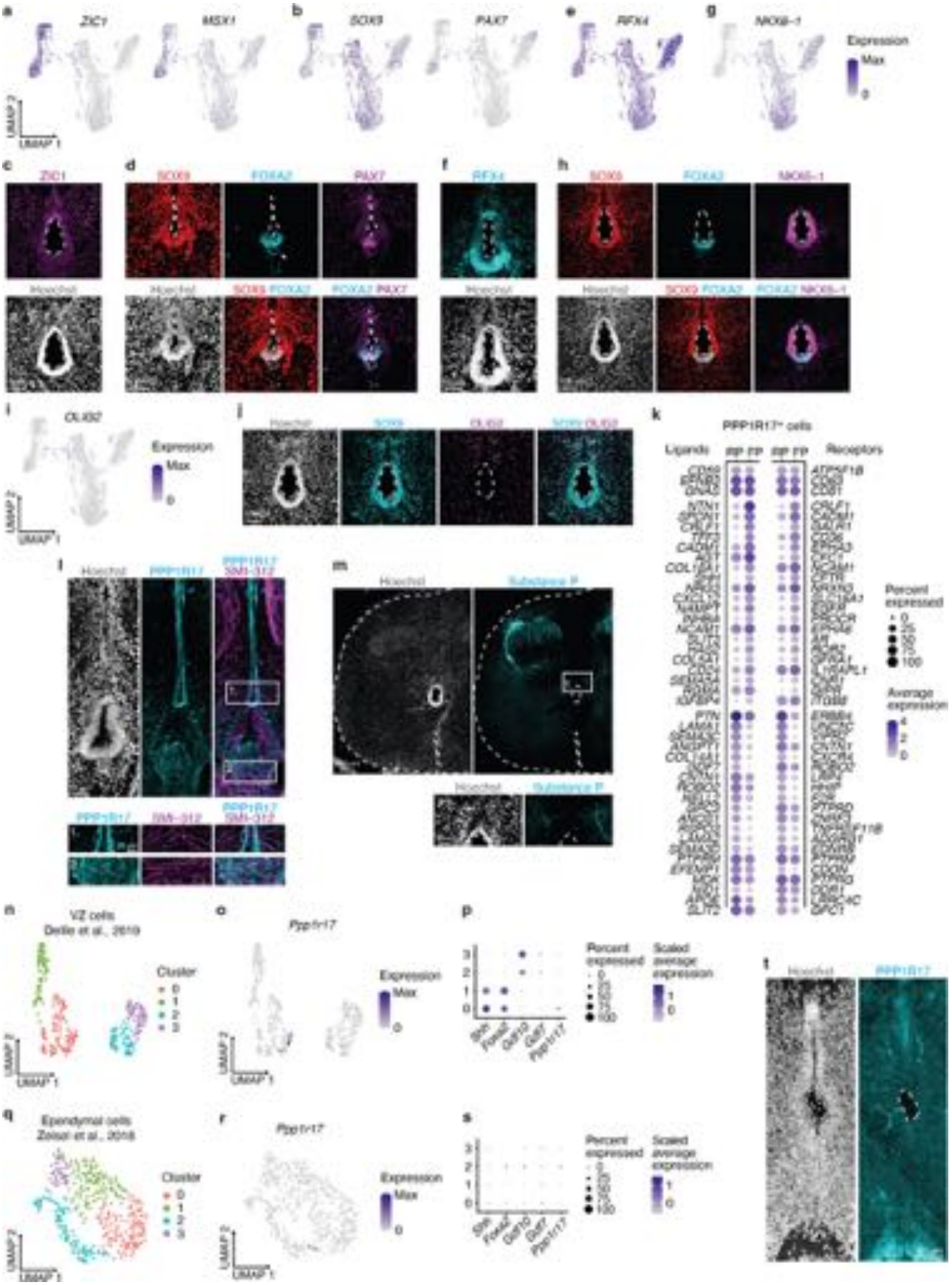


Extended Data Fig. 5 | See next page for caption.

