Resolving the transcriptional transitions associated with oligodendrocyte generation from adult neural stem cells by single cell sequencing

Running Title: Oligodendroglial transcriptional networks

Kasum Azim[1*], Filippo Calzolari[2], Martina Cantone[3], Rainer Akkermann[1], Julio Vera[4], Hans-Peter Hartung[1], Onur Basak[5], Arthur Morgan Butt[6], Patrick Küry[1*]

[1] Department of Neurology, Medical Faculty, Heinrich-Heine-University, 40225 Düsseldorf, Germany.

[2] Institute for Physiological Chemistry, University Medical Centre of the Johannes Gutenberg University Mainz, Mainz, Germany.

[3] Faculty of Mechanical Engineering, Specialty Division for Systems Biotechnology, Technische Universität München, Germany.

[4] Laboratory of Systems Tumor Immunology, Department of Dermatology, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany.

[5] Department of Translational Neuroscience, UMC Utrecht Brain Centre, University Medical Centre Utrecht, Utrecht University, Utrecht, The Netherlands.

[6] Institute of Biomedical and Biomolecular Sciences, School of Pharmacy and Biomedical Science, University of Portsmouth, St Michael's Building, White Swan Road, Portsmouth, PO1 2DT, UK.

* Corresponding Authors

Email: Kasum.Azim@med.uni-duesseldorf.de; kuery@uni-duesseldorf.de.

ACKNOWLEDGEMENTS

We wish to thank Dominic Grün at the Max Planck Institute of Immunobiology and Epigenetics in Freiburg, DE for his valuable expertise with the pseudotime analysis. We also wish to thank Dr Marco Antonio Mendoza at Genoscope, FR, for establishment, guidance and optimization of TETRAMER.

Word count: 13,928 (inc References and Figure Legends)

Abstract

The subventricular zone (SVZ) is the largest neurogenic niche in the adult forebrain. Notably, neural stem cells (NSCs) of the SVZ generate not only neurons, but also oligodendrocytes, the myelin-forming cells of the central nervous system. Transcriptomic studies have provided detailed knowledge of the molecular events that regulate neurogenesis, but little is understood about adult oligodendrogenesis from SVZ-NSCs. To address this, we performed in-depth single-cell transcriptomic analyses to resolve the major differences in neuronal and oligodendroglial lineages derived from the adult SVZ. A hallmark of adult oligodendrogenesis was the stage-specific expression of transcriptional modulators that regulate developmental oligodendrogenesis. Notably, divergence of the oligodendroglial lineage was distinguished by Wnt-Notch and angiogenesis-related signaling, whereas G-protein-coupled receptor signaling pathways were the major signature observed in the neurogenic lineage. Moreover, in-depth gene regulatory network analysis identified key stage-specific master regulators of the oligodendrocyte lineage and revealed

new mechanisms by which signaling pathways interact with transcriptional networks to control lineage progression. Our work provides an integrated view of the multi-step differentiation process leading from NSCs to mature oligodendrocytes, by linking environmental signals to known and novel transcriptional mechanisms orchestrating oligodendrogenesis.

Word count: 186

Keywords: Oligodendrogenesis; Gene Regulatory Networks; Transcription Factors; Signaling Pathways; Neural Stem Cell; Single Cell Sequencing.

Main points:

- Distinct adult NSC populations giving rise to either oligodendrocytes or neurons can be identified by the expression of transcription factors.
- Gene regulatory control of oligodendrogenesis is a major fate-determinant for their generation.

INTRODUCTION

Oligodendrocytes are the myelin-forming cells of the central nervous system (CNS), and provide insulation for rapid axonal conduction and support to axons. In the forebrain, most oligodendrocytes are generated shortly after birth from neural stem cells (NSCs) of the subventricular zone (SVZ) (Azim, Berninger, & Raineteau, 2016; Kessaris et al., 2006), which are also responsible for neurogenesis (Figueres-Onate, Sanchez-Villalon, Sanchez-Gonzalez, & Lopez-Mascaraque, 2019), and this activity is retained well into late adulthood (Fuentealba et al., 2015). The sequence of oligodendroglial differentiation is relatively similar in postnatal and adult SVZ (Azim et al., 2016; El Waly, Macchi, Cayre, & Durbec, 2014), whereby NSCs first generate transiently amplifying progenitors (TAPs), characterized by elevated expression of the transcription factor *Ascl1* (Nakatani et al., 2013), followed by progressive differentiation into oligodendrocyte precursor cells (OPCs) and maturation into myelinating oligodendrocytes (Azim et al., 2016; El Waly et al., 2014). In rodents, approximately 5% of all cells derived from the adult SVZ are oligodendrocyte lineage cells (Capilla-Gonzalez, Cebrian-Silla, Guerrero-Cazares, Garcia-Verdugo, & Quinones-Hinojosa, 2013; El Waly et al., 2014; Menn et al., 2006).

Fate-mapping studies using Cre drivers under the control of regulatory regions from pallial and subpallial transcription factors demonstrated that molecularly distinct SVZ microdomains derive from their embryonic counterparts (reviewed in (Azim et al., 2016)). The embryonic septum, the lateral ganglionic eminences, and cortex contain NSCs that populate the medial (i.e. septal), ventral (i.e. striatal; also called lateral) and dorsal (i.e. cortical) aspects of the adult SVZ, respectively (Azim et al., 2016; Fiorelli, Azim, Fischer, & Raineteau, 2015). Localized labeling of NSCs showed that

regionally segregated NSCs are robustly committed to generating distinct cell subtypes (Menn et al., 2006; Merkle et al., 2014; Merkle, Mirzadeh, & Alvarez-Buylla, 2007), and that environmental signals impinge on intrinsic fate control mechanisms to modulate specific aspects of NSC and progeny behavior.

The OPC and oligodendrocyte transcriptional programs in both adult and postnatal contexts have been comprehensively dissected using single cell sequencing technologies (Marques et al., 2018; Marques et al., 2016; Zeisel et al., 2015). A recent study revealed regional differences in lineage output from the SVZ with its ventral microdomain exhibiting predominantly pro-neuronal-, and the medial wall exhibiting pro-oligodendroglial hallmarks (Mizrak et al., 2019), in line with previous observations using fate-mapping and transcriptomic profiling methods demonstrating that ventral SVZ-NSCs during postnatal development (Azim, Fischer, et al., 2014; Azim et al., 2015; Zweifel et al., 2018) and in the adult (Azim, Raineteau, & Butt, 2012; Dulken, Leeman, Boutet, Hebestreit, & Brunet, 2017; Llorens-Bobadilla et al., 2015; Felipe Ortega et al., 2013) are biased towards olfactory bulb neurogenesis. The generation of oligodendrocytes from the SVZ dorsal wall during early postnatal development is, however, considered as the default source for oligodendrocyte generation (Azim et al., 2017; Azim, Fischer, et al., 2014; Azim, Rivera, Raineteau, & Butt, 2014; Kessaris et al., 2006). Likewise, the adult dorsal SVZ serves as an additional microdomain for their generation (Azim et al., 2017; Azim et al., 2012; Crawford, Tripathi, Richardson, & Franklin, 2016; Felipe Ortega et al., 2013). In contrast to the number of high throughput sequencing studies defining the molecular events underlying neuronal generation from adult NSC subpopulations (Marcy & Raineteau, 2019), far less is known about the regulatory context of adult SVZ-

dependent oligodendrogenesis. This is partly due to technical limitations in isolating relevant microdomains, e.g. typical SVZ whole-mount preparations are devoid of dorsal SVZ tissue (Mirzadeh, Doetsch, Sawamoto, Wichterle, & Alvarez-Buylla, 2010). To overcome such restrictions, a cruder approach was used by capturing both dorsal and lateral walls (Basak et al., 2018), which enabled us to sequence sufficient numbers of single cells and to examine lineage-specific cells in the adult SVZ. Here, we describe the molecular events that regulate oligodendrogenesis from the adult SVZ and unravel mechanisms by which transcription factor networks and signaling pathways interact to drive specification in the early oligodendrocyte lineage.

Word count: 880

MATERIALS AND METHODS

Determining the lineage memberships of SVZ-derived single NSCs and TAPs

The procedures for transgenic and FACs strategies used for capturing single cells, mapping, pre-processing and RACEID are applied as described previously (Basak et al., 2018). The following reporter mice were used: quiescent(q)NSCs, few active(a) and primed(p)NSCs were taken from Troy:::GFP mice; TAPs and neuroblasts were harvested from Ki67:::RFP mice, and qNSCs (including a few pNSCs and aNSCs) were extracted from Ki67:::creRT2:::Tom2 mice. FACs analysis (Figure S1a) allowed identification/isolation of GLAST+/EGF- qNSCs, GLAST+/EGF+ aNSCs, GLAST-/EGF+ TAPs and O4+ oligodendrocytes. Parameters for cell clusters identification were adapted as done previously (Basak et al., 2018) to obtain SVZ-lineage specific clusters.

We predicted that the transcriptional hallmarks of oligodendrogenesis and neurogenesis from the postnatal SVZ microdomains are retained to a certain extent into adulthood. Bulk datasets of our previous study (Azim et al., 2015) were used for identifying transcriptional signatures of the entire oligodendrocyte lineage by comparing dorsal NSCs, dorsal TAPs and oligodendrocytes and for ventral SVZ-derived neurogenesis. Transcription factors used as markers for lineage identification in adult single cell were (i) highly expressed during early postnatal development and (ii) uniquely expressed in either oligodendrocyte or neuronal lineages. These included among other genes *Notch2, Epas1, Foxo1,* and *Esrrg2*. A first characterization was performed with RACEID to identify rarer cell populations in single cell sequencing (Grun et al., 2015), then PartekFlow® (http://www.partek.com/partek-flow/) was used to assign single cells to defined SVZ lineages, including the rarer ones from RACEID. Visual inspection of t-Distributed Stochastic Neighbor Embedding (tSNE) plots displaying transcript levels for selected lineage-specific markers was used to sub-cluster cells found in close proximity in the plots. Figure S1b contains examples of some of the markers used in this oligodendrocyte lineage detection step.

Transcription factor markers obtained from bulk datasets were used to identify single cells of the oligodendroglial lineage (Azim et al., 2015), aside from OPCs and mature oligodendrocytes within subpopulations of NSCs/TAPs (Figure S1b). The soundness of our lineage membership assessment was further evaluated by comparing the gene expression profile of the remaining cells, expected to belong to the larger neurogenic population, with bulk datasets of adult qNSC/aNSC/TAP/NB populations. Adult TAPs and neuroblasts we described in a recent study (Azim et al., 2018).

PartekFlow was used to generate heat PCA plots describing pallial and subpallial marker expression by cells in the early stages of the oligodendrocyte lineage (qNSCI, qNSCII and pNSC). Gene Specific Analysis algorithm from PartekFlow was used with default parameters for acquiring gene lists defining the distinct cell lineages. Lineage-specificities were considered statistically relevant if the corresponding *p*-values were below 0.05. An additional analysis for determining the expression signatures of common genes (shared between oligodendroglial and neuronal lineages) was performed by identifying cluster-enriched genes for the first 5 stages of differentiation (qNSCI-to-TAP) without separation of lineage-specific clones. The above steps allowed the definition of a total of 13 clusters: qNSCI, qNSCII, pNSC, aNSC, TAP, neuroblasts for the neuronal lineage and OLqNSCI, OLqNSCII, OLpNSC, OLaNSC, OLTAP, OPC and mature oligodendrocytes for oligodendroglial lineage.

Differential expression analysis was performed using datasets mentioned above by applying methods as described in (Azim et al., 2018) in addition to the cluster enriched transcripts using the Seurat V3 package as described in the next paragraph. Genes were filtered to select for transcription factors or transcriptional cues identified with PartekFlow. Macro-groups accounting for distinct differentiation stages were characterized as follows: (i) early oligodendrocyte group and the neurogenic group were defined by their expression of transcription factors expressed in single cells of (OL)qNSCI, (OL)qNSCII and (OL)pNSC; (ii) mid group for either lineage included transcription factors expressed in aNSCs and TAPs. The same datasets were used to identify enriched transcription factors when specific differentiation stages were compared. Precisely, the following stages were examined

as follows: (1) dorsal NSCs vs dorsal TAPs, OPCs, immature oligodendrocytes and mature oligodendrocytes; (2) dorsal TAPs vs dorsal NSCs, OPCs, immature oligodendrocytes and mature oligodendrocytes; (3) vNSCs vs vTAPs, TAPs and neuroblasts; (4) vTAPs vs vNSCs, TAPs and neuroblasts; (5) qNSCs vs aNSCs, TAPs and neuroblasts; (6) aNSCs vs qNSCs, TAPs and neuroblasts. Statistical analysis was performed using ANOVA for determining the differentially expressed (DE) transcription factors. Comparison of adult transcription factor expression in postnatal NSCs was summarized by assigning each transcription factor to a category: “1” indicates upregulated and highly expressed transcription factor, “-1” downregulated transcription factors, while “0” represents the remaining ones.

Clustering and visualization of single cell data using the Seurat R package

The R package Seurat V3 was used to process single cell transcriptomic data and elucidate their heterogeneity (https://satijalab.org/seurat/vignettes.html (Butler, Hoffman, Smibert, Papalexi, & Satija, 2018)). Single cells were assigned a unique identifier, then the default pipeline for analysis was performed: data were log-normalized and cells containing mitochondrial genes with a *p*-value greater than 0.05 were filtered out. The obtained matrix was scaled so that JackStraw function could be applied for comparing the *p*-values distribution for all principal components. Using the genes displaying higher variation among the 13 clusters identified with PartekFlow, the Seurat function FindVariableFeatures determined 7 oligodendroglial and 6 neuronal clusters that were employed for the downstream pseudotime analysis. Graphical representations were created using Seurat package. Differential gene expression analysis was performed with FindAllMarkers function applying Wilcoxon rank sum test to find transcriptional markers of the 13 clusters. Of note,

several statistical tests were implemented to ensure the reliability of the set of most relevant genes. Aside from DESeq2, differentially expressed genes across these methods were comparably similar (further detailed in https://satijalab.org/seurat/v3.0/de_vignette.html).

