Title

Direct lineage conversion of postnatal mouse cortical astrocytes to oligodendrocyte lineage cells.

Authors

Justine Bajohr[1-2], Erica Y. Scott[1,3-5], Arman Olfat[1], Mehrshad Sadria[5-6], Kevin Lee[1], Maria Fahim[1], Hiba T. Taha[1,2], Daniela Lozano Casasbuenas[1,6], Ann Derham[1], Scott A. Yuzwa[6], Gary D. Bader[5,8], Maryam Faiz[1,2,6 # $]

Affiliations

1. Department of Surgery, Temerty Faculty of Medicine; University of Toronto; Toronto, Ontario, M5S 1A8, Canada
2. Institute of Medical Science, Temerty Faculty of Medicine; University of Toronto; Toronto, Ontario, M5S 1A8, Canada
3. Institute of Biomedical Engineering, Faculty of Applied Science and Engineering; University of Toronto; Toronto, Ontario, M5S 3G9, Canada
4. Department of Chemistry, Faculty of Arts and Science; University of Toronto; Toronto, Ontario, M5S 3G3, Canada
5. Donnelly Centre for Cellular and Biomolecular Research, Temerty Faculty of Medicine; University of Toronto; Toronto, Ontario, M5S 3E, Canada
6. Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada.
7. Department of Laboratory Medicine and Pathobiology, Temerty Faculty of Medicine; University of Toronto; Toronto, Ontario, M5S 1A8, Canada
8. Department of Molecular Genetics, Temerty Faculty of Medicine; University of Toronto; Toronto, Ontario, M5S 1A8

#Correspondence: maryam.faiz@utoronto.ca

$Lead contact: maryam.faiz@utoronto.ca

SUMMARY

Oligodendrocyte lineage cells (OLCs) are lost in many CNS diseases. Here, we investigate the generation of new OLCs via ectopic expression of *Sox10*, *Olig2* or *Nkx6.2* in mouse postnatal astrocytes. Using stringent analyses including, Aldh1l1-astrocyte fate mapping and live cell imaging we confirm that *Sox10* and *Olig2*, but not *Nkx6.2*, directly convert Aldh1l1^{pos} astrocytes to MBP+ and PDGFRα+ induced OLCs (iOLCs), respectively. With single cell RNA sequencing (scRNA-seq) we uncover the molecular signatures of iOLCs. Transcriptomic analysis of *Sox10*- and control cultures over time reveals a clear trajectory from astrocytes to iOLCs. Finally, perturbation models CellOracle and Fatecode support the idea that *Sox10* drives cells towards a terminal iOLC fate. Altogether, this multidimensional analysis shows bonafide conversion of astrocytes to iOLCs using *Sox10* or *Olig2* and provides a foundation for astrocyte DLR strategies to promote OLC repair.

KEYWORDS

Astrocyte, oligodendrocyte, reprogramming, deep learning, single cell RNA sequencing, fate mapping

INTRODUCTION

Oligodendrocytes (OLs) are best known as the myelinating cells of the central nervous system (CNS). OLs ensheath neuronal axons to enable fast propagation of action potentials. Consequently, the loss or dysfunction of OLs results in impaired neurological function that is characteristic of many types of CNS disease and injury. Multiple sclerosis (MS), Alzheimer's Disease, spinal cord injury, white matter stroke and cerebral palsy are all characterized by oligodendrocyte failure. Thus, therapeutic strategies aimed at replacing OLs are of significant clinical interest[1].

Direct lineage reprogramming (DLR) aims to generate new target cells lost to disease via the forced conversion of donor cells. Typically, DLR is performed by the overexpression of transcription factors. Early pioneering work from the Tesar Lab demonstrated fibroblast to induced oligodendrocyte progenitor cell (iOPC) conversion by ectopic expression of *Olig2*, *Sox10*, and *Nkx6.2*, determinants of OL cell fate in the embryonic brain[2]. The combination of *Sox10* and *Olig2* was also later used to generate iOPCs from pericytes[3]. Therapeutically, however, these newly generated iOPCs still require transplantation into the brain. Therefore, methods to reprogram endogenous CNS cells would be advantageous.

In parallel, astrocytes, CNS-resident cells, have emerged as donor cells in DLR strategies aimed at generating new neurons[4–7]. Astrocytes are an attractive donor cell type for OL conversion given their shared neural origin[8,9]. Astrocytes may already have relevant epigenetic marks and active TFs that could make DLR faster or more efficient[10–13]. In addition, closely related cells may require fewer TFs for conversion. Indeed, conversion of astrocytes to an induced oligodendrocyte-like cell was suggested using *Sox10* alone[14]. Therefore, we reasoned that single 'Tesar' factors, *Olig2*, *Sox10*, or *Nkx6.2,* could be used to force astrocyte conversion to new, induced oligodendrocyte lineage cells (iOLCs).

The OL lineage is comprised of oligodendrocyte progenitor cells (OPCs) that give rise to committed oligodendrocyte progenitors (COPs). These COPs differentiate into mature OLs (mOLs), comprising at least 6 different states[15]. Of interest, during OL development, each of these 'Tesar' factors shows different temporal expression and plays a different role in OL fate specification. *Olig2,* long considered an OL lineage fate determinant[16], is also expressed in astrocytes[17,18], which suggests a broad role in early glial commitment[19,20]. *Sox10* is important throughout OL development; *Sox10* promotes early OL lineage specification by re-inducing *Olig2*[21] and by inhibiting *Sufu*[22], but is also required for OL survival following myelination[23,24]. In contrast, *Nkx6.2* is expressed late in OL development, with myelin genes *Mbp* and *Mog,* and plays a role in regulating myelination[25,26]. Given the unique roles of each of these TFs in development, we further hypothesized that single Tesar factors could be used to create distinct types of iOLCs that ranged from OPCs to mature, myelinating OLs.

Recent studies have highlighted the need for rigorous reporting of DLR outcomes, following controversy of astrocyte to neuron DLR *in vivo*. A landmark study using astrocyte fate mapping strategies suggested that conversion was misrepresented as a result of AAV and promoter confounds[27]. Therefore, it is important that new DLR paradigms utilize fate mapping, and stringent, multi-faceted analysis to determine the origin of the newly generated cells.

Here, we used lentiviral delivery of *Olig2*, *Sox10*, or *Nkx6.2* to investigate single TF conversion of postnatal (P0-P5) GFAP+ cortical astrocytes to different types of iOLCs. Lineage tracing

experiments using Aldh1l1-CreERT2;Ai14 mice demonstrated that *Sox10* and *Olig2* convert Aldh1l1+ astrocytes to iOLCs. Moreover, live cell imaging, single cell RNA sequencing (scRNA seq) and deep learning methods further support the findings that iOLCs can be generated from astrocytes following TF delivery. Altogether, our findings show bonafide astrocyte to iOLC DLR and lay the groundwork for future studies utilizing DLR for diseases involving OLC dysfunction and loss.

METHODS

Animals

All experiments were performed in accordance with approved Animal Use Protocols (AUP 20012151, 25-0389H) from the Division of Comparative Medicine at the University of Toronto. P0-P5 Ai14 **(**B6;129S6- *Gt(ROSA)26Sor*$^{tm14(CAG\text{-}tdTomato)Hze}$/J, RRID:IMSR_JAX:007908) and Ai14;Ald1l1-CreERT2 mice (Ai14 crossed to B6N.FVB-Tg(Aldh1l1-cre/ERT2)1Khakh/J [RRID:IMSR_JAX:031008]) were used to generate postnatal astrocyte cultures.

Cell Culture

Cortical astrocytes were isolated from male and female P0-P5 mice as previously described[28]. Briefly, mice were decapitated, followed by the removal of the skull and meninges. Cortices were dissected, pooled, and mechanically dissociated in astrocyte media [DMEM (Gibco Catalogue No. 10569-010), 10% fetal bovine serum (FBS) (Gibco Catalogue No. 10082147) and 1% penicillin/streptomycin (Gibco Catalogue No. 15140122)]. Cells were cultured in flasks pre-coated with 10 µg/ml poly-d-lysine (Sigma Catalogue No. P6407), and incubated at 37˚C, 5% CO2. Media was changed the day following isolation and every other day thereafter. Once the cells reached 80% confluency, typically after 6 days, flasks were placed on an orbital shaker for 30 minutes at 180rpm to remove contaminating microglia. Astrocyte media was replaced. For Ai14;Aldh1l1-CreERT2 cultures, 1uM 4-OHT was added to the astrocyte media at this step. Flasks were returned to the orbital shaker overnight at 180rpm, followed by vigorous shaking for one minute to remove contaminating OLCs. Media was removed and astrocytes were incubated in TrypLE Express Enzyme (Gibco Catalogue No. 12604013) for 5 minutes at 37˚C, 5% CO2 to lift off the astrocytes. To inactivate the enzyme, astrocyte media was added at a 3:1 ratio (media:TrypLE). Cell suspension was collected and centrifuged at 300 x g for 5 minutes. Following removal of the supernatant, the pellet was resuspended in astrocyte media. Cells were then plated on poly-l-ornithine/laminin coated coverslips at either 50,000 or 70,000 cells/well in 24 well plates and incubated at 37˚C, 5% CO2. For live cell analysis, cells were plated at 10,000 cells/ well in 96 well plates with poly-l-ornithine/laminin coating. For poly-l-ornithine/laminin coating, 0.1mg/ml poly-l-ornithine (Sigma Catalogue No. P4957) was added to dishes overnight at 37˚C, washed 3 times with 1X phosphate buffered saline (PBS), and then dishes were incubated for two hours at 37˚C with 10 µg/ml laminin (Sigma Catalogue No. L2020).

Reprogramming

Lentiviral particles were purchased from VectorBuilder. For Ai14 reprogramming, LV-hGFAP::*Sox10*-P2A-Cre, LV-hGFAP::*Olig2*-P2A-Cre, LV-hGFAP::*Nkx6.2*-P2A-Cre, and control LV-hGFAP::BFP-T2A-Cre were used. For Ai14;Aldh1l1-CreERT2 reprogramming, LV-hGFAP::*Sox10*-P2A-zsGreen, LV-hGFAP::*Olig2*-P2A-zsGreen, LV-hGFAP::*Nkx6.2*-P2A-zsGreen and control LV-hGFAP::zsGreen were used. A multiplicity of infection (MOI) of 100 was used for all experiments. Virus-containing astrocyte media was placed on the cells and left overnight. Viral media was replaced with fresh astrocyte media one day post transduction (DPT). Three DPT, cells

were switched to OPC differentiation media[2] [DMEM/F12/Glutamax (Gibco Catalogue No. 11330032), 1X N2 (Gibco Catalogue No. 17502048), 1X B27 without vitamin A (Gibco Catalogue No. 17504044), 200 ng/ml SHH (R&D Systems Catalogue No. 464-SH), 20 ng/ml FGF (R&D Systems Catalogue No. 3139-FB), 4 ng/ml PDGF (Sigma Catalogue No. SRP3228)]. At 10 DPT, the cells were switched to OL differentiation media[2] [DMEM/F12/Glutamax, 1X N2, 1X B27 without vitamin A, 40 ng/ml T3 (T2877 Catalogue No. Sigma), 200 ng/ml SHH, 100 ng/ml Noggin (R&D Systems Catalogue No. 1967-NG), 10 μM cAMP (Sigma Catalogue No. A9501), 100 ng/ml IGF (R&D Systems Catalogue No. 791-MG), 10 ng/ml NT3 (Sigma Catalogue No. SRP6007)].

Live Cell Analysis

Astrocytes isolated from Ai14;Aldh1l1-CreERT2 mice were plated in 96 well plates, transduced and imaged every hour from 7 to 12DPT using the Apotome live cell system (Zeiss). 25 z-stack tiled images per well were captured with brightfield as well as the 488nm and 568nm fluorescent wavelength. Images were stitched to create a continuous video for each well. At 12DPT, cells were fixed and stained for OLC markers to confirm fate. Each well was then re-imaged at this final timepoint and OLC+ reprogrammed cells were matched to the live cell video and retrospectively analyzed for starting cell morphology and fluorescent expression.