https://doi.org/10.1038/s41593-023-01311-w

Extended Data Fig. 5 | Fibrous and protoplasmic astrocytes in the spinal cord. a. MA plot showing differential expression between protoplasmic and fibrous astrocytes. Red dots indicate genes in the top one percent of differentially expressed genes by log fold change. **b**. MA plot showing differential expression between dorsal and ventral protoplasmic astrocytes. Red dots indicate genes in the top one percent of differentially expressed genes by log fold change. **c**. MA plot showing differential expression between dorsal and ventral fibrous astrocytes. Red dots indicate genes in the top one percent of differentially expressed genes by log fold change. **d**. Representative immunohistochemistry images of the fibrous astrocyte marker CRYAB and colocalization with astrocyte markers AQP4 and GFAP in a coronal GW19 spinal cord cryosection. Immunohistochemistry replicated 3 times in 1 independent sample. **e**. Representative immunohistochemistry images of the protoplasmic astrocyte markers GLUL and SLC1A2 and colocalization with the astrocyte marker GFAP in a coronal GW19 spinal cord cryosection. Immunohistochemistry combination replicated 1 time in 1 independent sample, and independently 4, 4 and 16 times, respectively, in 1-2 independent samples. **f**. Representative immunohistochemistry images of AQP4 and SMI-312 showing the location of astrocytes and axons in the GW19 spinal cord. **g**. Representative immunohistochemistry images of the fibrous astrocyte marker CRYAB and the myelinating oligodendrocyte marker MBP showing the location of white matter in the GW19 spinal cord. **h**. UMAP plot showing expression of CD38 in the astroglia subcluster. **i**. Dot plot showing the expression of selected genes associated with astrocyte functions in fibrous and protoplasmic subtypes. The size of the dots represents the percent of cells expressing each gene while the color depicts the scaled average expression per subtype. **j**. Heatmap showing the sum of specificity scores for all interactions between fibrous and protoplasmic astrocytes and OPC/Oligo cell types as computed by NATMI (see Methods). **k**. Network plot showing interactions between sender and receiver cell types for ligand and receptor pairs. Only fibrous astrocytes and late-stage OPC/Oligo cell types are shown. The size of the dots represents the percent of cells expressing each gene, while the color of the dots depicts their average expression. The color of the line represents the specificity of each interaction as computed by NATMI. Scale bars: 50 μm (insets in **d**, **e**), 200 μm (**d**, **e**, **f**, **g**).

https://doi.org/10.1038/s41593-023-01311-w

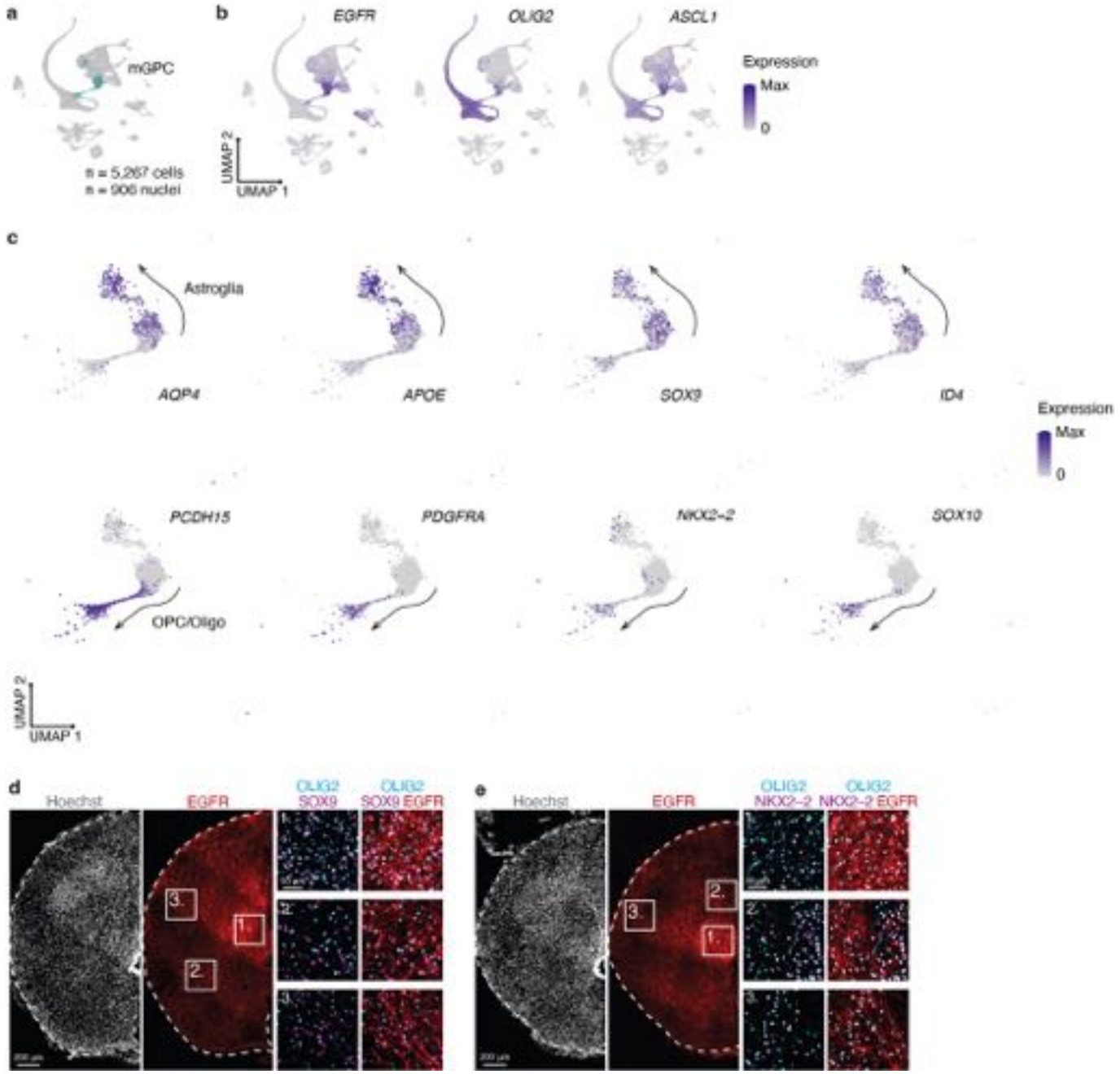


Extended Data Fig. 6 | See next page for caption.