FateID cellular trajectory construction

The genes specific for the 13 clusters were processed for assessing lineage biases by using FateID package for R environment. The goal is to resolve early lineage priming via iterative random forest classification (Herman, Sagar, & Grun, 2018). The default settings of the package were applied. In order to determine the relationship between terminal differentiation states and naïve stem/progenitor states, mature oligodendrocytes were equated as the heterotypic outlier of the neurogenic lineage. The variable genes in the neurogenic clusters and mature oligodendrocytes were used as a subset expression data frame. The top 10 genes belonging to each cluster as determined by the above Seurat analysis were used as learning set for the classification method hence defining the v matrix, while neuroblast and mature oligodendrocyte clusters were indicated as target set. Using fateBias function, a few iterations were operated by comparing consecutive stages of differentiation by setting the training set as the previous stage and the test set as the following one. For example, the defined NSC/TAP pool of the oligodendroglial lineage (including neuroblasts) was used as training set and mature oligodendrocytes, OPCs and neuroblasts were classified as the test set. Differentiation trajectories were computed using the principal curve function “prc”. The dimensionality reduction coordinates of the analyses of the cells and differentiation trajectories (data contained in the “dr”

and “pr” lists) were exported and plotted using the ggplot2 R package for improved graphical representation.

Cell cycle regression

NSCs/TAPs/neuroblasts, quiescent or activated, were further investigated to determine their position within the cell cycle, using CellCycleScoring function of Seurat package. Full details are available at https://satijalab.org/seurat/v3.0/cell_cycle_vignette.html (Butler et al., 2018). Seurat estimates the mean expression of the marker genes from 43 S-phase and 54 G2/M-phase thus generating a uniform rank for S- and G2/M-phase for each cell, which results in the assignment of each single cell to a particular phase. The score was gauged by deducting the mean expression of the 10 nearest neighbor-cells in dimensionality reduction space computed by Seurat. Cell cycle scorings of cell populations were plotted according to their pseudotime trajectory coordinates as previously calculated. The data are visualized as cell cycle marker signal distribution in ridge plots (Figure 3). The aNSCs and TAPs of both lineages were further classified according to the different phases of the cell cycle, therefore leading to the identification of 15 distinct clusters (while the initial aNSC and TAP clusters were replaced). Seurat was run generating cluster-enriched expression profiles and those of the mid stages were used for further GO analysis as described below.

Gene ontology and pathway annotation of cell clusters

Gene lists for the original 13 clusters and the additional 4 ones (Figure 3e) were used for overrepresentation analysis using Protein ANalysis THrough Evolutionary Relationships (PANTHER) analysis tools ((http://www.pantherdb.org) (Mi, Poudel,

Muruganujan, Casagrande, & Thomas, 2016)). Comparing the overrepresented pathways of the oligodendroglial and neuronal clusters, PANTHER pathways function was used for pointing out the mechanisms that differ or are shared between stages within a lineage or between the same stage in oligodendroglial and neuronal lineages. In this analysis, a minimum of 4 pathways for each cluster was retained if the corresponding *p*-value was below 0.05. TAP clusters or the later stage OPCs/mature oligodendrocytes contained a greater number of highly expressed genes, and so the top 10 most significantly expressed genes found with Seurat were used. Identification of genes that are unique to the oligodendrocyte lineage, the neuronal lineage and the shared ones was performed using the R Package Venn Diagram and the online tool http://www.biovenn.nl/index.php. Pathway and Protein Class lists obtained from PANTHER were downloaded for the analyses described above and presented as dot plots using the ggplot2 R package. Similarly, the time courses for GO SLIM Biological Processes were constructed for the oligodendrocyte lineage using the 5 most significant terms. Pathway time course was represented applying the same method and using selected key pathways from the prior analysis.

Due to the potential reversibility of early (i.e. NSC) cell state transitions (Basak et al., 2018; Calzolari et al., 2015; Obernier et al., 2018), a considerable overlap among cell clusters in terms of expression profiles, Reactome and PANTHER Pathways, and Protein Class led to the definition of 3 broader groups. For this, the earlier stages constituting OLqNSCI, OLqNSCII and OLpNSC, the mid stages represented by OLaNSC and OLTAP, and the remaining OPCs and mature oligodendrocytes as the later stages were assembled. Results for signaling Pathways, Reactome and Protein Classes enrichment corresponding to the 3 groups were merged and the top

third most significant terms for each of the 3 groups were charted as dot plots and ranked in descending order based on their enrichment. Results for transcription factor Protein Class for the 7 oligodendrocyte lineage clusters were visualized with a heatmap representing the top 3 most enriched PANTHER Pathways related to "transcription factor function" across the 7 clusters. The heatmap was obtained with the pheatmap package in R.

Characterizing transcription factors for gene regulatory network (GRN) reconstruction

Once the most significant set of genes is determined, it is of importance to investigate if and how the genes are interacting at the molecular level. In addition to the genes previously identified as transcription factors, additional transcription factors known to play a central role in NSC differentiation were selected and their putative targets taken from http://genome.gsc.riken.jp/TFdb/tf_list.html, http://www.tfcheckpoint.org/index.php/browse, and PANTHER. Given the large numbers of transcription factors identified, a prioritization procedure was developed, relying on functional and transcription factor-transcription factor interaction data. Initially, the Cytoscape app GeneMANIA was used to identify functional interactions among the input genes (Franz et al., 2018). The data available for GeneMANIA for human gene orthologs are significantly more numerous and better characterized, therefore, mouse gene symbols were converted to human gene symbols. The genes used as input for GeneMANIA were limited to transcription factors and the queries were performed using the default parameter besides the type of interactions. In fact, the selected databases were "genetic" for downstream gene regulation and "physical" for protein-protein interactions. A first network (precursor to the one

depicted in Figure S7) was thus obtained, on which additional analyses were performed for prioritizing the large numbers of transcription factors included. For this, the heat diffusion algorithm was applied (Carlin, Demchak, Pratt, Sage, & Ideker, 2017) to facilitate identification of central transcription factors according to their level of interactions: the higher the number of interactions, the higher the calculated heat diffusion rank. The heat diffusion algorithm was applied 3 times separately, once focusing on genetic interactions, (Figure S7b), once on physical interactions (Figure S7c), and finally combining the newly obtained heat diffusion ranks (Figure S7a). The diffusion ranking values so obtained were plotted against the number of interactions of each transcription factor. As additional information, the nodes size of the included transcription factors was adjusted to their combined physical and genetic interaction ranks in Figure S7d.

Aiming to group the transcription factors to organize the network similarly to the clustering of FateID, the precursor network was organized to depict transcription factor expression according to their stage of expression. The primary goal is focusing on oligodendrogenesis, hence transcription factors whose annotation showed an enrichment towards neuronal lineage only were discarded. Genes common to both lineages at different stages were processed in order to define transcription factors that show specificity to either lineage. The specificity threshold was defined at linear 2-fold. This selection enabled emphasizing genes/transcription factors that are expressed in the oligodendrocyte lineage still including those ones sharing neuronal signatures. To assign transcription factors to the appropriate oligodendrocyte or neuronal clusters, the principle of higher expression was adapted. Nodes and labels sizes were organized based on the FDR values. A separate grouping of transcription

337 factors was made by selecting the transcription factors overlapping between aNSCs
338 and TAPs. Transcription factors that were significantly expressed in at least 2
339 different stages of the oligodendrocyte lineage were classified as "Pan transcription
340 factors". GeneMania networks presented in Figure S7a were used as visual aids for
341 facilitating the positioning of transcription factors according to the stages in
342 differentiation they are expressed in.
343
344 **TETRAMER reconstruction of GRNs of SVZ lineages**

345 The above step focuses on transcription factor functions in terms of their genetic
346 transcription factor-transcription factor interactions. This information was used as the
347 framework for a comprehensive inspection of transcription factor-target gene
348 interactions. For unravelling GRNs, TETRAMER (TEmporal TRAnscription regulation
349 ModellER) reconstructs fate transition-specific GRNs by integrating transcriptome
350 data (user provided) with inbuilt information (Cholley et al., 2018) (available gene
351 expression profiles, human genome-wide promoters and enhancers maps, and
352 ChIP-Seq available datasets) (see also http://ngs-qc.org/tetramer/.). The input data for
353 TETRAMER was constituted of a matrix containing 1467 genes – including
354 transcription factors – irregularly expressed in the oligodendrocyte lineage during
355 differentiation and, at the same time, expressed more in the neuronal lineage. Part of
356 the genes were included in the input matrix due to their association with GO
357 Biological Processes identified as major hallmarks of later stages of oligodendroglial
358 maturation. These hallmark GO Biological processes were: (i) Signal Transduction,
359 (ii) Ca Ion Transport, (iii) Synaptic Organization, (iv) CNS Development, and (v)
360 Multicellular Organismal Development. All of them presented a p-value < 0.0002 (as
361 estimated with Hypergeometric test). Specific hallmarks for mature oligodendrocytes

in the considered GO Biological Process were: (i) Cytoskeleton Organization, (ii) Myelination, (iii) Myelin Maintenance, (iv) Cell Adhesion, and (v) Transport. All of them presented a *p*-value < 0.007 (as estimated with Hypergeometric test). Genes consistently enriched more than 2.5 folds (aside from transcription factors) in OPCs and mature oligodendrocytes in the lists above and among DEGs found with Seuratwere considered for a further functional enrichment analysis using the freely available Functional Enrichment analysis tool (FunRich) http://funrich.org/index.html. GO Biological terms were reduced to fit, and a complete list has otherwise been made publicly available (see below).

The following settings were necessary prior to assessing the transcriptional signaling propagation in the GRNs with TETRAMER. The matrix consisting of 1467 genes, among which are transcription factors expressed in both linages (common transcription factors) or exclusively in either lineage, was used to construct 3 separate subnetworks (presented in Figure 4b-d). These networks were filtered accordingly to highlight transcription factors expressed in the neuronal lineage and excluding genes expressed in OPCs and mature oligodendrocytes. The nodes expressed in the qNSCI stages were named “start node” as they represent the initial time point of differentiation, and transcription factors expressed in more than 2 stages were termed “pan transcription factors”. The nodes expressed in final stages of differentiation, i.e. OPCs and mature oligodendrocytes for the oligodendrocyte lineage, were named “end node”. The time course for differentiation was arranged in order from qNSCI to mature oligodendrocytes, and all intermediate differentiation stages were selected for network propagation. In the TETRAMER results section, all transcription factors containing a yield were considered to be part of the main

signaling network. When the signal propagation network viewing was initiated, nodes in the network were colored from white/light red for the qNSCI nodes to darker red for mature oligodendrocytes, white to dark grey for transcription factors expressed in both lineage as ”common transcription factors”, and white to dark blue for transcription factors expressed in the neuronal lineage. Node sizes in these networks and subsequent analysis were graded based on the heat diffusion rank combined with additional parameters: node stress (calculated as the number of shortest paths passing through the considered node) indicating the extent of node activity, neighborhood connectivity, and the numbers of transcription factor-target gene interactions (edges). The values for each of these criteria were normalized in the range (0, 10) (low to high attribute values) and an average of all parameters was computed generating a final score termed as the connectivity index.

Reconstruction of the AdultOLgenesis GRN

The GRNs representing transcription factors expressed in the oligodendrocyte lineage and the ones common to both SVZ lineages were examined further by combining data of these two networks. The input matrix used for TETRAMER was expanded with genes that account for the GO Biological Processes for later stages of oligodendrogenesis. The TETRAMER analysis was performed as described above. The newly obtained larger network was expanded further to include the interactions with transcription factors known to play a central role in NSC differentiation. Additional ChIP-seq data for *Ascl1, Cebpa, Hdac2, Jun, Sox4, Sox9, Tgif2*, and *Tpr53* were obtained from the Chip Atlas database (Oki et al., 2018), while *Hopx* (He et al., 2016; Jain et al., 2015), *Zeb1* (Rosmaninho et al., 2018) were taken from individual studies. For *Trp53* and *Zeb1* original data, the potential targets were

annotated using human species gene names, hence to merge the results of the different datasets coherently, an R script was compiled for converting human-annotated to mouse-annotated genes (biomaRt package) (Durinck, Spellman, Birney, & Huber, 2009). For the potential gene targets extracted from the Chip Atlas database, a threshold on the score returned by the database was defined to select the targets with highest level of confidence (threshold values: 250). For *Hopx*, a different thresholding approach was applied, for the available data were processed differently: ChIP-seq data were processed with the Homer suit of tools, and therefore a different score was assigned to each identified peak. To discriminate between the relatively reliable targets, the average score was calculated and set as the discriminant threshold. For all the added transcription factors, a node was added to the precursor GRN representing the transcription factor itself. Interactions representing the ChIP-Seq data were created between the transcription factor and the targets already included in the precursor GRN. In order to select the target genes and generate this representation, a Python script was compiled. The code generated a text file associating to each transcription factor-target pair the source of the datum, the type of the interaction, and the type of the source data. To extend the number of interactions of the GRN specifically or oligodendrogenesis, the transcription factor-target gene networks were imported independently in Cytoscape and union-merged with the Exploration network recently published by Cantone and colleagues (Cantone et al., 2019). The union merge was repeated between the latest union merge and the AdultOLgenesis Network generated by TETRAMER, thus incorporating additional information. For the 11 selected transcription factors, targets were selected based on (i) a score describing the confidence of the actual formation of the transcription factor-target gene binding, thus increasing the reliability of the