Immunocytochemistry

Cells were fixed in 4% paraformaldehyde (PFA) (Sigma Catalogue No. P6148) for 20 minutes followed by three washes with 1X PBS. Cell membranes were permeabilized with 0.1% Triton-X-100 (Sigma Catalogue No. X100) for 10 minutes at room temperature, followed by three washes with 1X PBS, and then blocked with 5% milk for one hour at room temperature. Cells were incubated with primary antibodies in 1X PBS overnight at 4ºC, washed three times with 1X PBS, and then incubated with secondary antibodies and DAPI (Sigma Catalogue No. D9542) in 1X PBS at room temperature for one hour. Following three final 1X PBS washes, coverslips were mounted on glass slides (Fisher Scientific Catalogue No. 125523) with Mowiol mounting solution (Sigma Catalogue No. 81381). For staining of membrane bound proteins (O4, PDGFRα), no permeabilization step with Triton-X-100 was performed.
Primary antibodies: mouse anti-SOX10 (RRID:AB_10844002, 1:250), rabbit anti-PDGFRα (RRID:AB_2892065, 1:500), mouse anti-O4 (RRID:AB_357617, 1:1000) and rat anti-MBP RRID:AB_305869, 1:50).

Secondary antibodies: anti-mouse IgG 488 (Invitrogen Catalogue No. A32723) and 647 (Invitrogen Catalogue No. A32728), anti-mouse IgM heavy chain 488 (Invitrogen Catalogue No. A21042) and 647 (Invitrogen Catalogue No. A21238), anti-rabbit 488 IgG (Invitrogen Catalogue No. A11034) and 647 (Invitrogen Catalogue No. A32733), anti-rat IgG 488 (Invitrogen Catalogue No. A21208) and 647 (Invitrogen Catalogue No. A21247) all at 1:1000.

Microscopy and Image Analysis

Fluorescent images for quantification were taken on an LSM 880 Elyra Superresolution and LSM 900 (Zeiss) using a 20x objective and Zen Blue software (Zeiss). Post-acquisition linear adjustments of brightness for all channels were made to micrographs using the Zen Blue software in Figure 1C, Figure 2E-F and Figure S4D. For quantification, ten regions of interest were selected at random for each well. Images were analyzed using ImageJ software (National Institutes of

Health, RRID:SCR_003070). Reprogramming efficiency was calculated as a measure of total OLC marker$^+$reporter$^+$DAPI$^+$ cells over total reporter$^+$DAPI$^+$cells.

scRNA-seq capture and processing

At 14DPT, LV-hGFAP::*Sox10*-P2A-Cre, LV-hGFAP::*Olig2*-P2A-Cre, LV-hGFAP::*Nkx6.2*-P2A-Cre, and control LV-hGFAP::BFP-T2A-Cre cultures were processed using the BD Rhapsody System (BD Biosciences) and then sequenced. For single-cell isolation, an average of 9813.75 viable cells were captured in wells at cell load (Table S1). The BD Rhapsody scanner reported an average multiplet rate of 10.13% and an average number of wells with viable cells and a bead of 7081.5 (Table S1). Detailed metrics for each sample can be found in Table S1. Samples were down-sampled to 2500 cells and carried through and converted to cDNA using the BD Rhapsody WTA Reagent Kit (Becton Dickinson Canada, Catalogue No. 633802). Each cell was sequenced at approximately 100 million reads per cell (at least 2x150 bp paired-end reads) on a Novaseq (Donnelly Sequencing Centre, University of Toronto).

In addition, LV-GFAP::*Sox10* and control LV-GFAP::Cre cultures were collected prior to transduction, at 3DPT and 8DPT, and processed using the BD Rhapsody System (BD Biosciences) and then sequenced. For single-cell isolation, an average of 10263.8 viable cells were captured in wells at cell load (Table S2). The BD Rhapsody scanner reported an average multiplet rate of 6.86% and an average number of wells with viable cells and a bead of 7785 (Table S2). Detailed metrics for each sample can be found in Table S2. Samples were down-sampled to 2500 cells and carried through and converted to cDNA using the BD Rhapsody WTA Reagent Kit (Becton Dickinson Canada, Catalogue No. 633802). Each cell was sequenced at approximately 100 million reads per cell (at least 2x150 bp paired-end reads) on a Novaseq (Donnelly Sequencing Centre, University of Toronto).

scRNA-seq analysis

Fastq files were first demultiplexed with Kallisto[29] (RRID:SCR_016582)(v0.48.0) and Bustools[30] (RRID:SCR_018210)(V 0.41.0) using supplied whitelists (Data S1) with the -BDWTA option and aligning to GRCm38.96 with Cre sequence appended to the end. Bustools[30] was then used to generate gene count tables. Cells were plotted based upon UMI counts per barcode, thresholds were selected based on inflection point of UMI count per barcode plots. These thresholds produced read and gene count distributions that were comparable between all treatment groups (Supplemental Figure 1A, Supplemental Figure 5A). Gene count tables were made into S4 objects, scaled, normalized and dimensions reduced (PCA then UMAP) using the Seurat[31] package (RRID:SCR_016341) (v4.1). Clusters identified as microglia and as vascular and leptomeningeal cells (VLMCs) (Supplemental Figure 1B-C [cluster 8], Supplemental Figure 5B-C [cluster 6 and 8]) were removed prior to further analysis.

Gene markers for oligodendrocyte lineage cells were adopted from studies observing *in vivo* mouse oligodendrocyte lineage cells across several areas in young and mature CNS tissues[15]. These markers were converted into percent expression of each UMAP cluster using the PercentageFeatureSet function from Seurat[31], with further gene resolution displayed by heatmaps created using ComplexHeatmap[32] (RRID:SCR_017270) (v12.13.1). An additional set of gene markers, demonstrating similar, but less resolved, conclusions was also used from an *in vitro* rat study looking at OLCs from the cortex[33]. Stacked violin cluster dot plots were made with scCustomize[34](RRID:SCR_024675)(v2.1.2). For pseudotime and trajectory analysis Slingshot[35] (RRID:SCR_017012)(v2.3.1) and Monocle3[36–40] (RRID:SCR_018685) were used. For *in silico*

Sox10 perturbation, CellOracle[41] (v0.14.0) was used. Plots were produced using ggplot2[42] (RRID:SCR_014601)(v3.3.5) and figure generating scripts were run in R studio[43] (v4.2.0), with demultiplexing using Kallisto[29] and Bustools[30] run on a Compute Canada HPC cluster. All scripts used for processing of scRNA-seq data and for figure generation can be found at github.com/eyscott.

Deep-learning analysis

Fatecode[44] was used to identify key genes for cellular transition. To optimize the autoencoder and subsequent classifier configuration, a grid search was conducted to systematically evaluate various combinations of hyperparameters. These included: latent layer size, number of nodes in the first and second layers of the autoencoder, classifier architecture, and type of activation function. The grid search aimed to identify the hyperparameters that minimized a combined reconstruction and classification loss function, signifying the optimal performance for our specific dataset. Perturbations on the latent space were performed and cell classifications as well as genes associated with each perturbation were obtained.

All scripts used for processing and for figure generation of Fatecode analysis can be found at https://github.com/MehrshadSD.

Statistical Analysis

Percentage values were transformed using the arcsine square root transformation and assessed for normal distribution using the Shapiro-Wilks test. When distribution and variance were equal, a matched pairs one-way ANOVA ([Ai14:D8, D10, D12 O4, D12 MBP], [Ai14;Aldh1l1-CreERT2: D12 MBP zsGreen+, D12 MBP tdTomato+) or one-way ANOVA ([Ai14:D14], [Ai14;Aldh1l1-CreERT2: D12 MBP zsGreen+,tdTomato+]) , or paired t-test ([Ai14;Aldh1l1-CreERT2: D12 PDGFRα zsGreen+tdTomato+, D12 PDGFRα tdTomato+, O4 zsGreen+, O4 tdTomato+]) was performed to compare reprogramming efficiency of TF groups to a control group (Ai14: LV-GFAP::Cre, Ai14;Aldh1l1-CreERT2: LV-GFAP::zsGreen). When transformed values did not follow a Gaussian distribution, a Kruskal-Wallis test (Ai14: D12 PDGFRα, D14 PDGFRα) or Wilcoxon test (Ai14;Aldh1l1-CreERT2: PDGFRα zsGreen+, O4 zsGreen+tdTomato+) was performed to compare reprogramming efficiency to a control group (Ai14: LV-GFAP::Cre, Ai14;Aldh1l1-CreERT2: LV-GFAP::zsGreen). In both cases, Dunnett's post-hoc testing was performed to correct for multiple comparisons. Differences were considered significant at $p < 0.05$. Values are presented as mean ± SEM. The statistical software used for transformation, distribution, variance, ANOVA, Kruskal-Wallis, t-test, Wilcoxon test and Dunnett's analysis was GraphPad Prism version 9.0.1 (RRID:SCR_002798).

RESULTS

Different types of iOLCs are generated following expression of *Olig2*, *Sox10* or *Nkx6.2*

To investigate astrocyte to OLC conversion, we established postnatal, cortical astrocyte cultures from Ai14 mice. To understand the purity of our cultures, we quantified the numbers of contaminating OLCs. Quantification of SOX10+, O4+ and MBP+ OLCs showed less than 2.5% in our cultures (Supplementary Figure 2A).

To examine the reprogramming potential of *Olig2, Sox10,* and *Nkx6.2*, we transduced Ai14 astrocytes with either LV-GFAP::*Olig2,* LV-GFAP::*Sox10*, LV- GFAP::*Nkx6.2* or a control LV-GFAP::*Cre* (Figure 1A). When we quantified the numbers of tdTomato+ cells that co-expressed

OPC (PDGFRα), COP (O4) or OL (MBP) markers at 8 and 10 DPT, no differences were seen in any of the TF-transduced cultures compared to controls (Figure 1B). By 12 DPT, *Olig2*- cultures showed an increase in the percentage of tdTomato$^+$PDGFRα$^+$ OPCs ($p = 0.0309$, H-statistic: 7.514), whereas *Sox10*- and *Nkx6.2*- cultures showed an increase in tdTomato$^+$MBP$^+$ OLs (*Sox10:* $p = 0.0004$, *Nkx6.2:* $p = 0.0108$, F statistic = 13.28, degrees of freedom =3) (Figure 1B-C). No differences were seen in the percentage of tdTomato$^+$O4$^+$ COPs in any condition at 12DPT (Figure 1B). However, by 14DPT we observed an increase in the percentage of tdTomato$^+$O4$^+$ COPs in *Sox10*- cultures when compared to controls ($p = 0.0213$, F statistic = 3.512, degrees of freedom = 3) (Figure 1B-C). Finally, to understand whether a longer time in culture could increase the number of MBP+ OLs, we cultured cells for an additional 8 days and analyzed the percent of tdTomato$^+$MBP$^+$ cells at 22 DPT. No differences were seen in tdTomato$^+$MBP$^+$ OLs in any condition compared to controls (Supplementary Figure 2B). Altogether, these findings suggest that *Olig2*, *Sox10* and *Nkx6.2* increase different types of iOLCs at different times relative to controls in Ai14 astrocyte cultures.

Canonical OLC cluster found in TF-treated cells following scRNA-seq

To further characterize *Olig2-, Sox10-, Nkx6.2-* and control cultures, we performed scRNA-seq at 14DPT (Figure 1A). Clustering analysis showed the appearance of nine clusters (Figure 1D). We then used proportion analysis to identify clusters that were unique to TF-treated samples. Clusters 3 and 8 were predominantly comprised of cells from TF-induced cultures (Figure 1E-F). The top 10 genes marking each cluster in Figure 1D are highlighted in Figure 1G. Canonical OLC genes such as *Mbp, Plp1, Bcas1* and *Cldn11* were expressed in clusters 3 and 8 (Figure 1G), further suggesting that the increase in iOLCs is TF-dependent.

We then compared the molecular profiles of clusters 3 and 8 to established OL lineage datasets. When using OLC specific annotations of mouse fetal and adult OLCs derived from Marques *et al*[15] to bin the data we found that cluster 3 was characterized by COP signatures, while cluster 8 was characterized by COP, newly formed oligodendrocyte (nfOL) and myelin forming oligodendrocyte (mfOL) signatures (Supplemental Figure 3A). When using OLC annotations of *in vitro* OLCs derived from Dugas *et al*[33] to bin the data, we again saw that cluster 8 was characterized by OL signatures, but cluster 3 was not. (Supplemental Figure 3B). Altogether, these findings further support the observation of increased OLC generation in TF-induced Ai14 cultures (Figure 1B).