https://doi.org/10.1038/s41593-023-01311-w

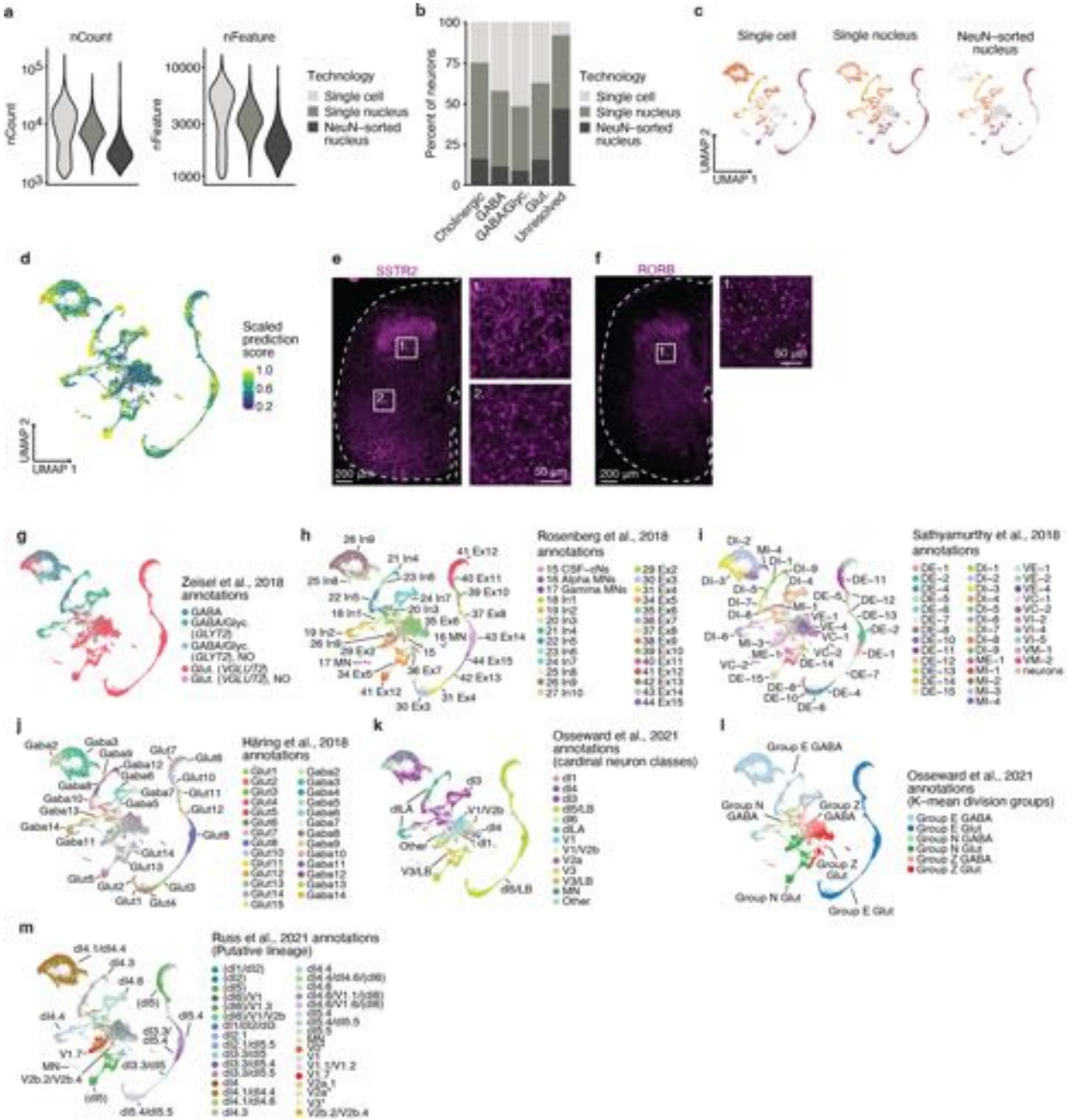
Extended Data Fig. 6 | Ventricular zone (VZ) cells in the human spinal cord at midgestation. a. UMAP plot showing expression of *ZIC1* in the VZ subcluster. **b**. UMAP plots showing expression of *SOX9* and *PAX7* in the VZ subcluster. **c**. Representative immunohistochemistry image showing ZIC1 in the VZ of GW19 spinal cord. Immunohistochemistry replicated 4 times in 2 independent samples. **d**. Representative immunohistochemistry image showing SOX9, FOXA2 and PAX7 in the VZ of GW19 spinal cord. Arrow shows dissociation of ventral cells from the ventricular wall. Immunohistochemistry replicated 2 times in 2 independent samples. Individual markers were replicated 16, 4 and 4 times. **e**. UMAP plot showing expression of *RFX4* in the VZ subcluster. **f**. Representative immunohistochemistry image showing RFX4 in the VZ of GW19 spinal cord. Immunohistochemistry replicated 4 times in 2 independent samples. **g**. UMAP plot showing expression of *NKX6-1* in the VZ subcluster. **h**. Representative immunohistochemistry image showing SOX9, FOXA2 and NKX6-1 in the VZ of GW19 spinal cord. Immunohistochemistry replicated 4 times in 2 independent samples. **i**. UMAP plot showing expression of *OLIG2* in the VZ subcluster. **j**. Representative immunohistochemistry image showing SOX9 and OLIG2 in the VZ of GW19 spinal cord. Immunohistochemistry replicated 7 times in 2 independent samples. **k**. Bubble plot showing expression of ligands and receptors in *PPP1R17*$^{+}$ cells of the roof plate (RP) and