437 interaction, and (ii) the biological context of the transcription factor-target gene
438 binding by limiting the set of genes to the ones previously identified in the Core GRN.
439 The confidence in the identification of the same transcription factor-target gene
440 binding in human and mouse species is ensured by a high level of genomic
441 similarity.
442
443 The obtained network was remodeled for positioning the nodes according to their
444 pseudotime trajectory coordinates, thus providing an improved overview for
445 transcriptional processes. The GeneMania network presented in Figure S7a was
446 used as reference to guide the correct placement of genes in individual stages of
447 differentiation. The Cytoscape plugin CoordinatesLayout version 3.0 and
448 copycatLayout version 0.0.9 were used. The average coordinates for defined cells at
449 each stage of differentiation was calculated along its x-y axis, and at each of the
450 estimated positions, a node was added to represent the corresponding differentiation
451 stage. The size of the node was scaled to reflect the relative abundance of
452 transcription factors associated to the differentiation stage. The “mock” nodes
453 representing oligodendrocyte cell clusters were organized according to their FateID
454 trajectories. Nodes representing transcription factors/genes associated with each of
455 the different stages of the oligodendrocyte lineage were grouped together and force
456 directed using Compound Spring Embedder (CoSE), and overlaid with the mock
457 nodes. Settings for CoSE were adjusted systematically for all oligodendrocyte stages
458 with edge lengths kept constant and spring strength/repulsion adjusted equally.
459 Nodes with higher heat diffusion ranks were permitted to overlay those genes with
460 lower heat diffusion ranks to highlight transcription factors with greater functional
461 relevance. Interactions were bundled using handles with the following parameters:

462 0.003 spring constant, compatibility threshold of 0.3 and 500 maximum iterations.
463 Nodes without interactions (that is, without edges) were excluded and hidden in the
464 background. Summaries highlighting downstream transcriptional regulation
465 (outwards transcriptional propagation) were constructed by adjusting node and edge
466 sizes reflecting the relative numbers of transcription factors/genes in the different
467 stages of differentiation and the downstream activation or repressive transcriptional
468 control, respectively. Five additional GO Biological Processes were manually
469 included close to later-stage oligodendrocytes. The entire reconstructed network was
470 termed “AdultOLgenesis GRN” and was thus an expansion of a recent study
471 describing GRNs during OPC-to-mature oligodendrocyte differentiation (Cantone et
472 al., 2019) by inclusion of additional regulatory features.
473
474 **Assessing transcriptional propagation by selected transcription factors and**
475 **Signaling Pathways**

476 Transcription factors expressed in defined stages of oligodendrogenesis (see also
477 Figure S7) or associated in PANTHER pathways were tested for transcriptional
478 propagation across the Core GRN (see also Section 7 of the online repository:
479 https://github.com/kasumaz/AdultOLgenesis). The transcription factors were selected
480 with their direct target genes and subsequent secondary and tertiary target genes.
481 All primary target genes were retained, whereas secondary and tertiary target genes,
482 if they were transcription factors, were filtered according to their protein-protein
483 interaction and connectivity index of >7.5 and >7, respectively. Biological Process
484 GO terms were preserved throughout. Interactions of transcription factors associated
485 with PANTHER pathways, secondary and tertiary target genes were sized in order of
486 line thickness in this order. Manual curation was operated for correcting the edges

between GO Biological Processes and PANTHER pathways. Transcription factors in each network were ranked in heatmaps for each pathway, according to the number of target gene interactions (activation/inhibition) across the different pathways selected for analysis. Gene regulatory data for the quantification of activation, inhibition and unspecified interaction were obtained from the edge section of the AdultOLgenesis GRN in Cytoscape.

Word count: 3860

RESULTS

Identification of the earliest stages of the SVZ-derived oligodendrocyte lineage

Transcriptional networks controlling oligodendrogenesis from adult SVZ-NSCs are currently unsolved. Here, NSCs/progenitors and subtypes of neural cells found within the young adult SVZ were isolated and FACS-enriched using genetic reporters and immunochemical markers, as described in a previous report on adult neurogenesis (Basak et al., 2018) (Figure S1a). Relying on well-established markers for the earlier stages in oligodendrogenesis as a guide, we identified putative lineage-specific subpopulations of NSCs (Figure S1b). Approximately 1200 single cells were visualized using tSNE plots and used for categorizing previously defined clusters and subclusters (Basak et al., 2018). We hypothesized that, as in the early postnatal SVZ, adult lineage-specific NSCs could be identified by their expression of oligodendroglial and neuronal hallmark genes. To this end, previously generated lists of genes associated with SVZ-oligodendrogenesis from bulk datasets were assembled and searched for oligodendrocyte lineage-enriched genes when compared to neurogenic populations (Azim et al., 2017). In this initial step, PartekFlow was used to input a number of landmark genes that define the earlier stages of the oligodendrocyte lineage. These include the essential transcription factors *Olig2*, *Olig1* and *Sox10*, and a cohort of other transcription factors as described in Materials and methods (Figure S1b). As an advantage, PartekFlow highlights cells based on input gene marker profiles, allowing the identification of neighboring cells that are transcriptionally similar. Following classification of putative oligodendrocyte lineage cells descending from NSCs/TAPs, gene lists derived from PartekFlow had initially revealed that the major differences in the 2 SVZ lineages are attributed to the expression of transcription factors (see below), in agreement with

the observed pre-eminence of transcription factors as lineage-specific markers within postnatal gliogenic lineages (Azim et al., 2015). Thus, expression levels of transcription factors were explored. In particular, single NSCs/TAPs expressing the above 3 essential transcription factors, as well as a cohort of others (Figure S1b) expressed in postnatal gliogenic NSCs were identified as likely to contribute to the oligodendrocyte lineage (Figure 1a) and amounted to approximately 7.2% of all NSC/TAPs across the 5 NSC and TAP stages of differentiation, which is in line with previous reports describing the relative scale of SVZ-oligodendrogenesis when compared to the generation of neurons (Menn et al., 2006). This particular subpopulation is identified by the co-expression of the known pan-NSC marker *Hes5* in qNSCs I/II expressing *Olig2* (Figure 1b). The first 3 stages of identified OLNSCs also expressed the pallial marker *Pax6* whereas they did not express the ventral SVZ markers *Gsx2* and *Nkx2-1* (Figure 1b). Most notably, an initial analysis of differentially expressed genes (DEGs) revealed patterns that are reminiscent of transcriptional programs found during postnatal oligodendrogenesis (Figure S1b). Identified cells of the oligodendrocyte lineage in the adult cluster closely with corresponding pooled populations (qNSCI-pNSC) in transcription factor expression, including postnatal gliogenic NSCs/TAPs (Figure S1c), reinforcing the view that subsets of adult NSCs are committed to an oligodendrogenic fate.

Next, FateID pseudotime analysis was performed to assess whether identified putative oligodendrogenic and neurogenic NSCs define distinct cell state trajectories (Herman et al., 2018). The genes that are significantly varying as determined by the Seurat analysis (see Materials and methods) in oligodendrogenic and neurogenic lineages, and the genes enriched in mature oligodendrocytes, were used for testing

the hypothesis that the selected oligodendrocyte lineage cells engage in a trajectory from earlier subpopulations of NSCs/TAPs. Oligodendrocytes and neuroblasts were used as the final lineage endpoints. Neuroblasts followed a continuous trajectory from qNSCI and many neurogenic TAPs resembled neuroblasts in pseudotime (Figure 1c). Addition of mature oligodendrocytes to this mostly neurogenic trajectory forces an early branching for the oligodendrocyte lineage, at qNSCI/II stages. However, when considering mature oligodendrocytes alongside putatively oligodendrogenic NSCs/TAPs and OPCs, mature oligodendrocytes branched off the main trajectory closer to OLaNSCs, TAPs and OPCs. Interestingly, addition of neuroblasts to the oligodendroglial trajectory resulted in a transcriptional path that overlapped partly with fewer OLTAPs compared to the neurogenic TAPs, highlighting the distinction between the two lineages. Genes detected in early- to mid-stage oligodendrocyte lineage cells (qNSCI to TAPs; Figure 1d-f) regulate signaling-to-transcriptional control of the oligodendrocyte lineage. Cluster-enriched markers (Figure 1f) include genes coding for the G-protein-coupled receptor genes *Lrig1* and *Adora2b* for OLqNSCI (Poth, Brodsky, Ehrentraut, Grenz, & Eltzschig, 2013; Simion, Cedano-Prieto, & Sweeney, 2014); the Bmp4-responsive gene *Htra1* serine protease (Chen et al., 2018), and the gliogenic transcriptional adaptor *Hopx* for OLqNSCII (Zweifel et al., 2018); the transcriptional regulator *Foxo1* (Kim, Hwang, Muller, & Paik, 2015), or the mitochondrial ketogenic enzyme *Hmgcs2* in OLpNSCs (Jebb & Hiller, 2018). OLaNSC and OLTAP markers include genes encoding for *Hspa1a* and *Hspa1b*, and *Cdk4* and *Hes6*, respectively (Arion, Unger, Lewis, Levitt, & Mirnics, 2007; Kim et al., 2015; Lukaszewicz & Anderson, 2011). This analysis revealed that many of the cluster-specific genes belong to families involved in signaling and transcriptional control, and demonstrated that cells representing the earliest stages

of the oligodendrocyte lineage are identifiable among NSCs, allowing the analysis of lineage-specific signatures. Further examples of pan lineage markers are given in Figure S2.

Stage-specific processes

Following determination of stage-/cluster-specific signatures, we focused on individual lineages (oligodendrocyte and neuronal, separately) and analyzed the cell clusters in each (7 oligodendroglial clusters, 6 neuronal clusters) for pathway enrichment using the PANTHER database. The most significant pathways for each stage of the 2 lineages are illustrated in dot plots showing the activity of a given pathway across the different clusters (Figure 2a). In this analysis, the genes exclusively belonging to 1 cluster were examined. Interestingly, the earliest stages of the oligodendrocyte lineage (qNSCI-pNSC) were uniquely characterized by the activity of pathways (e.g., “angiogenesis”, “cysteine biosynthesis”, “cytoskeletal regulation by RhoGTPases” and “pyruvate metabolism”), which also distinguished them from early NSC stages of the neuronal lineage. Contrastingly, a considerable overlap in pathways in the mid-stages (aNSC and TAPs) between the 2 lineages was evident, with few exceptions such as “Notch signaling” that was prominently enriched in OLaNSCs, and “p53 and cell cycle pathways” enriched in the same stages of the neuronal lineage.

All genes enriched in the oligodendrocyte lineage clusters (including those shared with the neuronal lineage) were next examined for characteristic biological processes (Slim Biological Processes in PANTHER) and plotted across the various stages of lineage progression. Machineries related to “lipid or energy metabolism”, “response

595 to signaling" or "developmental patterning molecules" were enriched in the earlier
596 stages, whereas mechanisms such as "translation", "gene expression" and "cell
597 division" were most prominent in OLaNSCs and OLTAPs (Figure 2b), consistent with
598 previous observations (Llorens-Bobadilla et al., 2015). At a later stage of
599 differentiation, OPCs engage in processes that were also active in OLqNSCII. For
600 example, GO biological process "cell communication" is enriched in OLqNSCII, and
601 then again at the OPC stage (Figure 2b), possibly accounting for the close proximity
602 of these 2 segregated populations in both the tSNE plots and pseudotime (Figure
603 1a,c). Many of the characteristic processes detected in mature oligodendrocytes
604 increase in expression during the course of lineage progression. We expanded this
605 analysis for both protein class and signaling pathways, to include the neuronal
606 lineage (Figure S3), observing that transcription factors and nucleic acid-binding
607 proteins are highly abundant in the early and mid-stages of the oligodendrocyte
608 lineage along with major developmental pathways. A time course for the expression
609 of the most significant signaling pathways in this analysis points to key processes
610 regulating transition along the oligodendrocyte lineage (Figure S4), with increasing
611 transcriptional activity suggested by the pattern of expression of transcription factors
612 and nucleic acid-binding proteins across the different oligodendrocyte lineage stages
613 (Figure S5). Altogether, these findings outline potential mechanistic differences in the
614 control of progression within the 2 SVZ-derived lineages, with notable differences in
615 inferred signaling pathway activities and in the expression of transcriptional
616 modulators.
617

618 **Cell cycle properties of SVZ-NSCs/progenitors**

Given the observed differences in cell cycle, transcriptional control and nucleic acid synthesis pathways in mid-stages (OLaNSC, OLTAP) of the oligodendrocyte lineage we further explored stage-specificity. To this end, cell clusters were analyzed by computing cell cycle phase scores based on canonical markers for the 3 main cycle phases (S, G2M and G1 phase) using the Seurat V3 package. Single cells´ expression levels of established cell cycle markers were compared and data were regressed out by modeling the relationship between expression signatures of each cell with cell cycle markers. Upon identification of the predicted cell cycle state for cells of all 13 clusters, those representing the oligodendrocyte lineage were plotted in accordance to their pseudotemporal coordinates and summarized in bar plots for both lineages (Figure 3a,b), and signal intensities as ridge plots (Figure 3c). In both lineages, the early stage NSCs (qNSCI-pNSC) and OPCs were almost entirely at the G1 phase with signal intensities shifted towards baseline, whereas S and G2M marker signals were detected in neuroblasts (Figure 3c). More than half of the aNSCs in either lineage were actively cycling and rarely at the G2M phase, whilst the remaining were quiescent. Although no proneuronal TAP was found to be quiescent, with just over 55% at the G2M phase and the remaining in S phase, a slightly greater proportion of OLTAPs were in S phase. Signal intensities for the S and G2M markers were relatively absent in the early neuronal lineage.