Bonafide astrocyte to OLC conversion is confirmed with astrocyte fate mapping.

We had previously observed contaminating OLCs in our cultures and the presence of tdTomato$^+$OLC marker$^+$ cells in our Cre controls. Therefore, to understand the origin of these newly generated OLCs, we performed astrocyte fate mapping using Ai14;Aldh1l1-CreERT2 cultures, the current gold standard in the field[27]. Prior to transduction, post-natal cortical Ai14;Aldh1l1-CreERT2 astrocyte cultures were treated with 1uM 4-OHT to permanently label all *Aldh1l1* expressing cells with tdTomato (Figure 2A). Cultures were transduced with LV-GFAP::*Sox10*-zsGreen, LV-GFAP::*Olig2*-zsGreen, LV-GFAP::*Nkx6.2*-zsGreen or a LV-GFAP::zsGreen control (Figure 2A). Any OLCs that were directly converted by a TF from these *Aldh1l1* expressing astrocytes would therefore express zsGreen and tdTomato (Figure 2A). To understand whether the OPCs generated from *Olig2* overexpression were the product of astrocyte reprogramming, we quantified the percentage of PDGFRα+zsGreen$^+$tdTomato$^+$ cells at 12DPT. We observed an increase in PDGFRα$^+$zsGreen$^+$tdTomato$^+$ OPCs compared to controls,

confirming the astrocytic origin of iOPCs (Figure 2B,E). To assess the origin of *Sox10*-induced COPs we quantified the percentage of O4$^+$zsGreen$^+$tdTomato$^+$ cells at 14DPT. No increase was observed in *Sox10* cultures compared to controls (Figure 2C). Finally, to determine the origin of OLs generated by *Sox10* or *Nkx6.2,* we quantified the percentage of MBP$^+$zsGreen$^+$tdTomato$^+$ OLs at 12DPT. We observed an increase of MBP$^+$zsGreen$^+$tdTomato$^+$ OLs generated by *Sox10*, but not from *Nkx6.2* (Figure D,F). This suggests that *Olig2* and *Sox10* reprogram astrocytes to iOPCs and iOLs, respectively but that the *Nkx6.2*-derived OLs and *Sox10*-derived COPs we previously observed (Figure 1B) were not the result of astrocyte conversion.

To further confirm these findings, we performed a live cell imaging experiment. Ai14;Aldh1l1-CreERT2 *Sox10*-treated and control cultures were imaged from 7DPT to 12DPT then fixed and stained for OLC markers. tdTomato$^+$zsGreen$^+$OLC marker$^+$ cells were then tracked retrospectively to confirm their origin from a Aldh1l1-tdTomato$^+$ cell. tdTomato$^+$zsGreen$^+$OLC marker$^+$ OLs could be tracked back to a tdTomato$^+$ cell with a characteristic astrocyte morphology (Figure 2G, Supplemental Video 1). Taken together, our findings further confirm bonafide astrocyte to iOLC conversion.

We also examined the numbers of Aldh1l1$^+$ non-transduced (tdTomato$^+$ only) cells and Aldh1l1^{neg} virally transduced (zsGreen$^+$ only) cells (Supplementary Figure 4). No difference in the percentage of OLC marker$^+$tdTomato$^+$zsGreen^{neg} was seen between TF-treated and control cultures (Supplementary Figure 4). However, when we examined zsGreen$^+$tdTomato^{neg} cells, we observed an interesting increase in the percentage of MBP$^+$zsGreen$^+$tdTomato^{neg} OLs at 12DPT in *Sox10*-treated cells compared to controls (Supplementary Figure 4C). When we analyzed our live cell imaging experiment to understand the origin of these cells, we found examples of tdTomato^{neg}zsGreen$^+$OLCmarker$^+$ cells that arose from a tdTomato^{neg} cell with an astrocyte morphology (Figure S4D, Supplemental Video 2), suggesting conversion of other astrocyte- like Aldh1l1^{neg} cells.

Characterization of *Sox10*-mediated DLR using scRNA-seq

To better resolve the process of reprogramming, we performed scRNA-sequencing at three time points (d0 prior to transduction, 3DPT, and 8DPT) on Cre-control and *Sox10*-treated cultures. For analysis, we combined cells from D0, 3DPT, 8DPT and 14DPT (Supplemental Figure 5B). We excluded cells with microglia and VLMC markers (Supplemental Figure 5C). Clustering analysis of the remaining cells showed the appearance of 10 distinct clusters (Figure 3A). We characterized these clusters based on expression of genes associated with astrocytes, oligodendrocytes, and NG2 cells, as well as proliferation markers (Figure 3B-C).

First, we examined the molecular identities of cells in our starting cultures. D0 cells were found in clusters 3, 5 and 9 and showed expression of canonical astrocyte markers, including *Slc1a3, Aqp4, Atp1b2, Glul* and *Gfap* (Figure 3B, D). Of interest, clusters 3 and 5 showed expression of *Aldh1l1*, but this gene was absent in cluster 9 (Figure 3B). This discrepancy in *Aldh1l1* expression, may explain the conversion of tdTomato^{neg} cells seen in our fate mapping experiments (Supplemental Figure 4). Altogether, this suggests that early postnatal cultures are heterogeneous and comprised of 3 distinct astrocyte-like populations (clusters).

To understand the cell fate transitions that occur over our reprogramming timecourse, we performed trajectory analysis with Slingshot[35] (Figure 3E) and Monocle3[36–40] (Figure 3F).

Slingshot and Monocle3 both reconstructed trajectories with clusters 6 and 4 as terminal branches (Figure 3E-F). We then examined the origin of cells comprising each cluster. As expected, the clusters formulating earlier roots of the Monocle3 branches (clusters 0, 1, 3) were primarily comprised of D0, 3DPT, and 8DPT timepoints (Figure 3G). Clusters 6 and 4, were predominantly comprised of 14DPT samples, and in particular, cells from *Sox10*-treated samples (D14_S) (Figure 3G). Analysis of the differentially expressed genes (DEGs) at these branches showed that cluster 6 and branch 14 of the Monocle3 trajectory were enriched in the OLC genes *Sox10*[21–24] (p- and q-value=0), *Bcas1*[45] (p- and q-value=0), and *Omg*[46,47] (p- and q-value=0 (Supplementary Figure 6A). In contrast, cluster 4 and the Monocle trajectory branch 11 showed expression of the genes *Lcn2* (p- and q-value=0) and *Igfbp5* (p- and q-value=0), (Supplementary Figure 6 B) previously shown to be expressed in reactive astrocytes[48–50]. All Monocle3-derived differentially expressed genes can be found in Data S2. Altogether, these analyses suggest that the trajectory to branch 14/cluster 6 represents the path of astrocyte to OLC conversion.

To determine how *Sox10* would influence the gene regulatory network of cells in clusters 6 and 4, we performed *in silico* perturbation of *Sox10* using CellOracle[41]. *in silico* knock out (KO) of *Sox10*, predicted a large shift away from a cluster 6 identity and little change away from the identity of cluster 4 (Figure 3H). In agreement, *in silico* knock in (KI) of *Sox10* showed a large shift towards a cluster 6 identify and little change towards a cluster 4 identity (Figure 3I). This suggests that leaf 14 and cluster 6 represent cells with a gene regulatory network most affected by *Sox10*.

As cluster 6 was comprised of both *Sox10*-treated and control cells, we then performed a differential gene expression analysis of 14DPT control and *Sox10*-treated cells within this cluster. 175 DEGs were found, with the top genes (expressed in at least 50% of Cluster 6, *Sox10*-treated cells and with a difference of at least 0.5 pct between control and *Sox10*-treated cultures) clustered and provided in a dotplot (Figure 3J). This analysis highlighted that genes including *Omg*, *Bcas1*, *Pdgfra*, *Cldn11*[51], and *Sox10* are enriched in *Sox10*-treated cells compared to control cells, suggesting that *Sox10*-treated cells are more representative of OLCs than the control cells exposed to OPC and OL media alone.

Understanding the genetic drivers of astrocyte to OLC DLR

Our analysis showed that *Sox10* was important for determining the OLC identity of cluster 6. We specifically chose *Sox10* based on its role in OLC fate specification in development[21,22]. To understand whether there might be other [better] candidate genes that would promote an OLC identity, we used an unbiased, deep learning perturbation model, Fatecode[44], to predict genes that would allow cells to shift from a cluster 4 identity (astrocytes at 14DPT, not fully reprogrammed) to a cluster 6 identity (OLCs at 14DPT, end state of DLR). Following training of the Fatecode model on our *Sox10* and control treated DLR dataset, we identified perturbation on node 16 of the latent layer as one which increased the number of cells in cluster 6 whilst simultaneously reducing the number of cells in clusters 0, 2, 4, 5 and 7. (Figure 3K). Given that the total number of cells in the dataset does not change, this suggests that the perturbation on node 16 of the latent layer was pushing cells that did not reprogram (astrocytes at 8 and 14DPT) towards a reprogrammed (OLCs at 14DPT) fate. To identify the genes driving this shift, we ranked the absolute value of gene expression changes to obtain a list of 3000 genes involved in this perturbation, with the top 40 genes listed in Table S3. Strikingly, we observed that *Sox10* was ranked as the ninth most correlated gene and the top TF driving this shift (Table S3). This supports our previous findings that suggest *Sox10* drives astrocyte to OLC DLR. Additionally, we observed

genes important for myelination (*Plp1, Mbp*)[52], as well as other TFs previously identified with OLC differentiation and myelination (*Sox3*[53], *Klf9*[54]) as integral to this shift. Taken together, these findings further support the importance of *Sox10* in reprogramming astrocytes to OLCs and identify additional genes and TFs that may be ideal candidates for astrocyte to OLC reprogramming.

DISCUSSION

Here, we show astrocyte to OLC conversion with *Sox10* or *Olig2* using a battery of experimental tools, including astrocyte fate mapping, live cell imaging, a scRNA-seq timecourse and unbiased deep learning. While previous studies demonstrated the generation of iOLCs from different types of somatic cells using combinations of *Sox10*, *Olig2* and *Nkx6.2*[2,3,55], ours is the first to compare the individual reprogramming ability of each these TFs in astrocytes. Using Aldh1l1-based fate-mapping, we found that that *Sox10* converts Aldh1l1$^+$ astrocytes to MBP$^+$ oligodendrocytes, whereas *Olig2* converts Aldh1l1$^+$ astrocytes to PDGFRα^+ OPCs.

In contrast to *Sox10* and *Olig2*, *Nkx6.2* was unable to convert astrocytes to new OLCs. Unlike *Sox10* and *Olig2*, which define oligodendrocyte identity and continue to shape their gene regulatory network throughout life[56], *Nkx6.2* may be unable to direct a stable OLC identity. In support of this, previous work investigating pericyte to OL reprogramming found that inclusion of *Nkx6.2* in their reprogramming cocktail was refractory to reprogramming. In addition to observing the activation of genes unrelated to oligodendrogenesis, genes required for OPC identity were downregulated when cells were transduced with *Nkx6.2*[3].

In a previous study, Khanghahi et al. investigated *Sox10* mediated astrocyte to OLC conversion. Ectopic expression of *Sox10* in astrocytes *in vitro* led to an increase in OPCs at 21DPT[14], a different OLC type and longer time to conversion than we observed in our study. Although similar experimental designs were used in both studies (cortical P3-P5 astrocytes cultured in OPC media in Khanghahi *et al*. versus P0-P5 cortical astrocytes cultured in OPC followed by OL media in our study), these discrepancies may be due to the use of different viral delivery strategies and metrics of reprogramming. In our study, we used lentiviral delivery of *Sox10*-P2A-zsGreen under the control of the long (2178bp) hGFAP promoter[57]. In addition, we used Ai14;Aldh1l1-CreERT2 mice and quantified only the iOLCs with an astrocyte origin (based on tdTomato$^+$OLC marker$^+$expression). In contrast, in the Khanghahi *et al.* study, the authors used a SFFV promoter to deliver *Sox10*-IRES-GFP and reported the number of GFP+ iOLCs in their *Sox10*-transduced cultures. Without an astrocyte specific promoter and stringent lineage tracking it is difficult to conclude that their iOPCs were the result of astrocyte conversion. The non-specific delivery strategy could have instead hit a contaminating, perhaps more distantly related cell that would need more time to convert.