floor plate (FP). The first 3 genes were differentially expressed in both FP and RP cells relative to midplate cells. The next 20 were differentially expressed in FP relative to RP cells, and the final 20 were differentially expressed in RP relative to FP cells. Genes were identified as ligands or receptors based on the NATMI database. **l**. Representative immunohistochemistry image showing sites of axon crossing (SMI-312) along the dorsal and ventral midline expressing PPP1R17. Immunohistochemistry replicated 4 times in 2 independent samples. **m**. Representative immunohistochemistry image showing substance P^{+} axons at the midline in GW19 spinal cord. Immunohistochemistry replicated 4 times in 2 independent samples. **n**. Subcluster of ventricular zone (VZ) cells from Delile et al.[60]. **o**. Feature plot showing expression of *PPP1R17* in Delile et al.[60]. **p**. Bubbleplot showing expression of floor plate markers (*Shh*, *Foxa2*), roof plate markers (*Gdf10*, *Gdf7*) and *Ppp1r17* in Delile et al.[60]. **q**. Subcluster of ependymal cells from Zeisel et al.[50]. **r**. Feature plot showing expression of *Ppp1r17* in Zeisel et al.[50] ependymal cells. **s**. Bubbleplot showing expression of floor plate markers (*Shh*, *Foxa2*), roof plate markers (*Gdf10*, *Gdf7*) and *Ppp1r17* in Zeisel et al.[50]. **t**. Representative immunohistochemistry image showing Ppp1r17 in mouse embryonic spinal cord. Immunohistochemistry replicated in n = 3. Scale bars: 20 µm (inset in **l**), 50 µm (**f**, **l**, **t**, inset in **m**), 100 µm (**c**, **d**, **h**, **j**), 200 µm (**m**).