Following identification of the mid-stage cell cycle properties, Seurat V3 was used to identify cluster-specific genes in OLaNSCs and OLTAP by subclassification according to their predicted cell cycle phase. Therein, transcripts significantly expressed by the 9 cell clusters (OLqNSCI, OLqNSCII, OLpNSC, OLaNSC in G1, OLaNSC in S Phase, OLTAP in G2M, OLTAP in S Phase, OPC and mature

644 oligodendrocytes) were extracted and a further pathway analysis was performed on
645 the mid-stages of the oligodendroglial lineage (OLaNSC in S Phase, OLTAP in G2M
646 and OLTAP in S Phase). Pathways and Protein Class determination was performed
647 for OLaNSCs in the G1 and S phase in the intercept (genes overlapping in the 2
648 groups) in this analysis and the same for G2M and S phase for OLTAPs (Figure 3d,
649 e). Interestingly, aside from Notch and Hedgehog signaling, metabolic-like
650 mechanisms that include thiamine metabolism and threonine biosynthesis are also
651 features of quiescent OLaNSCs, whereas cycling aNSCs were characterized by
652 pathways regulating transcription, translation, and nucleic acid binding. In OLTAPs at
653 the G2M phase, EGF-, p38 MAPK-, p53 and interferon gamma-signaling were
654 significantly enriched. Furthermore, as expected, many of the pathways expressed in
655 actively cycling TAPs included regulation of mRNA, replication, and cell cycle-related
656 processes (Figure 3d). Cell cycle-related analysis of mid-stage oligodendrocytes
657 accordingly revealed that the major protein classes expressed in G2M and S phases
658 were “nucleic acid-binding” and “transcription factor”, whereas OLaNSCs in G1 were
659 characterized by extracellular matrix and vesicle-associated SNARE proteins (Figure
660 3e). Altogether, examining heterogeneity of cell cycle states in proliferative cells of
661 the oligodendroglial lineage suggests that transcriptional cues are a major driver in
662 OLqNSCII and in mid-stages of the lineage, which warrants further investigation.
663

664 **Gene regulatory networks in NSCs/TAPs driving myelination**

665 The previous analysis disclosed major classes of genes expressed in earlier
666 oligodendroglial lineage cells, notably with mid-proliferative cell stages featuring
667 abundant expression of transcriptional modulators. Therefore, the expression of
668 transcription factors (including transcriptional adapters and transcription cofactors)

669 was further explored in both the oligodendroglial and neuronal lineages.
670 Transcription factors were prioritized via a GeneMania-based genetic and physical
671 interaction analysis. This step facilitates predicting the effects of genetically
672 perturbing a given transcription factor, in terms of its impact on neighboring genes,
673 based on predicted transcription factor-to-targetgene interactions and protein-protein
674 interactions (see Materials and methods; Figure S6). To determine which
675 transcription factors are most relevant according to their genetic perturbation, genes
676 were prioritized by applying the “heat diffusion” algorithm that tests the input query
677 data and the functional interaction of each gene for propagation across the network.
678 Strongly connected nodes in the network, supporting sufficient regulatory activity
679 propagation, are uncovered and allow creation of subnetworks, enabling exclusion of
680 nodes with lower functional relevance. Prioritization steps were applied for both
681 physical (protein-protein interactions) and genetic interaction (Figure S6a and
682 expanded further in Figure S7) and identified transcription factors with potentially
683 central roles in specifying cell states.
684
685 We next employed TETRAMER, a tool integrating inferred differentiation states
686 during lineage progression with gene regulatory networks (GRNs) from publicly
687 available datasets (Cholley et al., 2018), for the reconstruction of GRNs whose
688 activity may define cell states within the oligodendroglial and neuronal lineages.
689 Briefly, lists of transcription factors expressed in the identified clusters of the
690 oligodendroglial and neuronal lineages as well as transcription factors that are
691 expressed in both of these lineages (“common” transcription factors), were used as
692 input by TETRAMER to explore gene expression profiles of mammalian tissues/cells
693 with comprehensive analysis of deposited ChIP-sequencing data corresponding to

694 transcription factor-binding and further comprising data on actively transcribed
695 enhancers and promoters. In the preliminary step, transcription factor-transcription
696 factor and transcription factor-target gene interactions were compiled for determining
697 regulatory interactions between shared neurogenesis- and oligodendrogenesis-
698 restricted (Figure S6b-c) transcription factors. A large proportion (89.3-91.8% of all
699 interactions) of transcription factor-target gene interactions for transcription factors
700 common to both lineages and those regulating neurogenesis consisted in predicted
701 target gene activation. Interestingly, however, activating and inhibiting regulatory
702 interactions were predicted to occur at similar frequencies between
703 oligodendrogenesis-specific transcription factors, highlighting qualitative differences
704 in how cell state transitions are controlled at the transcriptional level in the 2
705 lineages. Transcription factor interactions with target genes were quantified using the
706 Cytoscape network analysis parameters by probing the entire training GRN
707 reconstructed (as described in Figure 4) and presented as intensity heatmaps for the
708 top 25 transcription factors with the most global, activating and inhibiting target gene
709 interactions in Figure S6e (ranked transcription factor examples in Figure S6f).
710 Amongst the most highly ranked transcription factors were *Cebpa,* expressed in
711 OLqNSCII and best described in the context of myelopoiesis, where it inhibits
712 proliferation (see (Calella et al., 2007) for proposed roles during neural
713 development), and *Ezh2*, which promotes embryonic NSC proliferation and
714 differentiation into OPCs (Sher et al., 2008). Altogether, these analyses enable
715 ranking transcription factors that are likely to act as “master regulators” based on a
716 number of parameters that include their target gene regulation, protein-protein
717 interactions and diffusive effects downstream. To further investigate the major gene
718 regulatory cascades controlling adult oligodendrogenesis, the GRN assembled for

719 common transcription factors and oligodendrogenesis (Figure S6b,d) were combined
720 and merged with a recently assembled GRN for post-OPC stages of oligodendroglial
721 maturation (Cantone et al., 2019) (Figure 5), which defined the role of transcription
722 factors such as *Olig2*, *Sox10*, and *Tcf7l2* during oligodendrocyte differentiation and
723 maturation. Here, genes overlapping in later-stage oligodendrocytes from the
724 present study with those published recently (Cantone et al., 2019), were filtered and
725 represented as GO Biological Processes in a larger merged network (Figure 5).
726 Approximately, a quarter of the differentially expressed genes, as determined by the
727 Seurat analysis, were used in this TETRAMER analysis. Of note, modulation of
728 many of these genes has previously been reported during postnatal and adult
729 oligodendrogenesis from NSCs (Azim et al., 2015). This final network termed as
730 “AdultOLgenesis GRN” was used for further analysis as described below. The
731 individual stages were then positioned relative to the average positioning of single
732 cells/clusters in FateID pseudotime (Figure 1c) (see Materials and methods for
733 network assembly).
734

735 **Stage-specific GRN**

736 Focusing on target gene activation or repression in substages of the oligodendrocyte
737 lineage reveals varying and contrasting degrees of gene regulation (Figure 5a-f).
738 Notably, the OLqNSCII, OLaNSC and OLTAP stages exhibited prominent gene
739 regulatory activities, with OLaNSCs seemingly progressively activating TAP-specific
740 target genes (Figure 5b,e,f), whilst also inhibiting the expression of genes typical for
741 the most differentiated (OPC/mature oligodendrocyte) stages. This suggests that,
742 following activation, oligodendrocyte lineage cells would steadily progress to the TAP
743 stage, which could be temporarily stabilized by active repression of differentiation-

associated gene modules. Interestingly, the uniquely low expression levels of transcription factors active in OLqNSCI and OLpNSCs are predicted to only modestly result in GRN state restructuring at these stages. These findings demonstrate that transcription factors expressed in distinct stages during oligodendrogenesis have contrasting modes of gene regulation.

Next, the highest ranked transcription factors for each of the selected oligodendrocyte lineage stages were probed for the extent to which their inferred transcriptional regulatory activity propagates within the assembled GRN. The downstream gene regulatory effects of *Olig2* and *Sox10*, which modulate the expression of a large number of genes involved in gliogenic and neurogenic processes (Cantone et al., 2019), provided a quantitative reference to assess the GRN-modulating activity of newly identified transcription factors. The immediate downstream target genes for *Olig2* and *Sox10* are presented in a subnetwork (Figure 5g), and downstream propagation cumulatively modulates 75% of all network components within one further tier. The regulatory activity of *Mitf* and *Rreb1*, expressed in OLqNSCI, spread the least within the GRN, supporting the previous analysis (Figure S6e) that shows *Rreb1* as the highest ranked transcriptional repressor (Figure 5h). The impact of OLqNSCII transcription factors *Cebpa* and *Epas1* quantitatively resembles that of *Olig2* and *Sox10* (Figure 6i). Interestingly, *Cebpa* has been described for myelopoiesis as inhibiting cell cycle-related genes (Calella et al., 2007), which fits with its predicted inhibitory interactions with major cell cycle genes expressed in TAPs, supporting its role as a quiescent NSC-specific transcription factor. The 2 highest-ranked OLpNSC transcription factors, *Foxo1* and *Nr1d1*, induce the expression of a number of cell cycle regulators expressed in

OLaNSCs and OLTAPs (Figure 5j). Like OLqNSCIIs, OLaNSCs and OLTAPs express transcription factors whose predicted influence on oligodendrocyte GRN activity resembles that exerted by *Olig2* and *Sox10*. These findings identify putative core regulators of stage-specific oligodendrocyte GRNs, thus complementing known pan-oligodendrocyte transcription factors *Olig2* and *Sox10* and providing a view into stage-specific control of oligodendroglial lineage progression.

PANTHER Pathway Integration of GRNs

Having defined GRN state transitions controlling oligodendrocyte lineage progression, we attempted integrating this knowledge with stage-specific signaling pathway activity predictions from our previous PANTHER analysis (Figure 2b), reasoning that the identified signaling pathways would impact cell fate by ultimately modulating transcriptional networks. We explored this hypothesis using the reconstructed AdultOLgenesis GRN, examining signaling pathways known to modulate at least 2 transcription factors expressed at a given stage, and then assessing the predicted propagation of their transcriptional modulatory activity within the GRN. The direct target genes and additional secondary target genes regulated by both *Olig2* and *Sox10*, as performed in the prior analysis in Figure 5g, were used as a gauge for transcriptional coverage of each tested transcription factor associated with PANTHER pathway terms. In addition, the transcription factors induced by selected PANTHER signaling pathways were ranked based on the number of transcription factor-target gene interactions. These were further classified according to their gene-activating, -inhibitory, and unspecified effects onto the oligodendrocyte lineage in red, green, and grey, respectively (Figure 6a). In the earlier stages of the oligodendrocyte lineage, Wnt signaling is predicted to affect approximately half as

many GRN components as those modulated by *Olig2* and *Sox10*. Our previous reports demonstrated that manipulating the Wnt signaling pathway by pharmacologic or genetic strategies increases the expression of *Ascl1*, *Olig2* and *Sox10* transcripts by at least 2.5-fold (Azim, Fischer, et al., 2014; Azim, Rivera, et al., 2014). In the present study, these essential transcription factors were indeed predicted to be secondary target genes of the Wnt pathway (Figure 6a,b), thus potentially accounting for the previously described enhancements of myelination following Wnt signaling modulation (Azim et al., 2017; Azim, Fischer, et al., 2014; Azim, Rivera, et al., 2014) and validating this strategy. PDGF signaling on the other hand induces similar gene regulatory events as *Olig2* and *Sox10*, although initially repressing genes expressed in later stage oligodendrocytes (Figure 6c), in line with its role as an OPC mitogen. However, *Stat1* and *Stat5b*, transcription factors modulated by the PDGF pathway, activate transcriptional networks in TAPs that enhance the expression of genes in OPCs and mature oligodendrocytes, leading to extensive downstream transcription factor induction that included a number of master regulators (see heatmap in Figure 6c). Interestingly, the pathways active in OLaNSCs (“Transcriptional regulation by bZIP transcription factors”, “Gonadotropin signaling pathway”) and in OLTAPs (“General Transcriptional Regulation”) are involved in a number of gene regulatory events larger even than those elicited by *Olig2* and *Sox10*, suggesting extensive GRN reorganization at these stages, possibly reflecting a lineage watershed on the way to differentiation. Indeed, a number of the above identified master regulators (Figure S6e) are key effectors of PANTHER pathway terms enriched in OLaNSCs and OLTAPs. Other potent signaling pathways examined include Notch signaling and p53 active in OLaNSCs (G1 and S phase) and TAPs (G2/M and S phase). Other pathways identified in the

earlier PANTHER analysis, including “Angiogenesis” (OLqNSCI), “Circadian clock” (OLqNSCII/qNSCII) and “p53 Feedback Loops 2” (G2M and S phase OLaNSC/OLTAP), had overall fewer predicted transcriptional effects. Altogether, these findings demonstrate that large cohorts of transcription factors in the reconstructed Core GRN are effectors of multiple signaling pathways and provide useful insights into how environmental cues could affect transcriptional activities, and thus lineage progression, during SVZ-derived oligodendrogenesis. Future studies will aim at expanding these observations by testing the functionality of identified signaling pathways and to confirm changes in target expression experimentally.

Word count: 3500

DISCUSSION

In the present study, a meta-analysis of single cell RNAseq data from the adult SVZ was performed to resolve the transcriptional signatures that distinguish between adult neurogenesis and oligodendrogenesis. Recent transcriptomic analyses focusing on early postnatal development and adulthood (Azim et al., 2017; Azim et al., 2015; Mizrak et al., 2019), provided some insights into the gene expression profiles associated with oligodendrogenesis. A number of newer transcriptomics studies of adult neurogenesis have shed light onto the molecular processes regulating the generation of olfactory bulb neurons (reviewed in (Marcy & Raineteau, 2019)). In addition, landmark studies using single cell profiling of murine oligodendrocytes from different stages of life in the CNS have immensely contributed to our understanding of the heterogeneity and molecular regulation of oligodendrocyte differentiation (Marques et al., 2018; Marques et al., 2016; Zeisel et al., 2015). A recent elegant single-cell sequencing survey of the lateral and medial walls of the young adult SVZ demonstrated that NSCs located in microdomains other than the lateral wall generate oligodendrocytes (Mizrak et al., 2019), but an in-depth investigation into the mechanisms controlling adult oligodendrogenesis had not yet been performed. To this end, we analyzed recently generated single-cell data from the SVZ (Basak et al., 2018) focusing on NSC/progenitor heterogeneity, in order to identify putative oligodendrocyte- and neuron-restricted progenitors (Felipe Ortega et al., 2013). Our findings reveal transcriptional regulators as major protein classes modulating all stages of the SVZ-NSC-derived oligodendrocyte lineage.