In this regard, a recent study suggested that the conversion reported in studies of *in vivo* astrocyte to neuron DLR[6,58,59] were not true reprogramming, but rather the result of erroneous labelling of endogenous neurons due to technical confounds[27]. As a result, the DLR community has advocated for the stringent validation of DLR paradigms. In this study, we first examined astrocyte to iOLC conversion in Ai14 cells. This enabled the permanent labeling of transduced cells and therefore, the tracking of those cells through the DLR timecourse. To then validate our findings, we used Aldh1l1-astrocyte fate mapping, which highlighted a lack of conversion with *Nkx6.2.*

When performing our lineage tracing experiments, we also discovered a subpopulation of cells (Aldh1l1-tdTomato^{neg}zsGreen^{+}OLCmarker^{+}) that was converted by *Sox10* at a high reprogramming efficiency (Supplemental Figure 4C). scRNA-seq analysis of cells prior to conversion also showed a cluster with low *Aldh1l1* expression compared to the other astrocyte clusters, but with similar levels of *Gfap* expression (Figure 3B), which could explain its transduction by the LV-GFAP::*Sox10*. These *Aldh1l1*^{lo}*Gfap*^{hi} cells were predominantly characterized by mature astrocyte markers but also showed expression of genes found in NG2 glia (*Cspg, Pdgfra*) (Figure 3B-C). Curiously, this population did not show expression of proliferation marker *Mki67* (Figure 3C), in contrast to previous studies showing that astrocyte-like NG2 glia are proliferative[60,61]. Of interest, one study reported high co-expression of *Gfap Cspg4* and *Pdgfra* in early postnatal astrocytes[62]. This leads to the intriguing idea that these cells represent a population of Aldh1l1^{lo} expressing astrocytes, which could preclude our understanding of the true extent of astrocyte reprogramming possible when using Aldh1l1-fate-mapping. The presence of these *Aldh11l*^{lo} cells shows that astrocyte cultures are heterogeneous and that different types of astrocytes or astrocyte-like cells may be suitable targets for DLR. Further studies using scRNA barcoding and CellOracle to profile the different types of astrocytes that are particularly amenable to reprogramming will be beneficial in designing more specific and tailored DLR strategies.

Conversion of Aldh1l1^{pos} astrocytes to PDGFRα^{+}tdTomato^{+}zsGreen^{+} iOPCs using *Olig2* occurred at a relatively low rate, with an average conversion efficiency of 16.75% cells. This conversion was even lower for the generation of mature MBP^{+}tdTomato^{+}zsGreen^{+} OLs, with an average conversion efficiency of 2.83%. This may suggest that although *Sox10* generates more mature OLCs, it can only do this in a select number of ‘elite’ donor cells. Alternatively, the absence of a substrate to myelinate in our cultures may preclude the true reprogramming ability of *Sox10.* OL survival *in vitro* and *in vivo* has been shown to be dependent on the presence of axons[63]. Furthermore, previous studies have shown that mRNA expression of myelin genes is increased when OLs are cultured in the presence of neurons[64], and differentiation of OPCs can be induced with bead or nanofiber scaffolding[65,66]. However, additional factors, or the discovery of “better” TFs may also allow for increased reprogramming efficiency. It remains to be determined whether the genes predicted by Fatecode in this study will improve the efficiencies of OLC generation. Nonetheless, it is important to note that the extent of iOLC generation must be balanced with the extent of astrocyte loss as many studies have shown that widespread loss of astrocytes can be detrimental[67,68].

Investigation of the reprogramming trajectory of *Sox10* and control treated cells showed a bifurcation, where the cells either converted to iOLCs or remained as astrocytes (Figure 3F). It remains to be determined if this is a genetic hurdle or a metabolic hurdle, similar to studies of astrocyte to neuron conversion[69]. Future studies investigating the genes that characterize the ‘tipping point’ of conversion will be useful in identifying the mechanisms of astrocyte to iOLC reprogramming. A clearer understanding of how different astrocyte types convert to iOLCs and the mechanisms that underly the reprogramming processes will help us identify the best DLR paradigm for therapeutic applications.

ACKNOWLEDGMENTS
The authors are grateful to Dr. Lindsey Fiddes at the Microscope Imaging Facility for

assistance with live cell imaging. Next-generation sequencing was performed at the Donnelly Sequencing Facility. J.B. was supported by the Baden Havard Endowment Fund, the Ontario Graduate Scholarship and the MITACs Accelerate Fellowship. This work was supported by grants to M.F. from CIHR (PJT-175254), Medicine by Design (Cycle 2 Funding), Temerty Foundation (Pathway Grant), Ontario Institute of Regenerative Medicine (Kickstart Innovation Investment Program), the Stem Cell Network (Early Career Researcher Jump Start Award), the Connaught Foundation (Innovation Grant), CFA-JELF program for microscopy infrastructure, and to S.A.Y. from the CFI-JELF program for scRNA-seq related infrastructure.

DATA AVAILABILITY

Please contact the authors upon a decision to send for review for public release of our datasets. Raw fastq files will be provided with accession numbers have been deposited into NCBI's short read archives. Genetables are provided as a Cyvers link through the Github pages (see methods) housing the scripts used in the manuscript.

AUTHOR CONTRIBUTIONS

Conceptualization, M.F.; Methodology, J.B., G.D.B., M.F.; Software, E.Y.S., M.S.; Formal Analysis, J.B., E.Y.S., M.S., S.A.Y., M.F.; Investigation, J.B., E.Y.S., A.O., M.S., K.L., M. Fahim, H.T.T., D.L.C., A.D., and S.A.Y.; Writing – Original Draft, J.B. and E.Y.S.; Writing – Review and Editing, M.F.; Visualization, J.B., E.Y.S., M.S.; Supervision, G.D.B., M.F.; Funding Acquisition, S.A.Y., M.F.

COMPETING INTERESTS

MF, JB are co-inventors on a patent application (Applicants: Justine Bajohr, Maryam Faiz, The governing council of the University of Toronto; Inventors: Maryam Faiz, Justine Bajohr, US Provisional Filed #63/497,357, April 20, 2023).

FIGURE LEGENDS

Figure 1. *Sox10*, *Olig2* and *Nkx6.2* convert GFAP+ cells to oligodendrocyte lineage cells. **(A)** Experimental design and timeline. **(B)** Quantification of PDGFRα$^+$tdTomato$^+$ iOPCs, O4$^+$tdTomato$^+$ iCOPs and MBP$^+$tdTomato$^+$ iOLs at 8 (n=5), 10 (n=4), 12 (n=5) and 14 (n=7) days post transduction (DPT). Data are presented as mean ± SEM, each data point represents one individual cell culture experiment; at each time point and for each cell type marker, a matched pairs one-way ANOVA or Kruskal-Wallis (d8, d10, d12) or one- way ANOVA (d14) was performed with Dunnet's post testing (*= $p<0.05$, *** = $p< 0.001$). **(C)** Representative images of PDGFRα$^+$tdTomato$^+$,O4$^+$tdTomato$^+$ and MBP$^+$ tdTomato$^+$ cells at 12DPT. Single channel images are shown of the boxed cells (arrows indicate double positive cells). **(D)** UMAP clustering of *Olig2*-, *Sox10*-, *Nkx6.2*- and control (Cre) transduced cells at 14DPT **(E)** UMAP of (D) overlayed with treatment (*Olig2*-, *Sox10*-, *Nkx6.2*- and control (Cre)) **(F)** Proportion analysis of clusters found in *Olig2*-, *Sox10*-, *Nkx6.2*-, and control (Cre) transduced cells. **(G)** Heatmap of top upregulated genes from each cluster in (D).

Figure 2. Lineage tracing confirms true conversion of astrocytes to oligodendrocyte lineage cells.
(A) Experimental design, timeline and outcomes. **(B)** Quantification of tdTomato$^+$zsGreen$^+$PDGFRα$^+$ iOPCs at 12DPT (n= 3). Data are presented as mean ± SEM, each data point represents one individual cell culture, a paired t-test was performed (** = $p<0.01$). **(C)** Quantification of tdTomato$^+$zsGreen$^+$O4$^+$ iCOPs at 14DPT (n = 4). Data are presented as mean ±

SEM, each data point represents one individual cell culture, a Wilcoxon test was performed (ns). **(D)** Quantification of tdTomato$^+$zsGreen$^+$MBP$^+$ iOLs at 12DPT (n= 4 for *Sox10, Nkx6.2*, n= 3 for *Cre*). Data are presented as mean ± SEM, each data point represents one individual cell culture, a one-way ANOVA with Dunnet's post testing was performed (* = $p<0.05$). **(E)** Representative image of PDGFRα$^+$tdTomato$^+$zsGreen$^+$ cells 12DPT. Single channel images are shown of the boxed cells (arrows indicate triple positive cells, scale bar = 50um (merge) and 20um (single channel)). **(F)** Representative image of MBP$^+$tdTomato$^+$zsGreen$^+$ cells 12DPT. Single channel images are shown of the boxed cells (arrows indicate triple positive cells, scale bar = 50um (merge) and 20um (single channel)). **(G)** Representative tdTomato$^+$zsGreen$^+$ cell with astrocyte-like morphology at onset (7DPT) and OLC expression at the end (12DPT) of live cell tracking (arrow indicates tracked cell, scale bar = 50um).

Figure 3. Characterization of DLR using scRNA-seq shows terminal oligodendrocyte cluster of cells at day 14 driven by *Sox10*. (A) UMAP clustering of *Sox10* and *Cre* control treated cells from prior to transduction, as well as 3, 8 and 14DPT **(B)** Canonical astrocyte gene expression (log-normalized, y-axis) separated by cluster (x-axis) and coloured by timepoint and treatment group. **(C)** Canonical gene expression (log-normalized, y-axis) of NG2 glia, VLMC, OLC, microglia and proliferation, separated by cluster (x-axis) and coloured by timepoint and treatment group. **(D)** UMAP clustering from (A) overlayed with timepoint and treatment group **(E)** Slingshot lineage analysis of *Sox10* and *Cre* control treated clusters overlaid with UMAP embeddings. **(F)** Monocle3 lineage analysis of *Sox10* and *Cre* control treated clusters overlaid on UMAP plot and coloured by pseudotime predictions. **(G)** The number of cells (y-axis) per cluster (x-axis) originating from each of the coloured timepoint+treatment groups. **(H)** CellOracle modeling of *in silico Sox10* knock out (KO) overlaid onto UMAP plot. Arrows indicate trajectory prediction with *Sox10* KO. **(I)** Cell Oracle modeling of *in silico Sox10* knock in (KI) overlaid onto UMAP plot. Arrows indicate trajectory prediction with *Sox10* KI **(J)** Clustered, top differentially expressed genes (dots, y-axis) between 14DPT *Sox10 (D14_S,* Beige, left*)* and *Cre (D14_C,* Blue, right*)* control treated cells from cluster 6. Size of dot is scaled to percent of cells in the cluster expressing that gene, colour of dot represents the average scaled expression of the gene across cells. **(K)** Predicted cell types following Fatecode perturbation on node 16 of the latent layer in the *Sox10* and control treated dataset. Blue represents the number of cells prior to perturbation. Orange (overlayed) represents the number of cells following perturbation.