https://doi.org/10.1038/s41593-023-01311-w



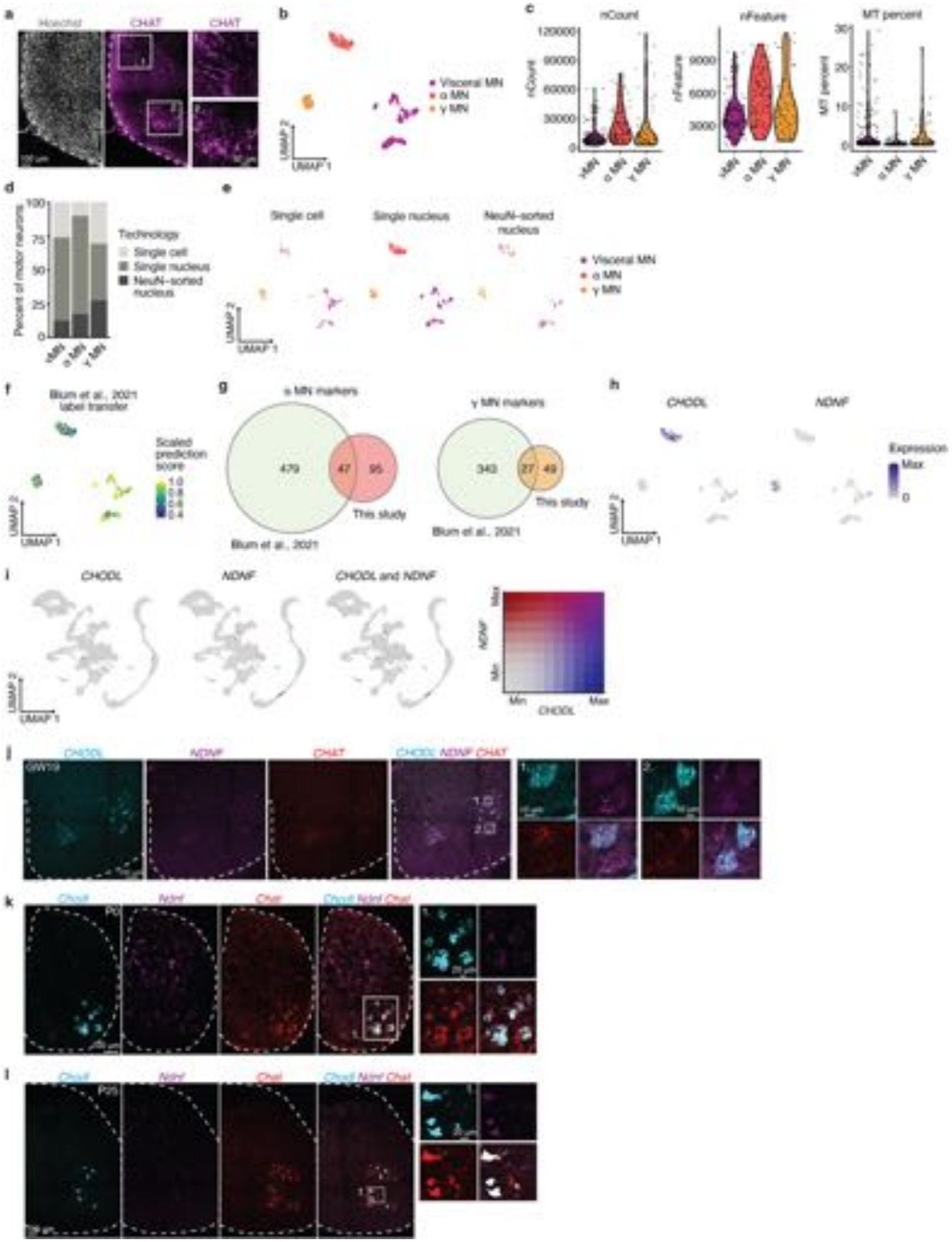
Extended Data Fig. 7 | mGPCs are split into the astrocyte and oligodendrocyte lineages in the spinal cord. a. Highlight of multipotent glial progenitor cells (mGPC) in main UMAP. **b**. UMAP plots showing expression of *EGFR, OLIG2* and *ASCL1* in the main UMAP. **c**. UMAP plots showing expression of astroglia and OPC/Oligo genes in mGPCs within the main UMAP. **d**. Representative immunohistochemistry images showing EGFR$^+$/OLIG2$^+$ mGPCs that colocalize with the astroglia-specific marker SOX9. Immunohistochemistry replicated 4 times in 2 independent samples. **e**. Representative immunohistochemistry images showing EGFR$^+$/OLIG2$^+$ mGPCs that colocalize with the oligodendrocyte lineage marker NKX2-2. Immunohistochemistry replicated 4 times in 2 independent samples. Scale bars: 50 µm (insets in **d**, **e**), 200 µm (**d**, **e**).

https://doi.org/10.1038/s41593-023-01311-w



Extended Data Fig. 8 | Diversity of neurons in the spinal cord. a. Violin plots showing nCount (left) and nFeature (right) in the neuron subcluster separated by single cell, single nucleus and NeuN-sorted nucleus samples. **b**. Bar plot showing the percent of single cells and single nuclei in the neuron subcluster. **c**. UMAP of the neuron subcluster to show single cell, single nucleus and NeuN-sorted nuclei samples separately. **d**. UMAP of the neuron subcluster colored by the scaled prediction score as computed by Seurat for label transfer annotations from Russ et al.[46]. **e**. Representative immunohistochemistry images of neuronal marker SSTR2 in a GW19 spinal cord cryosection. Immunohistochemistry replicated 5 times in 2 independent samples. **f**. Representative immunohistochemistry images of neuronal marker RORB in a GW19 spinal cord cryosection. Immunohistochemistry replicated 3 times in 2 independent samples. **g**. Label transfer showing neuronal annotations from Zeisel et al.[50] in the neuron subcluster. **h**. Label transfer showing neuronal annotations from Rosenberg et al.[47] in the neuron subcluster. **i**. Label transfer showing neuronal annotations from Sathyamurthy et al.[48] in the neuron subcluster. **j**. Label transfer showing annotations from Häring et al.[51] in the neuron subcluster. **k**. Label transfer showing annotations from Osseward et al.[49] based on cardinal spinal cord neuronal classes in the neuron subcluster. **l**. Label transfer showing annotations from Osseward et al.[49] based on K-means division groups in the neuron subcluster. **m**. Label transfer showing annotations from Russ et al.[46] based on Putative lineage groups in the neuron subcluster. Scale bars: 50 µm (insets in **e**, **f**), 100 µm (**e**, **f**).