We carefully inspected the transcriptomes of NSCs and TAPs for the expression of known early oligodendroglial lineage marker genes (Figure S1). These included members of the Sox transcription factor family and *Olig2/Olig1*, which were found to be expressed selectively in NSC/TAPs that also expressed other known gliogenic markers such as *Hopx, Notch* genes and their subsequent target genes of the *Hes* family, overall resembling the pattern of expression observed during postnatal oligodendrogenesis (Azim et al., 2015). Oligodendrocyte lineage NSCs/TAPs amounted to approximately 7% of all NSCs and TAPs, which is in line with earlier retroviral-fate mapping observations describing a similar oligodendrocyte/neuron output ratio between from the adult SVZ (Menn et al., 2006). Comparison of single cells at a defined stage of the two lineages showed only a modest degree of separation in tSNE plots. While this implies close similarity between the early steps of oligodendroglial and neuronal lineage progression, key differences were associated with specific classes of genes, such as transcription factors. Notably, more closely focusing on lineage-specific features by GO “Protein Class” and “Pathway” analyses revealed that from the earliest quiescent NSC stages to TAPs, the terms related to “transcriptional cues” accounted for most of the differences between the two lineages. In agreement with the findings of the present study, these broad classes of genes have recently been described as enriched in oligodendroglial clusters of adult NSCs/TAPs (Mizrak et al., 2019), similar to observations made during postnatal development in pre-OPC (TAPs) populations (Marques et al., 2018), and elsewhere during embryonic oligodendrogenesis (Klum et al., 2018). Our comprehensive side-by-side comparison of protein classes differentially expressed by the two SVZ lineages unequivocally demonstrates transcriptional control as a major determinant of fate specification and lineage progression and warranted

878 further in-depth characterization of the GRN controlling oligodendrogenesis (see
879 below). Pathway analysis of the most immature stages (qNSCI to TAP) in the two
880 lineages suggested previously unappreciated mechanisms controlling the generation
881 of adult oligodendrocytes from the SVZ. Important pathways detected in the qNSCI/II
882 and pNSC stages of the oligodendrocyte lineage included Wnt signaling,
883 angiogenesis and Notch signaling, all known as important pathways regulating adult
884 oligodendrogenesis (reviewed in (El Waly et al., 2014)). Other key pathway terms
885 derived from this analysis include “Circadian clock”, which has not previously been
886 described as a regulator of oligodendrogenesis, although an earlier study proposed
887 this pathway as a cell-intrinsic timer for inhibiting cell division in postnatal OPCs
888 (Gao, Durand, & Raff, 1997). Interestingly, PDGF signaling was uniquely enriched in
889 OLpNSCs in our analysis and we predict that this oligodendrocyte lineage stage
890 comprises the SVZ cell population previously reported to undergo *in vivo* expansion
891 upon exposure to infused PDGF-A (Jackson et al., 2006; Moore, Bain, Loh, &
892 Levison, 2014). Pathways detected in aNSC and TAP populations comprise “Cell
893 cycle”, “DNA/nucleotide synthesis” and “Transcriptional machineries”, consistent with
894 single cell profiling studies of adult neurogenesis (Dulken et al., 2019; Llorens-
895 Bobadilla et al., 2015). We also identified Notch signaling as a regulator of
896 oligodendrocyte-fated aNSCs and TAPs, which is in agreement with previous
897 observations in the context of early SVZ-derived gliomagenesis (Giachino et al.,
898 2015) and during OPC generation from embryonic NPCs (Cui et al., 2004).

899

900 Interestingly, scoring the early oligodendroglial and neuronal lineage cells for their
901 proliferative state revealed instead broad similarities between corresponding cell
902 stages, suggesting that lineage commitment and proliferative expansion are at least

in part independently controlled. These similarities were fewer at post-TAP stages, where OPCs and the very earliest neuroblasts (before emigration via the rostral migratory stream) comprised mostly of quiescent cells and proliferative states, respectively. The latter stage, upon pharmacogenomically instructed manipulation for directing specific cell fates in older adult mice, results in its expansion and subsequent dorsal-SVZ-derived OPCs (Azim et al., 2017). In support of these findings, “Nucleic acid binding” and “transcription factor” pathways were most prominent during S phase compared to G1 and G2M phases, suggesting transcriptional modulation associated with access to *de novo* synthesized DNA (e.g. propagation or passive loss of modifications on newly added histones) (Figure 3e). Thus, targeting transcription factor programs regulating cell cycle phases in adult OLTAPs as previously described (Azim et al., 2017), presents additional avenues in promoting the generating of adult-born OPCs.

Establishment of a pan-oligodendrocyte GRN allowed us to investigate transcriptional network organization during lineage progression, defining stage-specific GRN states and their core regulators, which add to well-established broad-oligodendrocyte transcription factors such as *Olig2* and *Sox10*. These transcription factors are well-characterized as binding to the promoters of a large number of genes regulating several aspects of NSC/progenitor behaviors. *Olig2*, aside from regulating transcriptional programs associated with lineage progression and myelination in the post-OPC stages, also represses proneuronal and quiescence genes and activates genes that promote cell cycle entry and oligodendrogenesis in NSCs/NPs (Mateo et al., 2015). A number of *Sox10* target genes overlap with those of *Olig2* in later stage oligodendrocytes (Cantone et al., 2019), while comparatively

928 little is known about oligodendrocyte transcription factor genome occupancy patterns
929 in SVZ-NSCs. In our analysis, *Sox10* mRNA was detected already in OLaNSCs,
930 consistent with its expression downstream of *Olig2* (Liu et al., 2007) and we predict
931 *Sox10* to be downregulating genes expressed in the neuronal lineage, including *Sufu*
932 that represses signaling pathways (e.g. Wnt and Shh signaling) which limit the
933 specification of OPCs (Pozniak et al., 2010).

934

935 Surprisingly, transcriptional control of lineage progression was most marked during
936 the actively proliferating oligodendrocyte lineage stages (OLaNSC and OLTAP) and
937 at the OLqNSCII stage, which we explored for predictive functions of transcription
938 factors. In our reconstructed GRN, analyzing transcription factor-target gene
939 interactions in the oligodendrocyte lineage aside from the well-characterized
940 transcriptional regulators, we uncovered novel master regulators. We used our
941 reconstructed GRN for predicting gene regulatory functions of these putative master
942 regulators. Amongst those identified were *Cebpa*, which has been described to
943 regulate genes expressed in the cycling population of the oligodendrocyte lineage
944 (Calella et al., 2007), and *Mitf*, which activates the transcription of the *Dct* gene (Jiao
945 et al., 2006), previously identified as highly expressed in gliogenic SVZ-NSCs (Azim
946 et al., 2015). Transcription factors expressed by OLaNSCs tended to positively
947 regulate TAP-specific genes, thus possibly promoting rapid lineage progression,
948 while the set of TAP-expressed transcription factors suggests concomitant
949 reinforcement of the TAP-specific gene-expression program and repression of
950 differentiation-associated genes (Figure 6e,f). Indeed, target gene repression
951 emerged as more widespread than in the neuronal lineage and may reflect the need

for stabilization of an intermediate undifferentiated stage after initial lineage progression and expansion downstream of NSCs (Menn et al., 2006).

Our work identified transcription factors/transcriptional networks in the earlier stages of the oligodendrocyte lineage that had been poorly characterized, and future mechanistic studies will aim to confirm their predicted gene regulatory control of oligodendrogenesis. Finally, given the predicted relevance of environmental signaling and transcriptional control over oligodendrocyte lineage progression, we investigated the ability of extracellular signaling events to modulate core transcriptional regulators of the oligodendrocyte GRN. Several well-established signaling pathways were predicted to stage-specifically impinge on core GRN components, suggesting a measure of environmental control over lineage progression. Thus, while SVZ-derived oligodendrocyte lineages are capable of remarkable intrinsic control over their maturation (F. Ortega et al., 2013), a series of known and novel environmental factors can confer flexibility to this process by regulating the expression of transcriptional modulators of oligodendrocyte lineage progression.

Word count: 1446

CONFLICTS OF INTEREST

The authors have no competing or financial interests to declare.

AUTHORS' CONTRIBUTIONS

KA was responsible for the conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, supervision. FC contributed to writing, validation and data analysis. MC carried out data curation, analysis, investigation validation, and methodology. RA contributed to writing, data curation, formal analysis, investigation, methodology. JV was responsible for funding acquisition, project administration, supervision; HPH: funding acquisition, supervision and project administration. OB was involved in data curation, acquisition and methodology. AB was responsible for writing, methodology, supervision and validation. HPH was responsible for funding acquisition and project administration. PK for funding acquisition, investigation, methodology, project administration, supervision, validation and writing.

OPEN RESEARCH

Scripts developed for the first time, Cytoscape files, cluster specific gene lists, gene matrices and any other raw data of this study are placed in the repository Github https://github.com/kasumaz/AdultOLgenesis. Assistance for TETRAMER usage can also be requested from Dr Marco A. Mendoza (mmendoza@genoscope.cns.fr). A user friendly web interface that allows investigators to examine gene expression across all single cells assembled will be made online upon acceptance with the

994 expertise of the lab of Julio Vera (see for example:

995 https://www.curatopes.com/melanoma/).

996

REFERENCES

Arion, D., Unger, T., Lewis, D. A., Levitt, P., & Mirnics, K. (2007). Molecular evidence for increased expression of genes related to immune and chaperone function in the prefrontal cortex in schizophrenia. *Biol Psychiatry, 62*(7), 711-721. doi: 10.1016/j.biopsych.2006.12.021

Azim, K., Akkermann, R., Cantone, M., Vera, J., Jadasz, J. J., & Kury, P. (2018). Transcriptional Profiling of Ligand Expression in Cell Specific Populations of the Adult Mouse Forebrain That Regulates Neurogenesis. *Front Neurosci, 12*, 220. doi: 10.3389/fnins.2018.00220

Azim, K., Angonin, D., Marcy, G., Pieropan, F., Rivera, A., Donega, V., . . . Raineteau, O. (2017). Pharmacogenomic identification of small molecules for lineage specific manipulation of subventricular zone germinal activity. *PLoS Biol, 15*(3), e2000698. doi: 10.1371/journal.pbio.2000698

Azim, K., Berninger, B., & Raineteau, O. (2016). Mosaic Subventricular Origins of Forebrain Oligodendrogenesis. *Front Neurosci, 10*, 107. doi: 10.3389/fnins.2016.00107

Azim, K., Fischer, B., Hurtado-Chong, A., Draganova, K., Cantu, C., Zemke, M., . . . Raineteau, O. (2014). Persistent Wnt/beta-catenin signaling determines dorsalization of the postnatal subventricular zone and neural stem cell specification into oligodendrocytes and glutamatergic neurons. *Stem Cells, 32*(5), 1301-1312. doi: 10.1002/stem.1639

Azim, K., Hurtado-Chong, A., Fischer, B., Kumar, N., Zweifel, S., Taylor, V., & Raineteau, O. (2015). Transcriptional Hallmarks of Heterogeneous Neural Stem Cell Niches of the Subventricular Zone. *Stem Cells, 33*(7), 2232-2242. doi: 10.1002/stem.2017

Azim, K., Raineteau, O., & Butt, A. M. (2012). Intraventricular injection of FGF-2 promotes generation of oligodendrocyte-lineage cells in the postnatal and adult forebrain. *GLIA, 60*(12), 1977-1990. doi: 10.1002/glia.22413

Azim, K., Rivera, A., Raineteau, O., & Butt, A. M. (2014). GSK3beta regulates oligodendrogenesis in the dorsal microdomain of the subventricular zone via Wnt-beta-catenin signaling. *GLIA, 62*(5), 778-779. doi: 10.1002/glia.22641

Basak, O., Krieger, T. G., Muraro, M. J., Wiebrands, K., Stange, D. E., Frias-Aldeguer, J., . . . Clevers, H. (2018). Troy+ brain stem cells cycle through quiescence and regulate their number by sensing niche occupancy. *Proc Natl Acad Sci U S A, 115*(4), E610-E619. doi: 10.1073/pnas.1715911114

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., & Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol, 36*(5), 411-420. doi: 10.1038/nbt.4096

Calella, A. M., Nerlov, C., Lopez, R. G., Sciarretta, C., von Bohlen und Halbach, O., Bereshchenko, O., & Minichiello, L. (2007). Neurotrophin/Trk receptor signaling mediates C/EBPalpha, -beta and NeuroD recruitment to immediate-early gene promoters in neuronal cells and requires C/EBPs to induce immediate-early gene transcription. *Neural Dev, 2*, 4. doi: 10.1186/1749-8104-2-4

Calzolari, F., Michel, J., Baumgart, E. V., Theis, F., Gotz, M., & Ninkovic, J. (2015). Fast clonal expansion and limited neural stem cell self-renewal in the adult subependymal zone. *Nat Neurosci, 18*(4), 490-492. doi: 10.1038/nn.3963

Cantone, M., Kuspert, M., Reiprich, S., Lai, X., Eberhardt, M., Gottle, P., . . . Vera, J. (2019). A gene regulatory architecture that controls region-independent dynamics of oligodendrocyte differentiation. *GLIA*. doi: 10.1002/glia.23569

Capilla-Gonzalez, V., Cebrian-Silla, A., Guerrero-Cazares, H., Garcia-Verdugo, J. M., & Quinones-Hinojosa, A. (2013). The generation of oligodendroglial cells is preserved in the rostral migratory stream during aging. *Front Cell Neurosci, 7*, 147. doi: 10.3389/fncel.2013.00147

Carlin, D. E., Demchak, B., Pratt, D., Sage, E., & Ideker, T. (2017). Network propagation in the cytoscape cyberinfrastructure. *PLoS Comput Biol, 13*(10), e1005598. doi: 10.1371/journal.pcbi.1005598