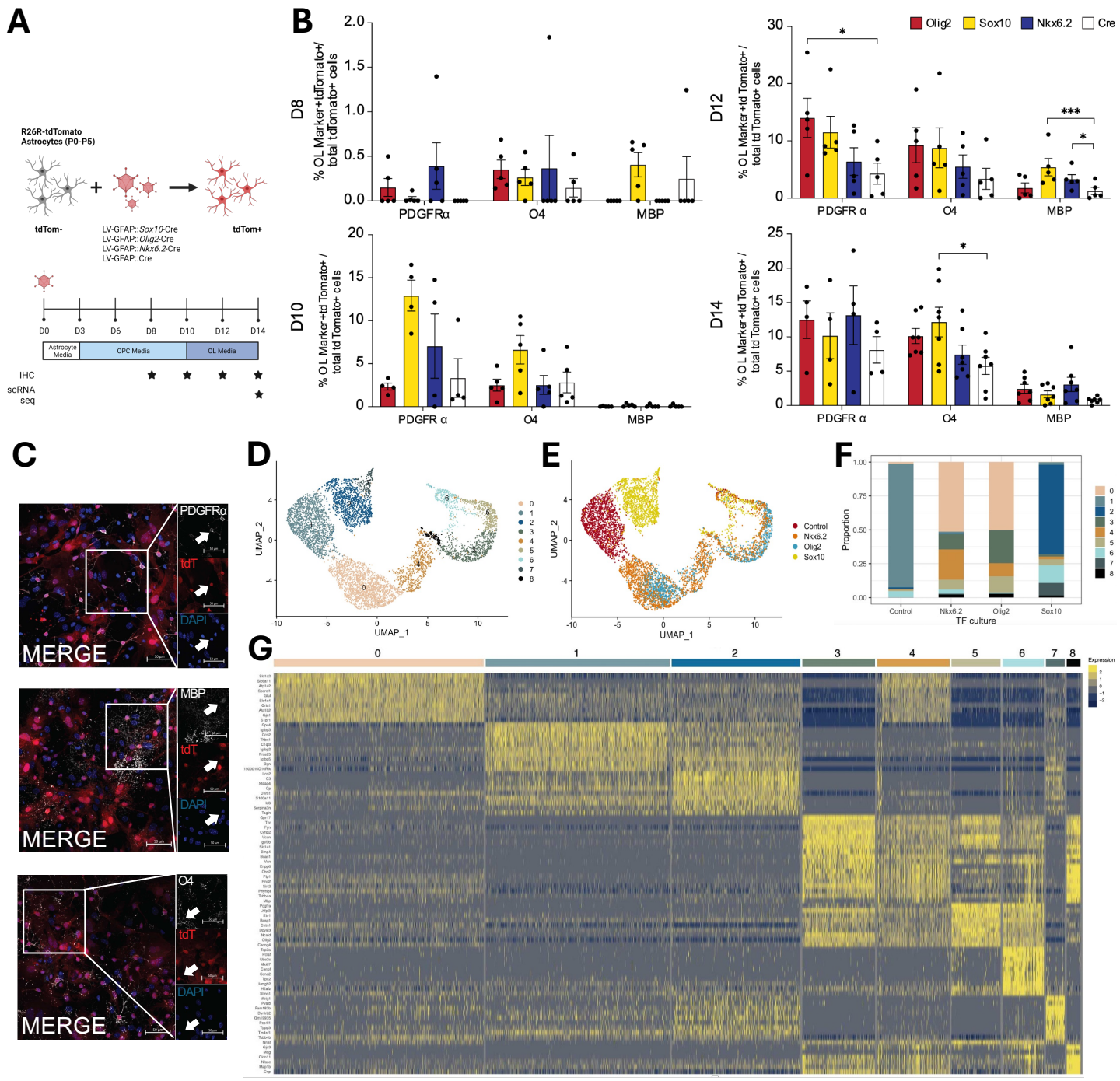

A
R26R-tdTomato Astrocytes (P0-P5)
tdTom-
LV-GFAP::Sox10-Cre
LV-GFAP::Olig2-Cre
LV-GFAP::Nkx6.2-Cre
LV-GFAP::Cre
tdTom+
D0
D3
D6
D8
D10
D12
D14
Astrocyte Media
OPC Media
OL Media
IHC
scRNA seq
B
D8
D10
D12
D14
% OL Marker+tdTomato+/ total tdTomato+ cells
% OL Marker+td Tomato+ / total td Tomato+ cells
PDGFRα
O4
MBP
Olig2
Sox10
Nkx6.2
Cre
C
PDGFRα
tdT
DAPI
MERGE
MBP
O4
D
UMAP_1
UMAP_2
E
Control
Nkx6.2
Olig2
Sox10
F
Proportion
TF culture
G
Expression

Figure 1

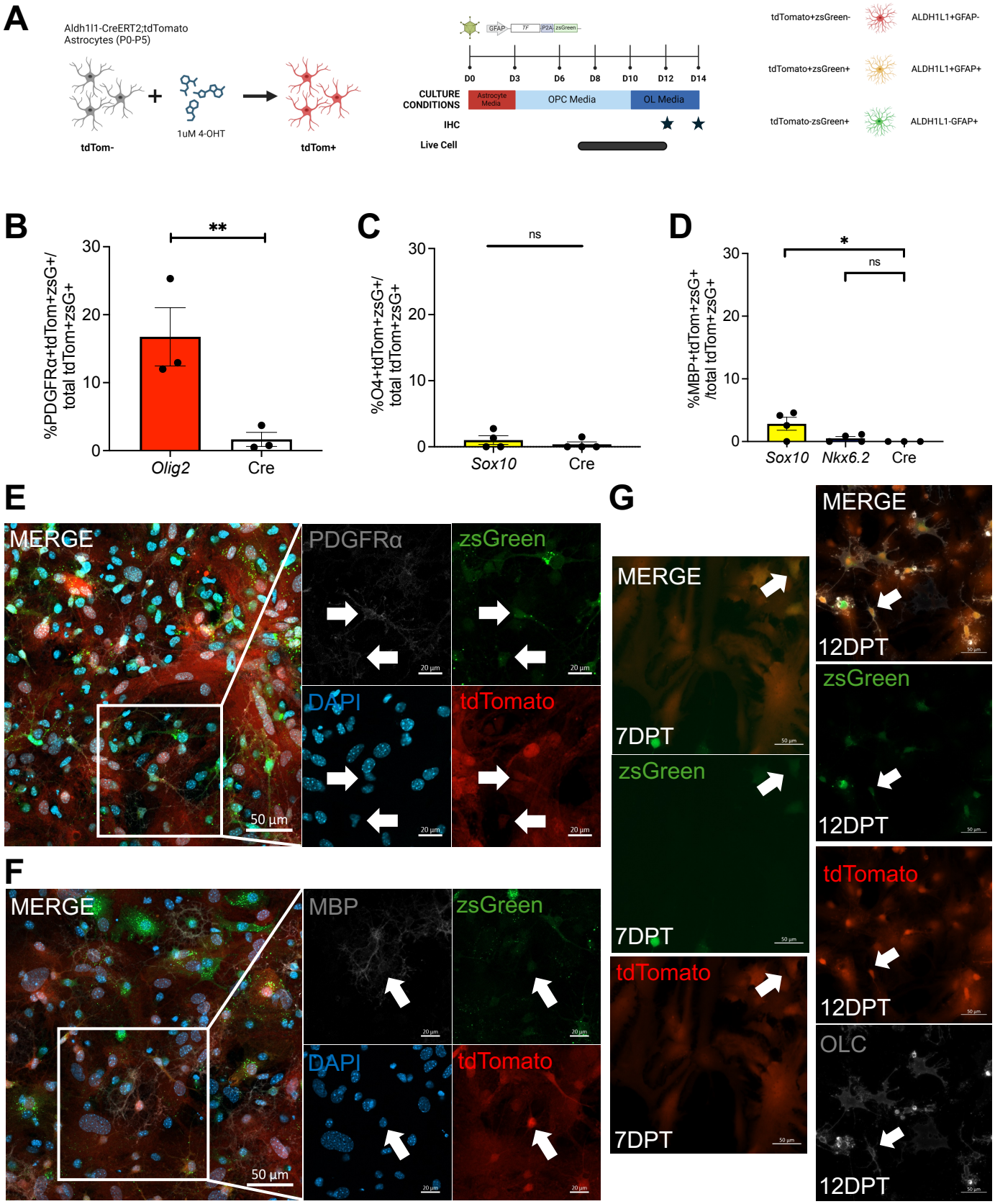

A
Aldh1l1-CreERT2;tdTomato
Astrocytes (P0-P5)
tdTom-
1uM 4-OHT
tdTom+
GFAP
TF
P2A
zsGreen
D0
D3
D6
D8
D10
D12
D14
CULTURE CONDITIONS
Astrocyte Media
OPC Media
OL Media
IHC
Live Cell
tdTomato+zsGreen-
ALDH1L1+GFAP-
tdTomato+zsGreen+
ALDH1L1+GFAP+
tdTomato-zsGreen+
ALDH1L1-GFAP+
B
%PDGFRα+tdTom+zsG+/ total tdTom+zsG+
**
Olig2
Cre
C
%O4+tdTom+zsG+/ total tdTom+zsG+
ns
Sox10
Cre
D
%MBP+tdTom+zsG+ /total tdTom+zsG+
*
ns
Sox10
Nkx6.2
Cre
E
MERGE
PDGFRα
zsGreen
DAPI
tdTomato
50 µm
20 µm
F
MERGE
MBP
zsGreen
DAPI
tdTomato
50 µm
20 µm
G
MERGE
7DPT
zsGreen
7DPT
tdTomato
7DPT
MERGE
12DPT
zsGreen
12DPT
tdTomato
12DPT
OLC
12DPT
50 µm

Figure 2

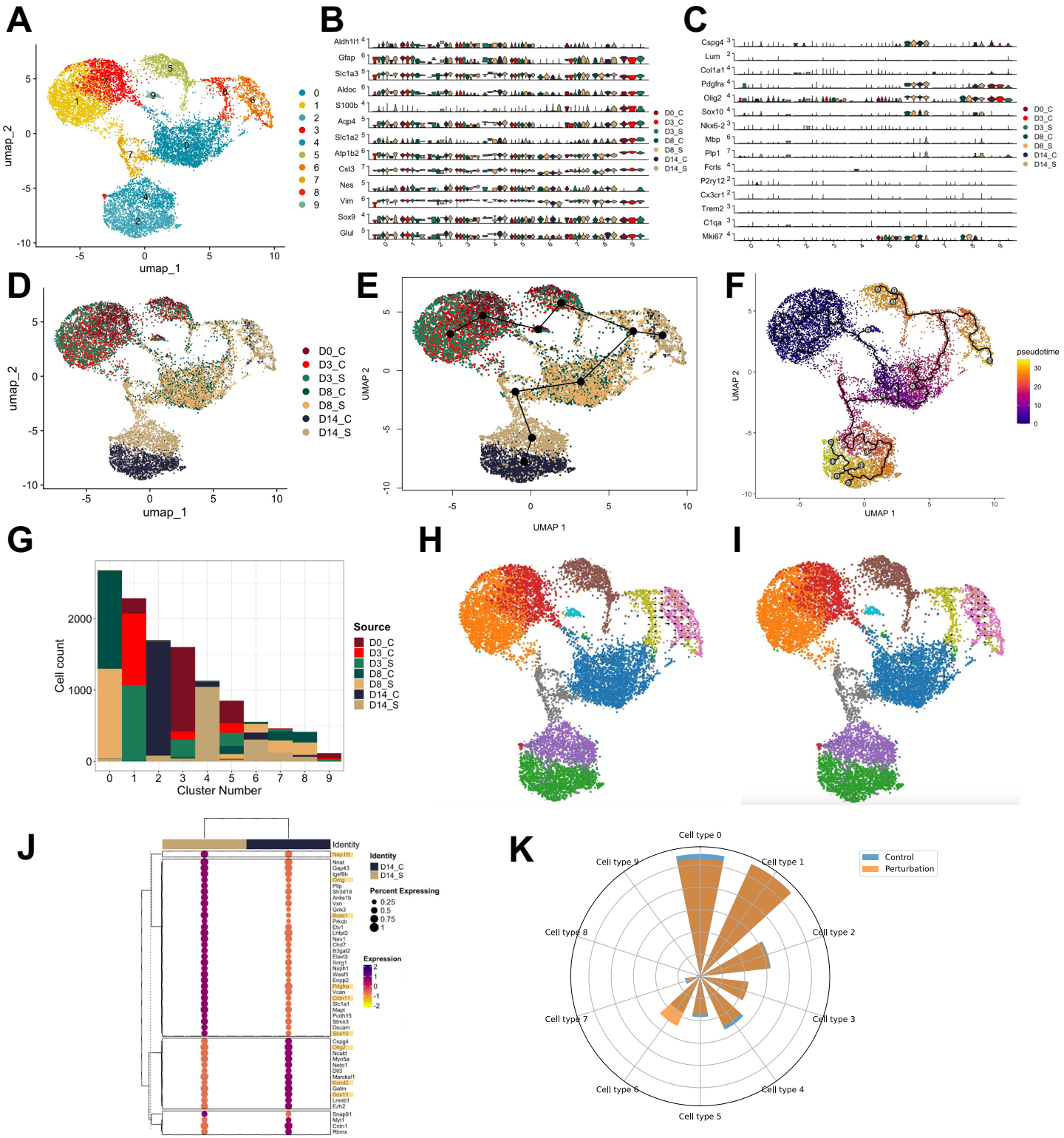

A
umap_1
umap_2
0
1
2
3
4
5
6
7
8
9
B
Aldh1l1
Gfap
Slc1a3
Aldoc
S100b
Aqp4
Slc1a2
Atp1b2
Cst3
Nes
Vim
Sox9
Glul
D0_C
D3_C
D3_S
D8_C
D8_S
D14_C
D14_S
C
Cspg4
Lum
Col1a1
Pdgfra
Olig2
Sox10
Nkx6-2
Mbp
Plp1
Fcrls
P2ry12
Cx3cr1
Trem2
C1qa
Mki67
D
E
UMAP 1
UMAP 2
F
pseudotime
G
Cell count
Cluster Number
Source
H
I
J
Identity
Percent Expressing
Expression
K
Cell type 0
Cell type 1
Cell type 2
Cell type 3
Cell type 4
Cell type 5
Cell type 6
Cell type 7
Cell type 8
Cell type 9
Control
Perturbation