https://doi.org/10.1038/s41593-023-01311-w

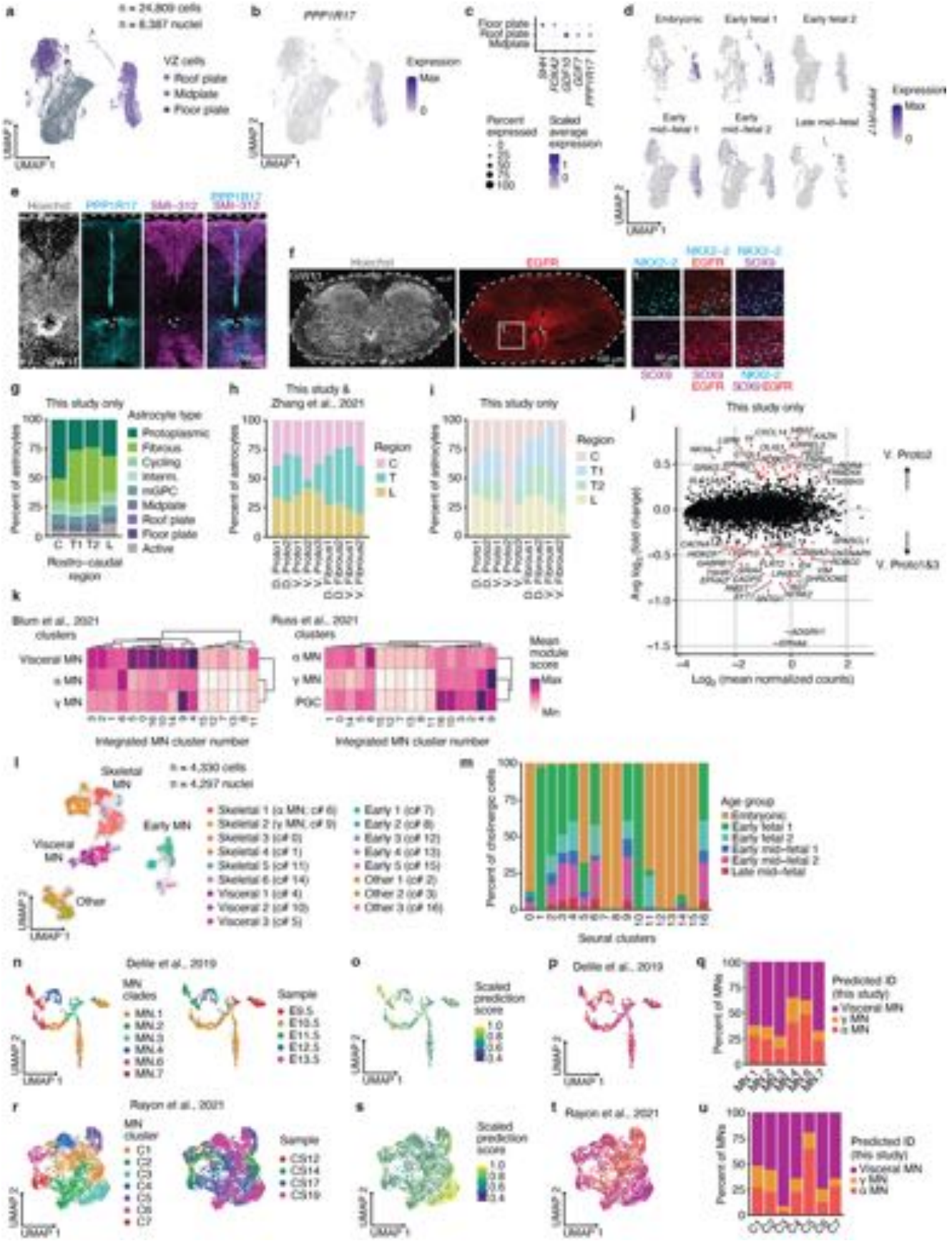


Extended Data Fig. 9 | See next page for caption.

https://doi.org/10.1038/s41593-023-01311-w

Extended Data Fig. 9 | Neuronal diversity and cell types and quality control metrics in the motor neuron subcluster. a. Representative immunohistochemistry images of CHAT motor neurons in a GW19 spinal cord cryosection. Immunohistochemistry replicated 9 times in 2 independent samples. **b**. UMAP of the neuron subcluster, colored by motor neuron subtype. **c**. Violin plots of nCount (left), nFeature (middle) and percent of mitochondrial genes (MT) separated by motor neuron subtype. **d**. Bar plot showing the percent of single cells, single nuclei and NeuN-sorted nuclei per motor neuron subtype. **e**. UMAP of the motor neuron subcluster split to show single cell, single nucleus and NeuN-sorted nucleus samples separately. **f**. UMAP plot showing the scaled prediction scores for the label transfer of Blum et al.[53] motor neuron annotations onto the motor neuron subcluster (this study). **g**. Venn diagrams showing the number of common markers between Blum et al.[53] and this study for alpha motor neurons (left) and gamma motor neurons (right). **h**. Feature plots showing expression of *CHODL* (left) and *NDNF* (middle) in the motor neuron subcluster. **i**. Feature plots showing expression of *CHODL* (left), *NDNF* (middle) and *CHODL* and *NDNF* combined (right) in the neuron subcluster. **j**. Representative *in situ* hybridization of *CHAT*, *CHODL* and *NDNF* in coronal cervical spinal cord cryosections at GW19. Insets show *CHAT*$^+$ motor neurons that express either the alpha marker *CHODL* or the gamma marker *NDNF*. Immunohistochemistry replicated in n = 2 biological replicates. **k**. Representative *in situ* hybridization of *Chat*, *Chodl* and *Ndnf* in coronal mouse spinal cord cryosections at P0. Inset shows *Chat*$^+$ motor neurons that express the alpha marker *Chodl* but not the gamma marker *Ndnf*. Immunohistochemistry replicated in n = 4 biological replicates. **l**. Representative *in situ* hybridization of *Chat*, *Chodl* and *Ndnf* in coronal mouse spinal cord cryosections at P25. Insets show *Chat*$^+$ motor neurons that express both the alpha marker *Chodl* and the gamma marker *Ndnf*. Immunohistochemistry replicated in n = 3 biological replicates. Scale bars: 10 µm (insets in **j**), 20 µm (insets in **k**, **l**), 50 µm (insets in **a**), 100 µm (**a**, **j**, **k**, **l**).