Chen, J., Van Gulden, S., McGuire, T. L., Fleming, A. C., Oka, C., Kessler, J. A., & Peng, C. Y. (2018). BMP-Responsive Protease HtrA1 Is Differentially Expressed in Astrocytes and Regulates Astrocytic Development and Injury Response. *J Neurosci, 38*(15), 3840-3857. doi: 10.1523/JNEUROSCI.2031-17.2018

Cholley, P. E., Moehlin, J., Rohmer, A., Zilliox, V., Nicaise, S., Gronemeyer, H., & Mendoza-Parra, M. A. (2018). Modeling gene-regulatory networks to describe cell fate transitions and predict master regulators. *NPJ Syst Biol Appl, 4*, 29. doi: 10.1038/s41540-018-0066-z

Crawford, A. H., Tripathi, R. B., Richardson, W. D., & Franklin, R. J. (2016). Developmental Origin of Oligodendrocyte Lineage Cells Determines Response to Demyelination and Susceptibility to Age-Associated Functional Decline. *Cell Rep*. doi: 10.1016/j.celrep.2016.03.069

Cui, X. Y., Hu, Q. D., Tekaya, M., Shimoda, Y., Ang, B. T., Nie, D. Y., . . . Xiao, Z. C. (2004). NB-3/Notch1 pathway via Deltex1 promotes neural progenitor cell differentiation into oligodendrocytes. *J Biol Chem, 279*(24), 25858-25865. doi: 10.1074/jbc.M313505200

Dulken, B. W., Buckley, M. T., Navarro Negredo, P., Saligrama, N., Cayrol, R., Leeman, D. S., . . . Brunet, A. (2019). Single-cell analysis reveals T cell infiltration in old neurogenic niches. *Nature, 571*(7764), 205-210. doi: 10.1038/s41586-019-1362-5

Dulken, B. W., Leeman, D. S., Boutet, S. C., Hebestreit, K., & Brunet, A. (2017). Single-Cell Transcriptomic Analysis Defines Heterogeneity and Transcriptional Dynamics in the Adult Neural Stem Cell Lineage. *Cell Rep, 18*(3), 777-790. doi: 10.1016/j.celrep.2016.12.060

Durinck, S., Spellman, P. T., Birney, E., & Huber, W. (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc, 4*(8), 1184-1191. doi: 10.1038/nprot.2009.97

El Waly, B., Macchi, M., Cayre, M., & Durbec, P. (2014). Oligodendrogenesis in the normal and pathological central nervous system. *Front Neurosci, 8*, 145. doi: 10.3389/fnins.2014.00145

Figueres-Onate, M., Sanchez-Villalon, M., Sanchez-Gonzalez, R., & Lopez-Mascaraque, L. (2019). Lineage Tracing and Cell Potential of Postnatal Single Progenitor Cells In Vivo. *Stem Cell Reports, 13*(4), 700-712. doi: 10.1016/j.stemcr.2019.08.010

Fiorelli, R., Azim, K., Fischer, B., & Raineteau, O. (2015). Adding a spatial dimension to postnatal ventricular-subventricular zone neurogenesis. *Development, 142*(12), 2109-2120. doi: 10.1242/dev.119966

Franz, M., Rodriguez, H., Lopes, C., Zuberi, K., Montojo, J., Bader, G. D., & Morris, Q. (2018). GeneMANIA update 2018. *Nucleic Acids Res, 46*(W1), W60-W64. doi: 10.1093/nar/gky311

Fuentealba, L. C., Rompani, S. B., Parraguez, J. I., Obernier, K., Romero, R., Cepko, C. L., & Alvarez-Buylla, A. (2015). Embryonic Origin of Postnatal Neural Stem Cells. *Cell, 161*(7), 1644-1655. doi: 10.1016/j.cell.2015.05.041

Gao, F. B., Durand, B., & Raff, M. (1997). Oligodendrocyte precursor cells count time but not cell divisions before differentiation. *Curr Biol, 7*(2), 152-155. doi: 10.1016/s0960-9822(06)00060-1

Giachino, C., Boulay, J. L., Ivanek, R., Alvarado, A., Tostado, C., Lugert, S., . . . Taylor, V. (2015). A Tumor Suppressor Function for Notch Signaling in Forebrain Tumor Subtypes. *Cancer Cell, 28*(6), 730-742. doi: 10.1016/j.ccell.2015.10.008

Grun, D., Lyubimova, A., Kester, L., Wiebrands, K., Basak, O., Sasaki, N., . . . van Oudenaarden, A. (2015). Single-cell messenger RNA sequencing reveals rare intestinal cell types. *Nature, 525*(7568), 251-255. doi: 10.1038/nature14966

He, D., Marie, C., Zhao, C., Kim, B., Wang, J., Deng, Y., . . . Lu, Q. R. (2016). Chd7 cooperates with Sox10 and regulates the onset of CNS myelination and remyelination. *Nat Neurosci, 19*(5), 678-689. doi: 10.1038/nn.4258

Herman, J. S., Sagar, & Grun, D. (2018). FateID infers cell fate bias in multipotent progenitors from single-cell RNA-seq data. *Nat Methods, 15*(5), 379-386. doi: 10.1038/nmeth.4662

Jackson, E. L., Garcia-Verdugo, J. M., Gil-Perotin, S., Roy, M., Quinones-Hinojosa, A., VandenBerg, S., & Alvarez-Buylla, A. (2006). PDGFR alpha-positive B cells are neural stem cells in the adult SVZ that form glioma-like growths in response to increased PDGF signaling. *Neuron, 51*(2), 187-199. doi: 10.1016/j.neuron.2006.06.012

Jain, R., Li, D., Gupta, M., Manderfield, L. J., Ifkovits, J. L., Wang, Q., . . . Epstein, J. A. (2015). HEART DEVELOPMENT. Integration of Bmp and Wnt signaling by Hopx specifies commitment of cardiomyoblasts. *Science, 348*(6242), aaa6071. doi: 10.1126/science.aaa6071

Jebb, D., & Hiller, M. (2018). Recurrent loss of HMGCS2 shows that ketogenesis is not essential for the evolution of large mammalian brains. *Elife, 7*. doi: 10.7554/eLife.38906

Jiao, Z., Zhang, Z. G., Hornyak, T. J., Hozeska, A., Zhang, R. L., Wang, Y., . . . Chopp, M. (2006). Dopachrome tautomerase (Dct) regulates neural progenitor cell proliferation. *Dev Biol, 296*(2), 396-408. doi: 10.1016/j.ydbio.2006.06.006

Kessaris, N., Fogarty, M., Iannarelli, P., Grist, M., Wegner, M., & Richardson, W. D. (2006). Competing waves of oligodendrocytes in the forebrain and postnatal elimination of an embryonic lineage. *Nat Neurosci, 9*(2), 173-179. doi: 10.1038/nn1620

Kim, D. Y., Hwang, I., Muller, F. L., & Paik, J. H. (2015). Functional regulation of FoxO1 in neural stem cell differentiation. *Cell Death Differ, 22*(12), 2034-2045. doi: 10.1038/cdd.2015.123

Klum, S., Zaouter, C., Alekseenko, Z., Bjorklund, A. K., Hagey, D. W., Ericson, J., . . . Bergsland, M. (2018). Sequentially acting SOX proteins orchestrate astrocyte- and oligodendrocyte-specific gene expression. *EMBO Rep, 19*(11). doi: 10.15252/embr.201846635

Liu, Z., Hu, X., Cai, J., Liu, B., Peng, X., Wegner, M., & Qiu, M. (2007). Induction of oligodendrocyte differentiation by Olig2 and Sox10: evidence for reciprocal

interactions and dosage-dependent mechanisms. *Dev Biol, 302*(2), 683-693. doi: 10.1016/j.ydbio.2006.10.007

Llorens-Bobadilla, E., Zhao, S., Baser, A., Saiz-Castro, G., Zwadlo, K., & Martin-Villalba, A. (2015). Single-Cell Transcriptomics Reveals a Population of Dormant Neural Stem Cells that Become Activated upon Brain Injury. *Cell Stem Cell, 17*(3), 329-340. doi: 10.1016/j.stem.2015.07.002

Lukaszewicz, A. I., & Anderson, D. J. (2011). Cyclin D1 promotes neurogenesis in the developing spinal cord in a cell cycle-independent manner. *Proc Natl Acad Sci U S A, 108*(28), 11632-11637. doi: 10.1073/pnas.1106230108

Marcy, G., & Raineteau, O. (2019). Concise Review: Contributions of single cell approaches for probing heterogeneity and dynamics of neural progenitors throughout life. *Stem Cells*. doi: 10.1002/stem.3071

Marques, S., van Bruggen, D., Vanichkina, D. P., Floriddia, E. M., Munguba, H., Varemo, L., . . . Castelo-Branco, G. (2018). Transcriptional Convergence of Oligodendrocyte Lineage Progenitors during Development. *Dev Cell, 46*(4), 504-517 e507. doi: 10.1016/j.devcel.2018.07.005

Marques, S., Zeisel, A., Codeluppi, S., van Bruggen, D., Mendanha Falcao, A., Xiao, L., . . . Castelo-Branco, G. (2016). Oligodendrocyte heterogeneity in the mouse juvenile and adult central nervous system. *Science, 352*(6291), 1326-1329. doi: 10.1126/science.aaf6463

Mateo, J. L., van den Berg, D. L., Haeussler, M., Drechsel, D., Gaber, Z. B., Castro, D. S., . . . Martynoga, B. (2015). Characterization of the neural stem cell gene regulatory network identifies OLIG2 as a multifunctional regulator of self-renewal. *Genome Res, 25*(1), 41-56. doi: 10.1101/gr.173435.114

Menn, B., Garcia-Verdugo, J. M., Yaschine, C., Gonzalez-Perez, O., Rowitch, D., & Alvarez-Buylla, A. (2006). Origin of oligodendrocytes in the subventricular zone of the adult brain. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 26*(30), 7907-7918.

Merkle, F. T., Fuentealba, L. C., Sanders, T. A., Magno, L., Kessaris, N., & Alvarez-Buylla, A. (2014). Adult neural stem cells in distinct microdomains generate previously unknown interneuron types. *Nat Neurosci, 17*(2), 207-214. doi: 10.1038/nn.3610

Merkle, F. T., Mirzadeh, Z., & Alvarez-Buylla, A. (2007). Mosaic organization of neural stem cells in the adult brain. *Science, 317*(5836), 381-384.

Mi, H., Poudel, S., Muruganujan, A., Casagrande, J. T., & Thomas, P. D. (2016). PANTHER version 10: expanded protein families and functions, and analysis tools. *Nucleic Acids Res, 44*(D1), D336-342. doi: 10.1093/nar/gkv1194

Mirzadeh, Z., Doetsch, F., Sawamoto, K., Wichterle, H., & Alvarez-Buylla, A. (2010). The subventricular zone en-face: wholemount staining and ependymal flow. *J Vis Exp*(39). doi: 10.3791/1938

Mizrak, D., Levitin, H. M., Delgado, A. C., Crotet, V., Yuan, J., Chaker, Z., . . . Doetsch, F. (2019). Single-Cell Analysis of Regional Differences in Adult V-SVZ Neural Stem Cell Lineages. *Cell Rep, 26*(2), 394-406 e395. doi: 10.1016/j.celrep.2018.12.044

Moore, L., Bain, J. M., Loh, J. M., & Levison, S. W. (2014). PDGF-responsive progenitors persist in the subventricular zone across the lifespan. *ASN Neuro, 6*(2). doi: 10.1042/AN20120041

Nakatani, H., Martin, E., Hassani, H., Clavairoly, A., Maire, C. L., Viadieu, A., . . . Parras, C. (2013). Ascl1/Mash1 promotes brain oligodendrogenesis during

myelination and remyelination. *J Neurosci, 33*(23), 9752-9768. doi: 10.1523/JNEUROSCI.0805-13.2013

Obernier, K., Cebrian-Silla, A., Thomson, M., Parraguez, J. I., Anderson, R., Guinto, C., . . . Alvarez-Buylla, A. (2018). Adult Neurogenesis Is Sustained by Symmetric Self-Renewal and Differentiation. *Cell Stem Cell, 22*(2), 221-234 e228. doi: 10.1016/j.stem.2018.01.003

Oki, S., Ohta, T., Shioi, G., Hatanaka, H., Ogasawara, O., Okuda, Y., . . . Meno, C. (2018). ChIP-Atlas: a data-mining suite powered by full integration of public ChIP-seq data. *EMBO Rep, 19*(12). doi: 10.15252/embr.201846255

Ortega, F., Gascon, S., Masserdotti, G., Deshpande, A., Simon, C., Fischer, J., . . . Berninger, B. (2013). Oligodendrogliogenic and neurogenic adult subependymal zone neural stem cells constitute distinct lineages and exhibit differential responsiveness to Wnt signalling. *Nat Cell Biol, 15*(6), 602-613. doi: 10.1038/ncb2736

Ortega, F., Gascón, S., Masserdotti, G., Deshpande, A., Simon, C., Fischer, J., . . . Berninger, B. (2013). Oligodendrogliogenic and neurogenic adult subependymal zone neural stem cells constitute distinct lineages and exhibit differential responsiveness to Wnt signalling. *Nat Cell Biol, 15*(6), 602-613. doi: 10.1038/ncb2736

Poth, J. M., Brodsky, K., Ehrentraut, H., Grenz, A., & Eltzschig, H. K. (2013). Transcriptional control of adenosine signaling by hypoxia-inducible transcription factors during ischemic or inflammatory disease. *J Mol Med (Berl), 91*(2), 183-193. doi: 10.1007/s00109-012-0988-7

Pozniak, C. D., Langseth, A. J., Dijkgraaf, G. J., Choe, Y., Werb, Z., & Pleasure, S. J. (2010). Sox10 directs neural stem cells toward the oligodendrocyte lineage by decreasing Suppressor of Fused expression. *Proc Natl Acad Sci U S A, 107*(50), 21795-21800. doi: 10.1073/pnas.1016485107