Figure 3

REFERENCES

1. Bajohr, J. & Faiz, M. Direct Lineage Reprogramming in the CNS. in *Cell Biology and Translational Medicine, Volume 6: Stem Cells: Their Heterogeneity, Niche and Regenerative Potential* (ed. Turksen, K.) 31–48 (Springer International Publishing, Cham, 2020). doi:10.1007/5584_2019_374.
2. Najm, F. J. *et al.* Transcription factor–mediated reprogramming of fibroblasts to expandable, myelinogenic oligodendrocyte progenitor cells. *Nat Biotechnol* **31**, 426–433 (2013).
3. Kim, K.-P. *et al.* Donor cell memory confers a metastable state of directly converted cells. *Cell Stem Cell* **28**, 1291-1306.e10 (2021).
4. Heinrich, C. *et al.* Directing astroglia from the cerebral cortex into subtype specific functional neurons. *PLoS Biol* **8**, e1000373 (2010).
5. Heins, N. *et al.* Glial cells generate neurons: the role of the transcription factor Pax6. *Nat Neurosci* **5**, 308–315 (2002).
6. Guo, Z. *et al.* In vivo direct reprogramming of reactive glial cells into functional neurons after brain injury and in an Alzheimer's disease model. *Cell Stem Cell* **14**, 188–202 (2014).
7. Liu, Y. *et al.* Ascl1 Converts Dorsal Midbrain Astrocytes into Functional Neurons In Vivo. *J Neurosci* **35**, 9336–9355 (2015).
8. Levison, S. W. & Goldman, J. E. Both oligodendrocytes and astrocytes develop from progenitors in the subventricular zone of postnatal rat forebrain. *Neuron* **10**, 201–212 (1993).
9. Kriegstein, A. & Alvarez-Buylla, A. The glial nature of embryonic and adult neural stem cells. *Annu Rev Neurosci* **32**, 149–184 (2009).
10. Masserdotti, G., Gascón, S. & Götz, M. Direct neuronal reprogramming: learning from and for development. *Development* **143**, 2494–2510 (2016).
11. Hörmanseder, E. Epigenetic memory in reprogramming. *Current Opinion in Genetics & Development* **70**, 24–31 (2021).

12. Bussmann, L. H. *et al.* A Robust and Highly Efficient Immune Cell Reprogramming System. *Cell Stem Cell* **5**, 554–566 (2009).
13. Xie, H., Ye, M., Feng, R. & Graf, T. Stepwise reprogramming of B cells into macrophages. *Cell* **117**, 663–676 (2004).
14. Mokhtarzadeh Khanghahi, A., Satarian, L., Deng, W., Baharvand, H. & Javan, M. In vivo conversion of astrocytes into oligodendrocyte lineage cells with transcription factor Sox10; Promise for myelin repair in multiple sclerosis. *PLoS One* **13**, e0203785 (2018).
15. Marques, S. *et al.* Oligodendrocyte heterogeneity in the mouse juvenile and adult central nervous system. *Science* **352**, 1326–1329 (2016).
16. Zhou, Q., Wang, S. & Anderson, D. J. Identification of a novel family of oligodendrocyte lineage-specific basic helix-loop-helix transcription factors. *Neuron* **25**, 331–343 (2000).
17. Wang, H. *et al.* Region-specific distribution of Olig2-expressing astrocytes in adult mouse brain and spinal cord. *Molecular Brain* **14**, 36 (2021).
18. Buffo, A. *et al.* Expression pattern of the transcription factor Olig2 in response to brain injuries: Implications for neuronal repair. *Proc Natl Acad Sci U S A* **102**, 18183–18188 (2005).
19. 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).
20. Tatsumi, K. *et al.* Olig2-Lineage Astrocytes: A Distinct Subtype of Astrocytes That Differs from GFAP Astrocytes. *Front Neuroanat* **12**, 8 (2018).
21. Liu, Z. *et al.* Induction of oligodendrocyte differentiation by Olig2 and Sox10: evidence for reciprocal interactions and dosage-dependent mechanisms. *Dev Biol* **302**, 683–693 (2007).
22. Pozniak, C. D. *et al.* Sox10 directs neural stem cells toward the oligodendrocyte lineage by decreasing Suppressor of Fused expression. *Proc Natl Acad Sci U S A* **107**, 21795–21800 (2010).

23. Takada, N., Kucenas, S. & Appel, B. Sox10 is necessary for oligodendrocyte survival following axon wrapping. *Glia* **58**, 996–1006 (2010).
24. Stolt, C. C. *et al.* Terminal differentiation of myelin-forming oligodendrocytes depends on the transcription factor Sox10. *Genes Dev* **16**, 165–170 (2002).
25. Awatramani, R. *et al.* Evidence That the Homeodomain Protein Gtx Is Involved in the Regulation of Oligodendrocyte Myelination. *J Neurosci* **17**, 6657–6668 (1997).
26. Cai, J. *et al.* Co-localization of Nkx6.2 and Nkx2.2 Homeodomain Proteins in Differentiated Myelinating Oligodendrocytes. *Glia* **58**, 458–468 (2010).
27. Wang, L.-L. *et al.* Revisiting astrocyte to neuron conversion with lineage tracing in vivo. *Cell* **184**, 5465-5481.e16 (2021).
28. McCarthy, K. D. & de Vellis, J. Preparation of separate astroglial and oligodendroglial cell cultures from rat cerebral tissue. *J Cell Biol* **85**, 890–902 (1980).
29. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* **34**, 525–527 (2016).
30. Melsted, P. *et al.* Modular, efficient and constant-memory single-cell RNA-seq preprocessing. *Nat Biotechnol* **39**, 813–818 (2021).
31. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587.e29 (2021).
32. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016).
33. Dugas, J. C., Tai, Y. C., Speed, T. P., Ngai, J. & Barres, B. A. Functional genomic analysis of oligodendrocyte differentiation. *J Neurosci* **26**, 10967–10983 (2006).
34. Marsh, S. scCustomize:Custom Visualizations & Functions for Streamlined Analyses of Single Cell Sequencing. (2021).
35. Street, K. *et al.* Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* **19**, 477 (2018).

36. Trapnell, C. *et al.* The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* **32**, 381–386 (2014).
37. Qiu, X. *et al.* Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods* **14**, 979–982 (2017).
38. Cao, J. *et al.* The single cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496–502 (2019).
39. Traag, V. A., Waltman, L. & van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Sci Rep* **9**, 5233 (2019).
40. Levine, J. H. *et al.* Data-driven phenotypic dissection of AML reveals progenitor-like cells that correlate with prognosis. *Cell* **162**, 184–197 (2015).
41. Kamimoto, K. *et al.* Dissecting cell identity via network inference and in silico gene perturbation. *Nature* **614**, 742–751 (2023).
42. Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis*. (Springer, New York, NY, 2009). doi:10.1007/978-0-387-98141-3.
43. RStudio: Integrated Development for R. (2020).
44. Sadria, M., Layton, A., Goyal, S. & Bader, G. D. Fatecode: Cell fate regulator prediction using classification autoencoder perturbation. 2022.12.16.520772 Preprint at https://doi.org/10.1101/2022.12.16.520772 (2022).
45. Fard, M. K. *et al.* BCAS1 expression defines a population of early myelinating oligodendrocytes in multiple sclerosis lesions. *Sci Transl Med* **9**, eaam7816 (2017).
46. Vourc'h, P. & Andres, C. Oligodendrocyte myelin glycoprotein (OMgp): evolution, structure and function. *Brain Res Brain Res Rev* **45**, 115–124 (2004).
47. Mikol, D. D. & Stefansson, K. A phosphatidylinositol-linked peanut agglutinin-binding glycoprotein in central nervous system myelin and on oligodendrocytes. *J Cell Biol* **106**, 1273–1279 (1988).

48. Bi, F. *et al.* Reactive astrocytes secrete lcn2 to promote neuron death. *Proceedings of the National Academy of Sciences* **110**, 4069–4074 (2013).

49. Liddelow, S. A. *et al.* Neurotoxic reactive astrocytes are induced by activated microglia. *Nature* **541**, 481–487 (2017).

50. Lee, W.-H., Wang, G.-M., Seaman, L. B. & Vannucci, S. J. Coordinate IGF-I and IGFBP5 Gene Expression in Perinatal Rat Brain after Hypoxia-Ischemia. *J Cereb Blood Flow Metab* **16**, 227–236 (1996).

51. Morita, K., Sasaki, H., Fujimoto, K., Furuse, M. & Tsukita, S. Claudin-11/OSP-based Tight Junctions of Myelin Sheaths in Brain and Sertoli Cells in Testis. *J Cell Biol* **145**, 579–588 (1999).

52. Kister, A. & Kister, I. Overview of myelin, major myelin lipids, and myelin-associated proteins. *Front. Chem.* **10**, (2023).

53. Hoffmann, S. A. *et al.* Stem cell factor Sox2 and its close relative Sox3 have differentiation functions in oligodendrocytes. *Development* **141**, 39–50 (2014).

54. Dugas, J. C., Ibrahim, A. & Barres, B. A. The T3-induced gene KLF9 regulates oligodendrocyte differentiation and myelin regeneration. *Mol Cell Neurosci* **50**, 45–57 (2012).

55. Yang, N. *et al.* Generation of oligodendroglial cells by direct lineage conversion. *Nat Biotechnol* **31**, 434–439 (2013).

56. Sock, E. & Wegner, M. Using the lineage determinants Olig2 and Sox10 to explore transcriptional regulation of oligodendrocyte development. *Dev Neurobiol* **81**, 892–901 (2021).

57. Lee, Y., Messing, A., Su, M. & Brenner, M. GFAP promoter elements required for region-specific and astrocyte-specific expression. *Glia* **56**, 481–493 (2008).