https://doi.org/10.1038/s41593-023-01311-w



Extended Data Fig. 10 | See next page for caption.

https://doi.org/10.1038/s41593-023-01311-w

Extended Data Fig. 10 | Developing human spinal cord integrated dataset. **a**. UMAP showing a subcluster of ventricular zone (VZ) cells in the integrated dataset colored by cell type. **b**. Feature plot showing expression of *PPP1R17* in the integrated dataset. **c**. Bubble plot showing expression of floor plate markers (*SHH*, *FOXA2*), roof plate markers (*GDF10*, *GDF7*) and *PPP1R17* in the integrated dataset. **d**. Feature plots showing expression of *PPP1R17* in the integrated dataset split by age. **e**. Representative immunohistochemistry images showing the midline glia marker PPP1R17 and the axonal marker SMI-312 in spinal cord cryosections at GW11. Immunohistochemistry replicated 2 times in 1 independent sample. **f**. Representative immunohistochemistry image showing EGFR, the astroglia marker SOX9 and the OPC/Oligo lineage marker NKX2-2 in spinal cord cryosections at GW11. Immunohistochemistry replicated 2 times in 1 independent sample. **g**. Bar plot showing the percent of astroglia clusters in cervical (C), thoracic (T1, T2) and lumbar (L) regions of the spinal cord from this study. **h**. Bar plot showing the percent of fibrous and protoplasmic astrocyte clusters in C, T and L regions of the spinal cord from this study and Zhang et al.[58]. **i**. Bar plot showing fibrous and protoplasmic astrocyte clusters separated by rostro-caudal region in this study only (GW18). **j**. MA plot showing differential expression between the V. proto2 astrocyte cluster and the rest of the ventral protoplasmic clusters (V. proto 1 and 3) from this study (GW18). Red dots indicate genes in the top one percent of differentially expressed genes by log fold change. **k**. Heatmaps showing the alpha/gamma/visceral or PGC mean module score per Seurat cluster in the integrated cholinergic neuron dataset for Blum et al.[53] (top) and Russ et al.[46] (bottom). Module scores were calculated using the 'AddModuleScore' function in Seurat with ≤50 genes with P-value < 0.01 and $Log_2FC > 1$ from either alpha, gamma or visceral/PGC clusters from each study. **l**. UMAP showing the cholinergic neuron subcluster in the integrated dataset showing cell type annotations and Seurat cluster numbers (c#). **m**. Bar plot showing the percent of cholinergic cells per Seurat cluster in each age group. **n**. UMAP plots showing motor neurons in Delile et al.[60], colored by their assigned clades (left) and sample age (right). **o**. UMAP plot showing the scaled prediction scores for the label transfer of this study's motor neuron type annotations onto Delile et al.[60] motor neurons. **p**. Label transfer showing predicted IDs from motor neuron clusters in this study onto Delile et al.[60] motor neurons. **q**. Bar plot showing the percent of predicted identities based on annotations from this study in Delile et al.[60] MN clades. **r**. UMAP plots showing the motor neuron subcluster in Rayon et al.[6], colored by Seurat clusters (left) and sample age (right). Approximately, Carnegie stage (CS) 12 corresponds to GW4; CS14 to GW5; CS17 to GW6; and CS19 to GW7. **s**. UMAP plot showing the scaled prediction scores for the label transfer of this study's motor neuron type annotations onto Rayon et al.[6] motor neurons. **t**. Label transfer showing predicted IDs from motor neuron clusters in this study onto Rayon et al.[6] motor neurons. **u**. Bar plot showing the percent of predicted identities based on annotations from this study in Rayon et al.[6] motor neuron clusters. Scale bars: 50 μm (**e**, insets in **f**),100 μm (**f**).

nature portfolio

Corresponding author(s): Sergiu Pasca

Last updated by author(s): 03/14/2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software or code was used for data collection
Data analysis	Cell Ranger (v5.0.0, 10x Genomics), velocyto (v0.17.17), Seurat (v3.9.9.9024), R (v3.6.1), R package fgsea (v1.19.2), NATMI, Python (3.6.1), scVelo toolkit (v.0.2.3), reticulate R-Python interface, ImageJ (v2.0.0), LASX software (Leica, v3.1.5.16308), R package FNN (v1.1.3), R package future (v1.21)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

De-identified spinal cord samples were obtained at Stanford University School of Medicine from elective pregnancy terminations under a protocol approved by the Research Compliance Office at Stanford University.
Data used for the analyses presented in this work are available under GEO accession number GSE188516. A website associated with the manuscript, including an interactive data browser, is available at https://devspinalcord.su.domains/.
Data for label transfer analysis performed in this study was obtained from GEO using the following accession numbers: GSE110823 (Rosenberg et al., 2018),