Rosmaninho, P., Mukusch, S., Piscopo, V., Teixeira, V., Raposo, A. A., Warta, R., . . . Castro, D. S. (2018). Zeb1 potentiates genome-wide gene transcription with Lef1 to promote glioblastoma cell invasion. *Embo J, 37*(15). doi: 10.15252/embj.201797115

Sher, F., Rossler, R., Brouwer, N., Balasubramaniyan, V., Boddeke, E., & Copray, S. (2008). Differentiation of neural stem cells into oligodendrocytes: involvement of the polycomb group protein Ezh2. *Stem Cells, 26*(11), 2875-2883. doi: 10.1634/stemcells.2008-0121

Simion, C., Cedano-Prieto, M. E., & Sweeney, C. (2014). The LRIG family: enigmatic regulators of growth factor receptor signaling. *Endocr Relat Cancer, 21*(6), R431-443. doi: 10.1530/ERC-14-0179

Zeisel, A., Munoz-Manchado, A. B., Codeluppi, S., Lonnerberg, P., La Manno, G., Jureus, A., . . . Linnarsson, S. (2015). Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science, 347*(6226), 1138-1142. doi: 10.1126/science.aaa1934

Zweifel, S., Marcy, G., Lo Guidice, Q., Li, D., Heinrich, C., Azim, K., & Raineteau, O. (2018). HOPX Defines Heterogeneity of Postnatal Subventricular Zone Neural Stem Cells. *Stem Cell Reports, 11*(3), 770-783. doi: 10.1016/j.stemcr.2018.08.006

References Word Count: 2244

FIGURE LEGENDS

Figure 1

Identification of lineage-specific NSCs/NPs. (a) tSNE plot of 1200 single cells (including microglia, choroid plexus, neurons and vascular cells; see Figure S1) expressing the transcription factors *Olig1*, *Olig2* and *Sox10*. PCA color plot of transcript expression for determining cells with elevated expression of core oligodendrocyte lineage markers. (b) Selected early oligodendrocyte lineage cell expression of pallial marker *Pax6* and subpallial markers *Nkx2.1* and *Gsx2* together with the NSC marker *Hes5* for determining a potential dorsal origin of oligodendrocyte lineage. (c) FateID pseudotime demonstrates that the mature oligodendrocyte (MOL) transcriptional trajectory branches off only at the very earliest stages in the neuronal lineage, whereas among identified oligodendrocyte cells mature oligodendrocytes branch via OPCs that are in close proximity to OLTAPs and OLaNSCs. (d) tSNE plots of proneuronal and proolidendroglial lineage cells and selected transcript expression in the early- and mid-stages of the oligodendrocyte lineage. e Heatmap of the 13 analyzed clusters illustrating the highly enriched stage-specific signatures. Stage color legend is presented in f. (f) Violin plots of selected marker expression across the 13 clusters further demonstrating markers specific/enriched in stage and lineage. Abbreviations: aNSC: activated neural stem cell; MOL: mature oligodendrocyte; NB: neuroblast; NSC: neural stem cell; OLaNSC: oligodendroglial activated neural stem cell; OLpNSC: oligodendroglial primed neural stem cell; OLqNSCI: oligodendroglial quiescent neural stem cell (subtype I); OLqNSCI: oligodendroglial quiescent neural stem cell (subtype II); OLTAP:

oligodendroglial transiently amplifying progenitor; OPC: oligodendrocyte precursor cell; pNSC: primed neural stem cell; qNSCI: quiescent neural stem cell (subtype I); qNSCII: quiescent neural stem cell (subtype II).

Figure 2

Pathway and mechanistic analysis of the oligodendrocyte lineage. (a) Dot plot of the most highly enriched/significant ("Sig" in plot) signaling pathways, organized alphabetically from top to bottom in the 2 lineages. (b) A time course of the top 5 most enriched biological processes in each of the 7 clusters in the oligodendrocyte lineage. Point size in plots reflects significance. Sig = significance (*p*-value). Abbreviations: aNSC: activated neural stem cell; MOL: mature oligodendrocyte; NB: neuroblast; NSC: neural stem cell; OPC: oligodendrocyte precursor cell; pNSC: primed neural stem cell; qNSCI: quiescent neural stem cell (subtype I); qNSCII: quiescent neural stem cell (subtype II).

Figure 3

Cell cycle heterogeneity of oligodendroglial and neuronal lineage cells. (a) Oligodendrocyte and neuronal lineage cells scored by Seurat for cell cycle regression using canonical markers. Oligodendrocyte lineage cells are plotted according to their pseudotime trajectory and overlaid displayed the phase of cycle of each cell. (b) A summary displaying the percentage of cells in the 3 phases of cell cycle and the dynamics of mid stages of either lineage shuttling between states. (c) Ridge plot of oligodendrocyte and neuronal lineage cells illustrating the distribution of signal in the S or G2M cell cycle phase. (d) The mid stages of oligodendroglial (light

red) and neuronal lineage (light blue) expanded for S (*Pcna, Mcm6*) and G2M (*Top2a, Mki67*) marker signal distribution in ridge plots. (e,f) Dot plots show pathway and protein class identification of the mid stages in the oligodendrocyte lineage. Green, light red and brown in OLaNSC labels quiescent, common to both stages and S phase, respectively; light purple, light red and light brown label OLTAPs in G2M, common to both and S phase, respectively. Terms are ranked according to their highest significance (Sig; *p*-value) from top to bottom. Pathways in red font are examined for revealing their signaling-to-transcriptional networks in Figure 7. Abbreviations: aNSC: activated neural stem cell; MOL: mature oligodendrocyte; NB: neuroblast; NSC: neural stem cell; OLaNSC: oligodendroglial activated neural stem cell; OLpNSC: oligodendroglial primed neural stem cell; OLqNSCI: oligodendroglial quiescent neural stem cell (subtype I); OLqNSCI: oligodendroglial quiescent neural stem cell (subtype II); OLTAP: oligodendroglial transiently amplifying progenitor; OPC: oligodendrocyte precursor cell; pNSC: primed neural stem cell; qNSCI: quiescent neural stem cell (subtype I); qNSCII: quiescent neural stem cell (subtype); TF: transcription factor.

Figure 4

Reconstruction of a GRN by merging of previously generated networks and organization of stages in the oligodendrocyte lineage according to their pseudotime coordinates. Networks assembled in Figure S6b and d, those published recently (see description in results section) and additional transcription factor-target gene (TG) interactions from CHIP-seq studies are processed for a larger network. transcription factor expressed in 7 distinct stages are arranged according to their pseudotime order and each node represents the relative numbers of transcription factors

expressed in a given phase in the lineage. A separate grouping of transcription factors expressed in both aNSCs and TAPs are plotted according to the average trajectory coordinates between these 2 phases. The top 5 GO Biological Processes in OPCs and mature oligodendrocytes are incorporated for identifying their regulatory pathways. Irregularly expressed transcription factors are those which are abundant in the neuronal lineage and additionally expressed along the oligodendrocyte lineage. The entire GRN of the oligodendrocyte lineage, termed as "AdultOLgenesis GRN". Nodes in bold outer surface are characterized by high protein-protein interactions. Green, red and grey interactions signify gene inhibition, activation and unspecified interaction onto their target genes (TGs), respectively. Legends for the AdultOLgenesis GRN are presented in Figure 5 and 6. Abbreviations: aNSC: activated neural stem cell; GRN: gene regulatory network; MOL: mature oligodendrocyte; NB: neuroblast; NSC: neural stem cell; OLaNSC: oligodendroglial activated neural stem cell; OLpNSC: oligodendroglial primed neural stem cell; OLqNSCI: oligodendroglial quiescent neural stem cell (subtype I); OLqNSCI: oligodendroglial quiescent neural stem cell (subtype II); OLTAP: oligodendroglial transiently amplifying progenitor; OPC: oligodendrocyte precursor cell; TF: transcription factor; TG: target gene.

Figure 5

Gene regulatory interactions in substages of the oligodendrocyte lineage and the predicted effects of key transcription factors in each stage. (a-f) Summaries of the GRNs in each stage showing downstream gene regulation propagation. Edges and nodes reflect the relative numbers of gene regulatory interactions and numbers of transcription factor expressed in each stage, respectively. (g-l) The downstream

gene regulatory interactions of selected transcription factors are quantified by propagating the transcription factor most highly ranked in the oligodendrocyte lineage and a transcription factor that is commonly expressed in a defined stage. In g, *Olig2* and *Sox10* are probed for their downstream target (designated as secondary, 2°), the subsequent downstream targets (including both genes in the GO Biological Processes and additional transcription factors) to the tertiary genes (3°). The transcription factor-target interactions quantified are plotted as a percentage versus the total numbers of genes/gene regulatory interactions from the entire reconstructed GRN. For each stage, the 2 highly ranked transcription factors and their direct target genes are shown as subnetworks from the entire GRN. Abbreviations: aNSC: activated neural stem cell; GRN: gene regulatory network; MOL: mature oligodendrocyte; NB: neuroblast; NSC: neural stem cell; OLaNSC: oligodendroglial activated neural stem cell; OLpNSC: oligodendroglial primed neural stem cell; OLqNSCI: oligodendroglial quiescent neural stem cell (subtype I); OLqNSCI: oligodendroglial quiescent neural stem cell (subtype II); OLTAP: oligodendroglial transiently amplifying progenitor; OPC: oligodendrocyte precursor cell; TF: transcription factor; TG: target gene.

Figure 6

Mapping the impact of Signaling/PANTHER pathways onto GRNs. (a) PANTHER pathways from previous GO analyses were used to filter out transcription factors associated with each pathway term and their propagation onto the GRN were quantified and represented as a percentage compared to *Olig2*/*Sox10* (Figure 5g). (b-g) Examples of PANTHER Pathway effects onto the reconstructed GRN where dashed edges show the secondary effects of propagation compared to the direct

effects of transcription factors associated with each pathway. Heatmaps show the relative ranking of key transcription factors of each pathway (both direct and secondary targets). Transcription factors highlighted in red, green and grey are those that activate expression of genes in later stage oligodendrocytes, inhibit the expression of genes in later stage oligodendrocytes and transcription factors that are essential to pathways as general regulators, respectively. Abbreviations: aNSC: activated neural stem cell; GRN: gene regulatory network; MOL: mature oligodendrocyte; NB: neuroblast; NSC: neural stem cell; OLaNSC: oligodendroglial activated neural stem cell; OLpNSC: oligodendroglial primed neural stem cell; OLqNSCI: oligodendroglial quiescent neural stem cell (subtype I); OLqNSCI: oligodendroglial quiescent neural stem cell (subtype II); OLTAP: oligodendroglial transiently amplifying progenitor; OPC: oligodendrocyte precursor cell; TF: transcription factor; TG: target gene.

Word Count: 1333

Funding

This work was supported by the German Research Council (DFG; SPP1757/KU1934/2_1, KU1934/5-1; AZ/115/1-1/Ve642/1-1), the German Academic Exchange Service (DAAD), the Swiss National Funds (P300PA_171224). PK: This study was supported by the Deutsche Forschungsgemeinschaft (DFG; AZ115/1-1). Research on myelin repair and neuroregeneration has also been supported by the Christiane and Claudia Hempel Foundation for clinical stem cell research, DMSG Ortsvereinigung Düsseldorf und Umgebung e.V., Stifterverband/Novartisstiftung and the James and Elisabeth Cloppenburg, Peek & Cloppenburg Düsseldorf Stiftung.

The MS Center at the Department of Neurology is supported in part by the Walter and Ilse Rose Foundation.FC was supported in part by NEURON-ERANET (01EW1604) and DFG (CRC1080) awards to Prof Benedikt Berninger (UMC Mainz) and by the Inneruniversitäre Forschungsförderung of the University Medical Center of Johannes Gutenberg University Mainz. AB was supported by grants from the MS Society UK, BBSRC and MRC.

List of abbreviations

aNSC: activated neural stem cell; CNS: central nervous system; DEGs: differentially expressed genes; GO: gene ontology GRN: gene regulatory network; MOL: mature oligodendrocyte; NB: neuroblast; NSC: neural stem cell; oligodendrocyte: oligodendrocyte; OLaNSC: oligodendroglial activated neural stem cell; OLpNSC: oligodendroglial primed neural stem cell; OLqNSCI: oligodendroglial quiescent neural stem cell (subtype I); OLqNSCI: oligodendroglial quiescent neural stem cell (subtype II); OLTAP: oligodendroglial transiently amplifying progenitor; OPC: oligodendrocyte precursor cell; PANTHER: protein annotation through evolutionary relationship; PCA: principal component analysis; pNSC: primed neural stem cell; qNSCI: quiescent neural stem cell (subtype I); qNSCII: quiescent neural stem cell (subtype II); SVZ: subventricular zone; TAP: transiently amplifying progenitor; TETRAMER: TEmporal TRAnscription regulation ModellER TF: transcription factor; TG: target gene; tSNE: t-Distributed Stochastic Neighbor Embedding.

Figure 1: Identification of lineage-specific NSCs/NPs.