58. Chen, Y.-C. *et al.* A NeuroD1 AAV-Based Gene Therapy for Functional Brain Repair after Ischemic Injury through In Vivo Astrocyte-to-Neuron Conversion. *Molecular Therapy* **28**, 217–234 (2020).
59. Qian, H. *et al.* Reversing a model of Parkinson's disease with in situ converted nigral neurons. *Nature* **582**, 550–556 (2020).
60. Janeckova, L. *et al.* Astrocyte-like subpopulation of NG2 glia in the adult mouse cortex exhibits characteristics of neural progenitor cells. *Glia* **72**, 245–273 (2024).
61. Kirdajova, D. *et al.* Transient astrocyte-like NG2 glia subpopulation emerges solely following permanent brain ischemia. *Glia* **69**, 2658–2681 (2021).
62. Rusnakova, V. *et al.* Heterogeneity of Astrocytes: From Development to Injury – Single Cell Gene Expression. *PLOS ONE* **8**, e69734 (2013).
63. Barres, B. A., Jacobson, M. D., Schmid, R., Sendtner, M. & Raff, M. C. Does oligodendrocyte survival depend on axons? *Current Biology* **3**, 489–497 (1993).
64. Macklin, W. b., Weill, C. I. & Deininger, P. I. Expression of myelin proteolipid and basic protein mRNAS in cultured cells. *Journal of Neuroscience Research* **16**, 203–217 (1986).
65. Rosenberg, S. S., Kelland, E. E., Tokar, E., De La Torre, A. R. & Chan, J. R. The geometric and spatial constraints of the microenvironment induce oligodendrocyte differentiation. *Proc Natl Acad Sci U S A* **105**, 14662–14667 (2008).
66. Lee, S. *et al.* A culture system to study oligodendrocyte myelination-processes using engineered nanofibers. *Nat Methods* **9**, 917–922 (2012).
67. Bush, T. G. *et al.* Leukocyte infiltration, neuronal degeneration, and neurite outgrowth after ablation of scar-forming, reactive astrocytes in adult transgenic mice. *Neuron* **23**, 297–308 (1999).
68. Davis, N. *et al.* Pharmacological ablation of astrocytes reduces Aβ degradation and synaptic connectivity in an ex vivo model of Alzheimer's disease. *J Neuroinflammation* **18**, 73 (2021).

69. Gascón, S., Masserdotti, G., Russo, G. L. & Götz, M. Direct Neuronal Reprogramming: Achievements, Hurdles, and New Roads to Success. *Cell Stem Cell* **21**, 18–34 (2017).

SUPPLEMENTARY INFORMATION

Supplementary Figure 1. 14DPT scRNA-seq processing for initial UMAP generation. (A) Distribution of genes (nFeature), reads (nCount) and percent of mitochondrial reads (percent.mt) per sample. **(B)** UMAP clustering prior to removal of microglia cluster. **(C)** Top upregulated genes characterizing each cluster observed in (B).

Supplementary Figure 2. Characterization of DLR cultures. (A) Quantification of contaminating iOLCs in post-natal astrocyte cultures 3DPT. **(B)** Quantification of MBP expression in tdTomato$^+$ transduced cells 22DPT (n=3, data are presented as mean ± SEM, each data point represents one individual cell culture, a matched pairs one way ANOVA was performed (ns)).

Supplementary Figure 3. Comparison of DLR clusters 14DPT to established datasets of OLCs. (A) Violin plots of cell types annotated by Marques *et al.* 2016 within each cluster (OPC = oligodendrocyte progenitor cells, COPs = committed oligodendrocyte progenitors, NFOL = newly formed oligodendrocyte, MFOL = myelin forming oligodendrocyte, mOL = mature oligodendrocyte).**(B)** Violin plots of cell types annotated by Dugas *et al.* 2006 within each cluster (OL = oligodendrocyte, OPC = oligodendrocyte progenitor cells).

Supplemental Figure 4. (A) Quantification of (i) tdTomato$^+$ and (ii) zsGreen$^+$ PDGFRα$^+$ iOPCs at 12DPT (n= 3). Data are presented as mean ± SEM, each data point represents one individual cell culture, a paired t-test for (i) and a Wilcoxon test for (ii) was performed (** = $p<0.01$). **(B)** Quantification of (i) tdTomato$^+$ and (ii) zsGreen$^+$ O4$^+$ iCOPs at 14DPT (n = 4). Data are presented as mean ± SEM, each data point represents one individual cell culture, a paired t-test was used. **(C)** Quantification of (i) tdTomato$^+$ (ii) zsGreen$^+$ MBP$^+$ OLs at 12DPT (n= 3). Data are presented as mean ± SEM, each data point represents one individual cell culture, a matched pairs one way ANOVA with Geisser –Greenhouse correction and Dunnet's post testing was used (* = $p<0.05$). **(D)** Representative tdTomato^{neg}zsGreen$^+$ cell with astrocyte-like morphology at onset (7DPT) and OLC expression at the end (12DPT) of live cell tracking (arrow indicates tracked cell, scale bar = 50um).

Supplementary Figure 5. DLR timecourse scRNA-seq processing for initial UMAP generation. (A) Distribution of genes (nFeature), reads (nCount) and percent of mitochondrial reads
(percent.mt) per sample. **(B)** UMAP clustering prior to removal of microglia and VLMC clusters. **(C)** Top upregulated genes characterizing each cluster observed in (B).

Supplementary Figure 6. Differential gene expression of clusters in the Monocle3 trajectory. (A) Genes enriched in leaf 14 of Monocle3 trajectory. Cells in UMAP colored according to log-normalized gene expression values. **(B)** Genes enriched in leaf 11 of Monocle3 trajectory. Cells in UMAP colored according to log-normalized gene expression values.

Video S1. Live cell conversion of tdTomato$^+$zsGreen$^+$ astrocyte to iOLC

Video S2. Live cell conversion of tdTomato^{neg}zsGreen$^+$ astrocyte to iOLC

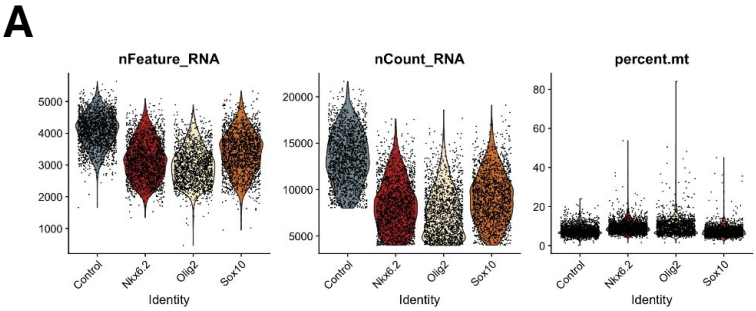
Table S1. 14DPT scRNA seq cell collection metrics

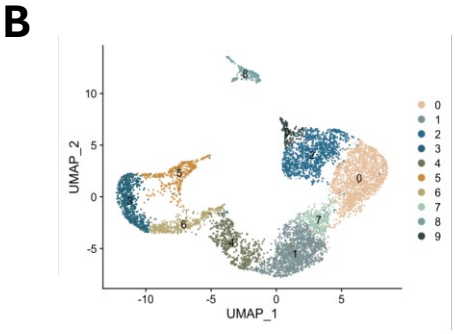
Table S2. Pre-transduction, 3 and 8DPT scRNA seq cell collection metrics

Table S3. Top 40 genes associated with Fatecode perturbation on node 16 of the latent layer

Data S1. BD Rhapsody Whole Transcriptome Analysis whitelist

Data S2. Monocle3-derived differentially expressed genes


A
nFeature_RNA
nCount_RNA
percent.mt
Control
Nkx6.2
Olig2
Sox10
Identity


B
UMAP_1
UMAP_2
0
1
2
3
4
5
6
7
8
9

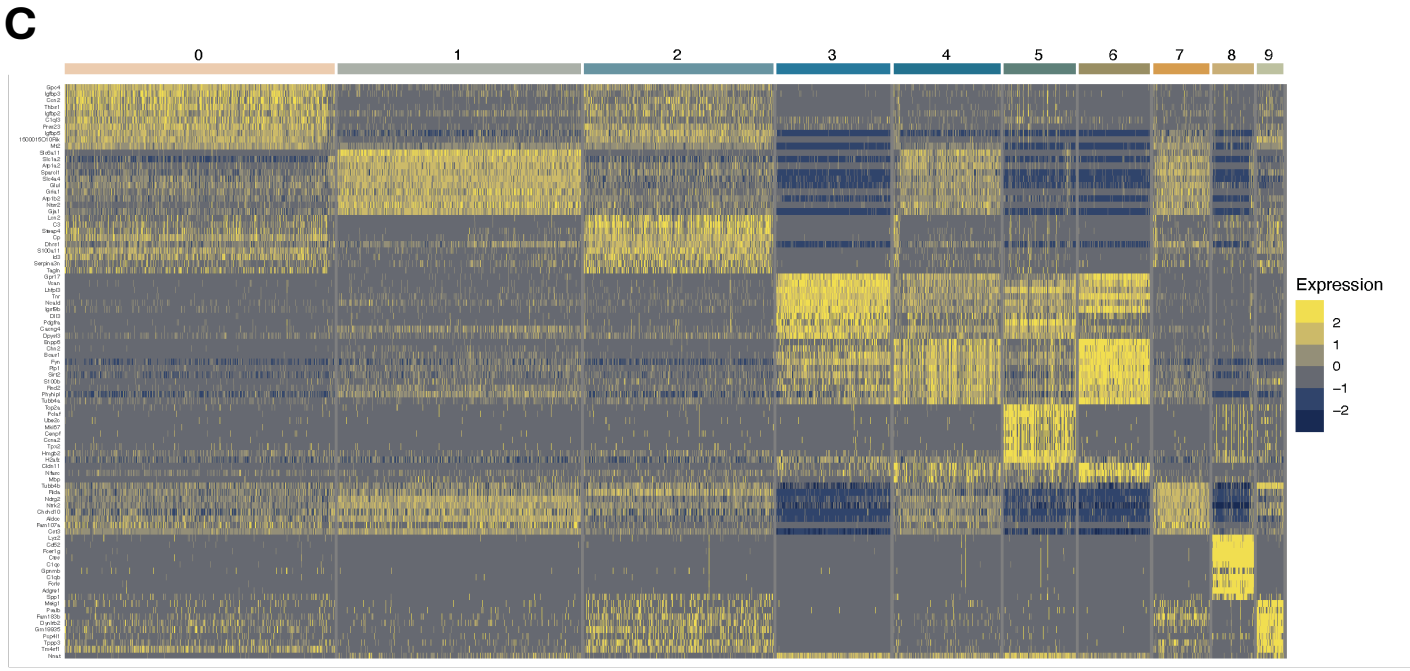

C
0
1
2
3
4
5
6
7
8
9
Expression
2
1
0
−1
−2

Figure S1

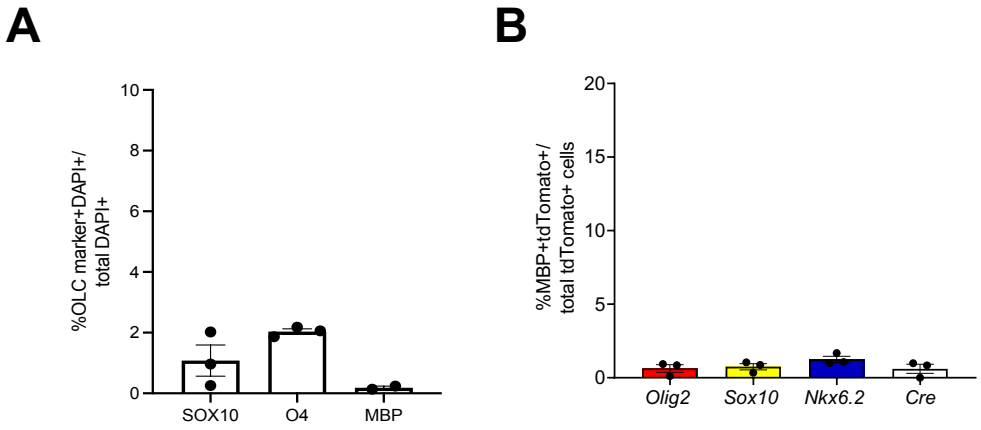
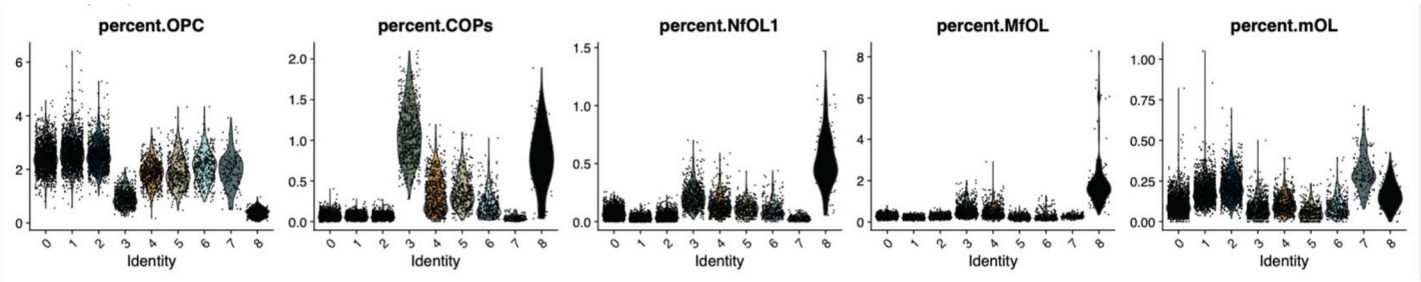