GSE103892 (Sathyamurthy et al., 2018), GSE103840 (Häring et al., 2018), GSE161621 (Blum et al., 2021), and GSE171892 (Rayon et al., 2021). Data for Zeisel et al., 2018 was obtained from mousebrain.org. Data for Delile et al., 2019 was obtained from E-MTAB-7320. RNA count data was obtained from the authors from Russ et al., 2021.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	A total of five PCW17-18 spinal cord samples were used for this study with a total of 157,108 single cell/nuclei transcriptomes collected. Samples were delivered on ice and processed for single cell analyses or immunocytochemistry within three hours of the procedure. No sample size calculation was performed.
Data exclusions	Low quality cells/nuclei and doublets were removed from the analysis. Cells/nuclei were considered low quality if they had low nCount and high mitochondrial percentage. Cells/nuclei were considered doublets if they expressed well known cell-type specific markers for two or more cell types. A detailed description of any exclusions and their rationale is shown in the Methods section.
Replication	The main findings in the transcriptomic analysis were validated using immunohistochemistry and/or in situ hybridization in spinal cord cryosections. Immunohistochemistry and in situ hybridization experiments were independently repeated 1 to 16 times.
Randomization	No treatment/intervention was performed and therefore randomization is not applicable
Blinding	No treatment/intervention and/or quantification was performed and therefore blinding is not applicable

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	AQP4 (GeneTex ,GTX133151), CD38 PE (Abcam, ab36422), CHAT (Millipore, AB144P), CRYAB (Abcam, ab13496), EGFR (Abcam, ab231), FOXA2 (Abcam, ab108422), GABA (Sigma-Aldrich, A2052), GFAP (DAKO, Z0334), GFAP (Thermo Fisher Scientific, 13-0300), GLUL (Abcam, ab176562), GLYT2 (Synaptic Systems, 272 003), KI67 (BD Biosciences, 550609), MBP (Millipore, MAB386), NeuN (Millipore, MAB377), NKX2-2 (DSHB, 74.5A5), NKX6-1 (DSHB, F55A10), OLIG2 (Millipore, AB9610), PAX3 (DSHB, Pax3), PAX7 (DSHB, PAX7), PPP1R17 (Atlas Antibodies, HPA047819), RFX4 (Atlas Antibodies, HPA050527), SLC1A2 (Millipore, AB1783), SMI-312 (BioLegend, 837904), SOX9 (R&D Systems, AF3075), Substance P (Abcam, ab106291), VGLUT2 (Synaptic Systems, 135 403), ZIC1 (DSHB, PCRP-ZIC1-1E3) Alexa Fluor donkey secondary antibodies conjugated to 488, 564 or 647 fluorophores (Invitrogen or Jackson Immunoresearch)
Validation	AQP4: validation and references (2) on manufacturer's website CD38 PE: validation on manufacturer's website CHAT: validation and references (>30) on manufacturer's website. See also Andersen et al 2020 CRYAB: validation and references (42) on manufacturer's website EGFR: validation and references (35) on manufacturer's website. See also Trevino et al 2021 FOXA2: validation and references (64) on manufacturer's website. GABA: validation and references (522) on manufacturer's website. See also Birey et al 2017, Andersen et al 2020 GFAP, DAKO: validation and references (345) on manufacturer's website. See also Birey et al 2017, Trevino et al 2020, Andersen et al

2020, Trevino et al 2021
GFAP, Thermo Fisher Scientific: validation and references (117) on manufacturer's website.
GLUL, validation and references (11) on manufacturer's website.
GLYT2: validation and references (8) on manufacturer's website. See also Andersen et al 2020
KI67: validation and references (14) on manufacturer's website. See also Trevino et al 2021
MBP: validation and references (>30) on manufacturer's website. See also Marton et al 2019, Andersen et al 2020
NeuN: validation and references (>50) on manufacturer's website
NKX2-2: validation and references (29) on manufacturer's website. See also Andersen et al 2020
NKX6-1: validation and references (40) on manufacturer's website
OLIG2: validation and references (>30) on manufacturer's website. See also Andersen et al 2020, Trevino et al 2021
PAX3: validation and references (62) on manufacturer's website
PAX7: validation and references (112) on manufacturer's website
PPP1R17: validation and references (4) on manufacturer's website. See also Trevino et al 2021 for validation in developing human cortex
RFX4: validation and references (1) on manufacturer's website. See also Trevino et al 2020
SLC1A2: validation and references (109) on manufacturer's website.
SMI-312: validation and references (61) on manufacturer's website
SOX9: validation and references (48) on manufacturer's website. See also Trevino et al 2020, Trevino et al 2021
Substance P: validation and references (3) on manufacturer's website
VGLUT2: validation and references (92) on manufacturer's website. See also Andersen et al 2020
ZIC1: validation and references (2) on manufacturer's website. See also Trevino et al 2020