A
B
%OLC marker+DAPI+/
total DAPI+
10
8
6
4
2
0
SOX10
O4
MBP
%MBP+tdTomato+/
total tdTomato+ cells
20
15
10
5
0
Olig2
Sox10
Nkx6.2
Cre

Figure S2

A


percent.OPC
percent.COPs
percent.NfOL1
percent.MfOL
percent.mOL
Identity

B

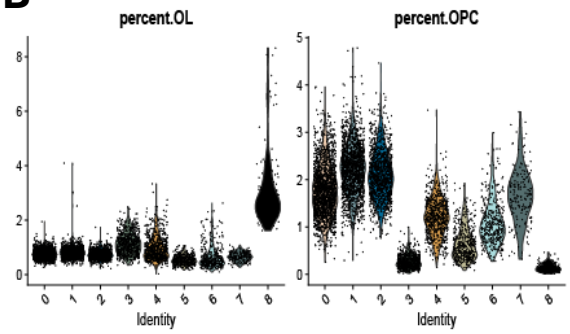

percent.OL
percent.OPC
Identity

Figure S3

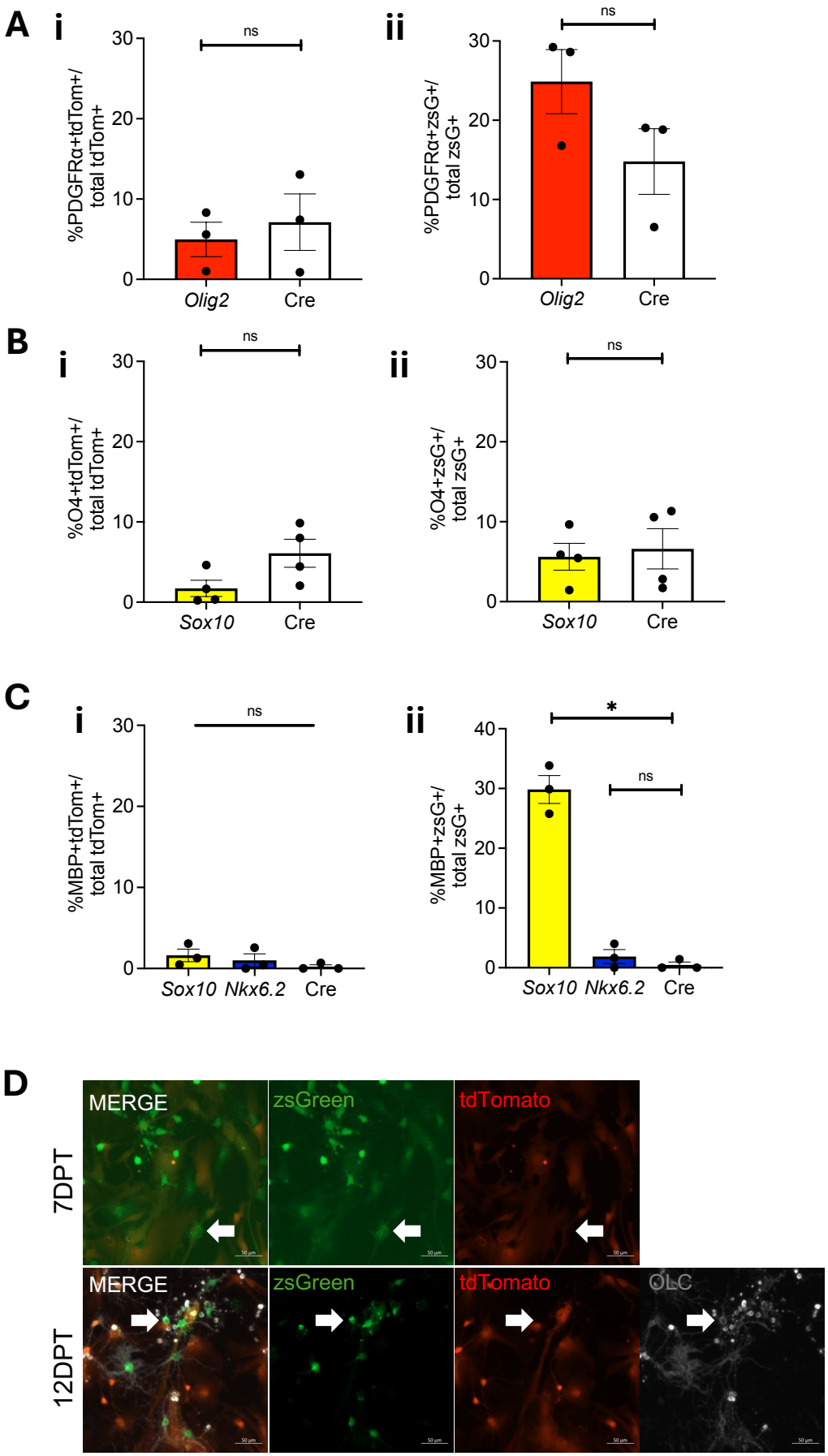

A
i
ns
%PDGFRα+tdTom+/ total tdTom+
30
20
10
0
Olig2
Cre
ii
ns
%PDGFRα+zsG+/ total zsG+
30
20
10
0
Olig2
Cre
B
i
ns
%O4+tdTom+/ total tdTom+
30
20
10
0
Sox10
Cre
ii
ns
%O4+zsG+/ total zsG+
30
20
10
0
Sox10
Cre
C
i
ns
%MBP+tdTom+/ total tdTom+
30
20
10
0
Sox10
Nkx6.2
Cre
ii
*
ns
%MBP+zsG+/ total zsG+
40
30
20
10
0
Sox10
Nkx6.2
Cre
D
7DPT
MERGE
zsGreen
tdTomato
12DPT
MERGE
zsGreen
tdTomato
OLC

Figure S4

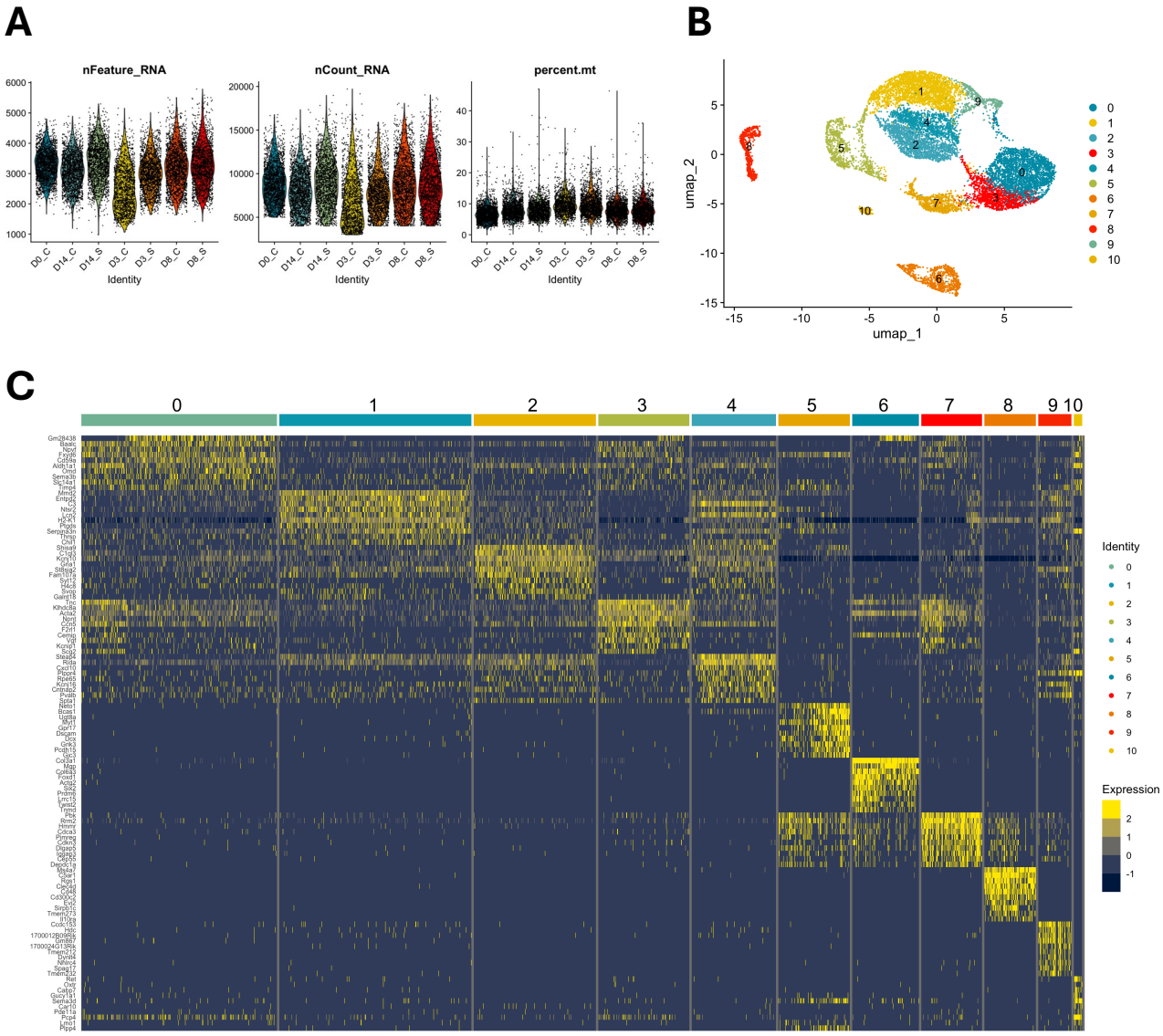
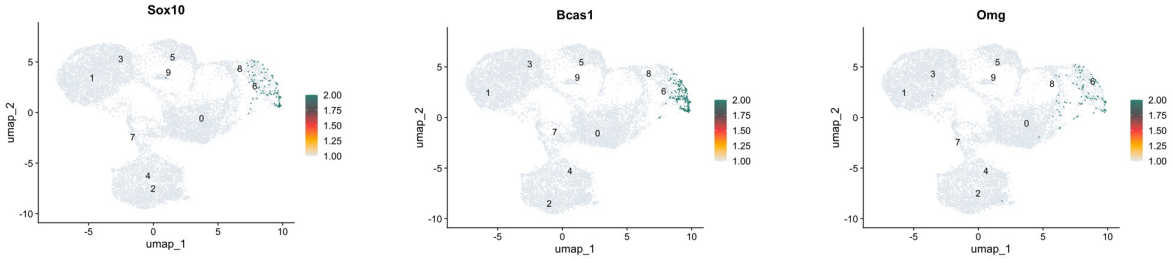

A
nFeature_RNA
nCount_RNA
percent.mt
D0_C
D14_C
D14_S
D3_C
D3_S
D8_C
D8_S
Identity
B
umap_1
umap_2
0
1
2
3
4
5
6
7
8
9
10
C
Identity
Expression

Figure S5

A


Sox10
Bcas1
Omg
umap_1
umap_2
2.00
1.75
1.50
1.25
1.00

B

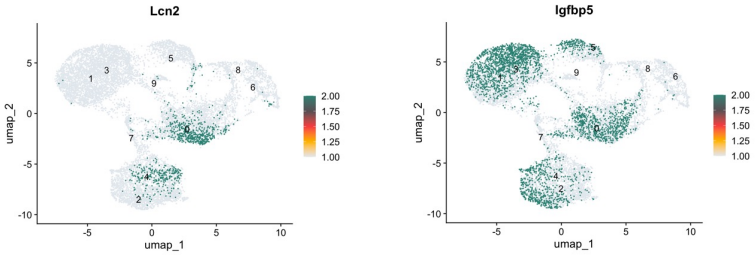

Lcn2
Igfbp5
umap_1
umap_2
2.00
1.75
1.50
1.25
1.00

Figure S6