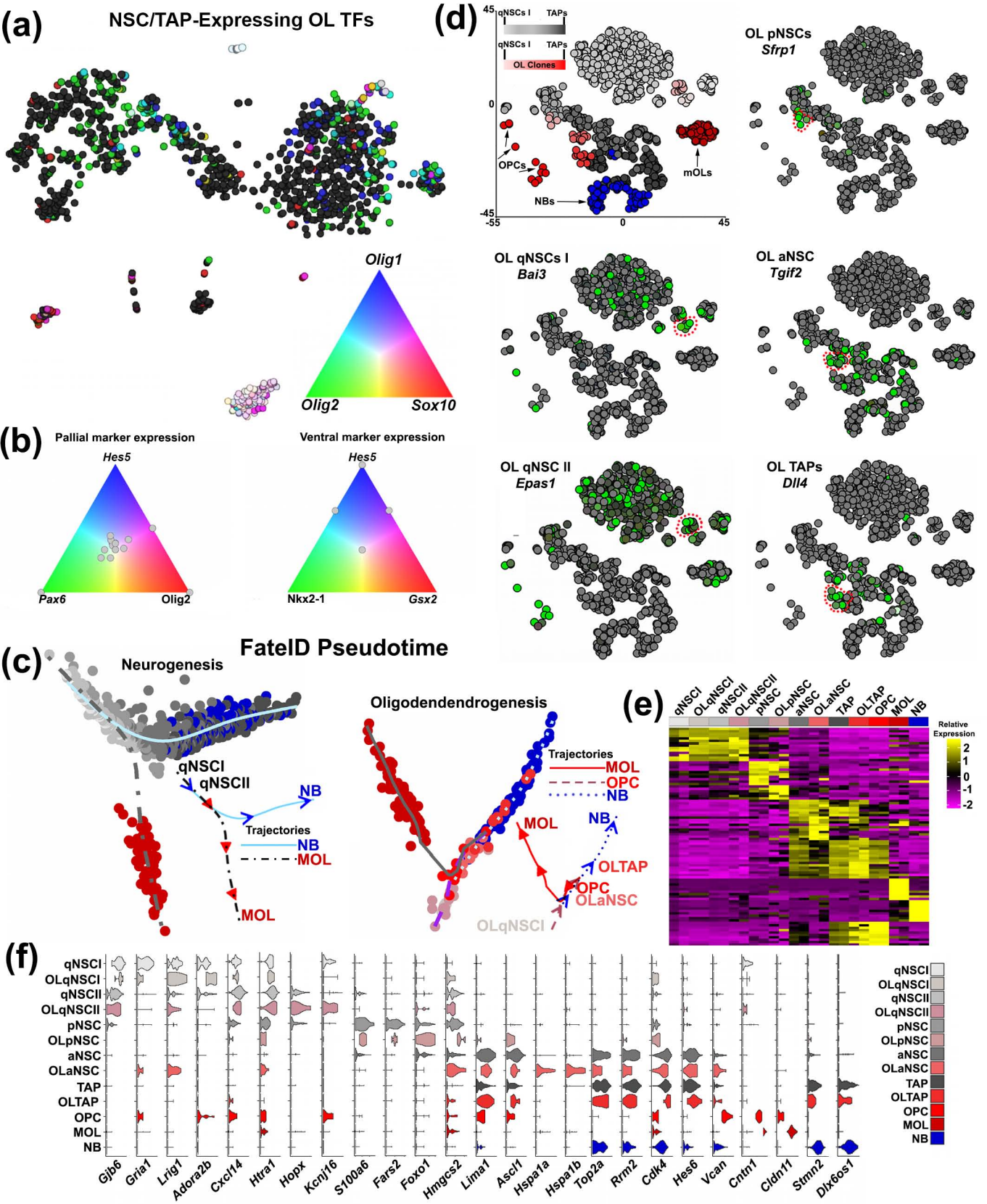
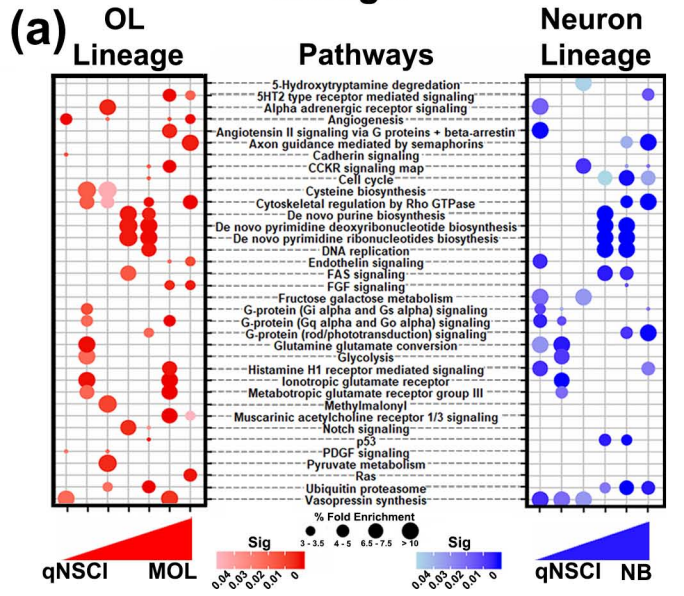


Figure 2: Pathway and mechanistic analysis of the oligodendrocyte lineage.



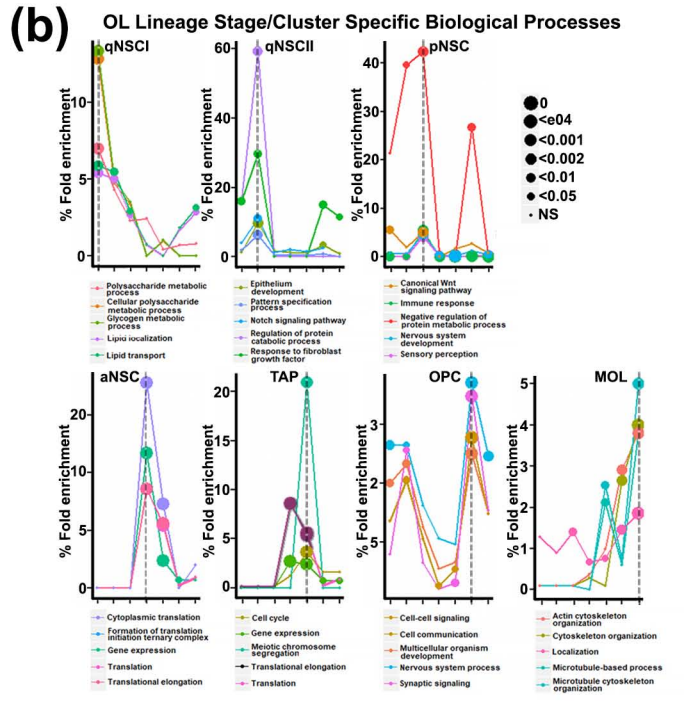


Figure 3: Cell cycle heterogeneity of oligodendroglial and neuronal lineage cells.

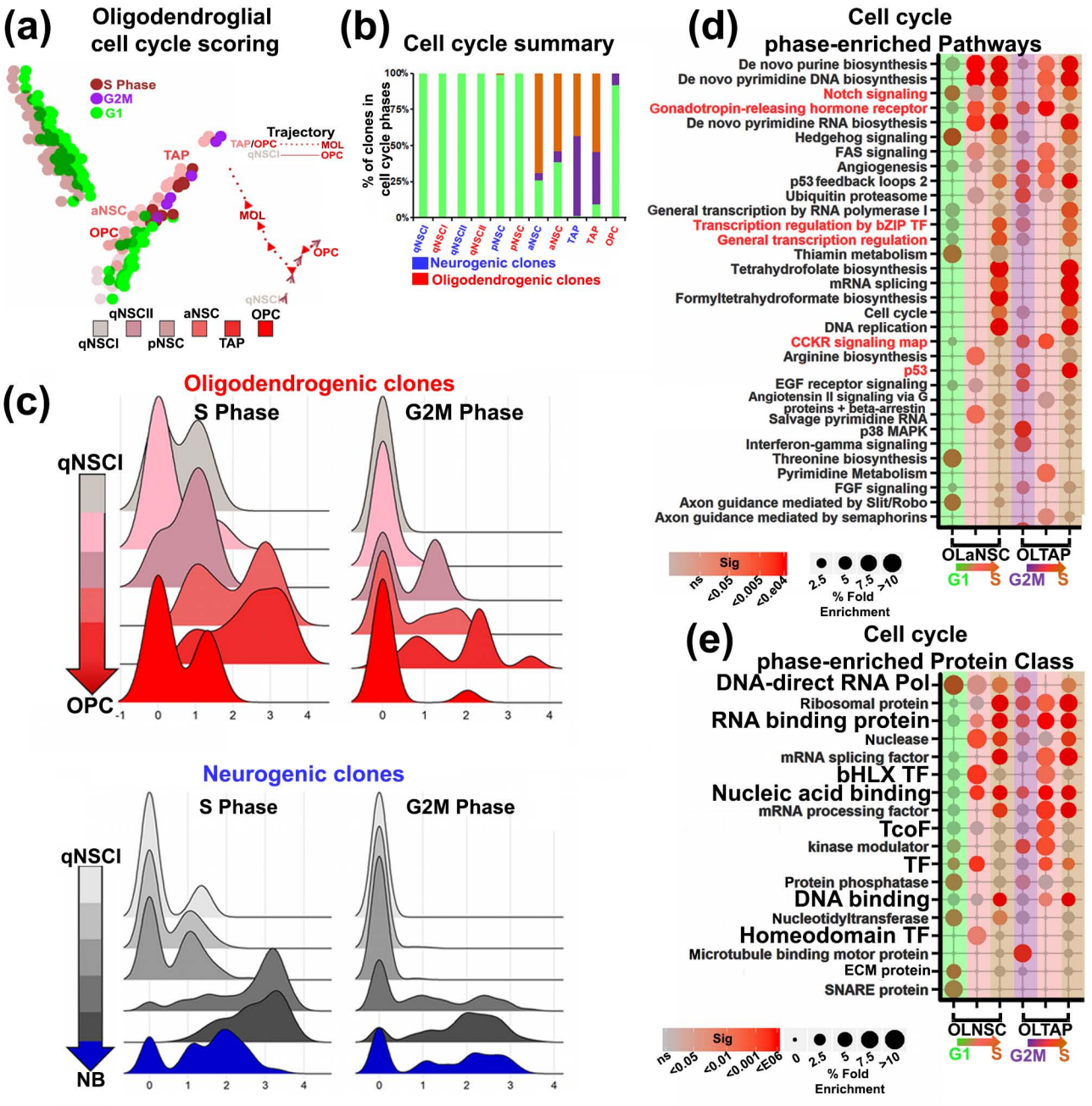


Figure 4: Reconstruction of a GRN by merging of previously generated networks and organization of stages in the oligodendrocyte lineage according to their pseudotime coordinates.

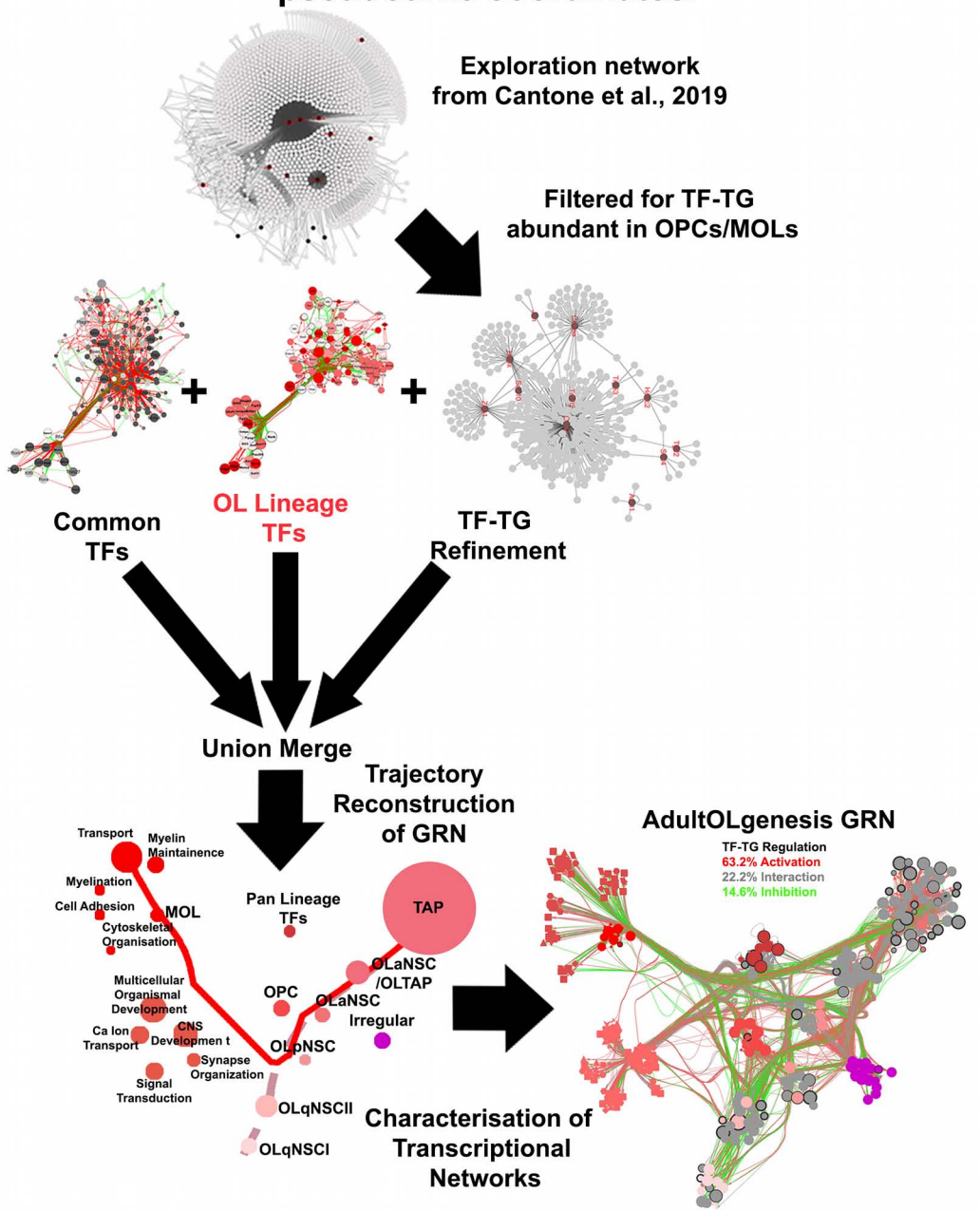


Figure 5: Gene regulatory interactions in substages of the oligodendrocyte lineage and the predicted effects of key transcription factors in each stage.

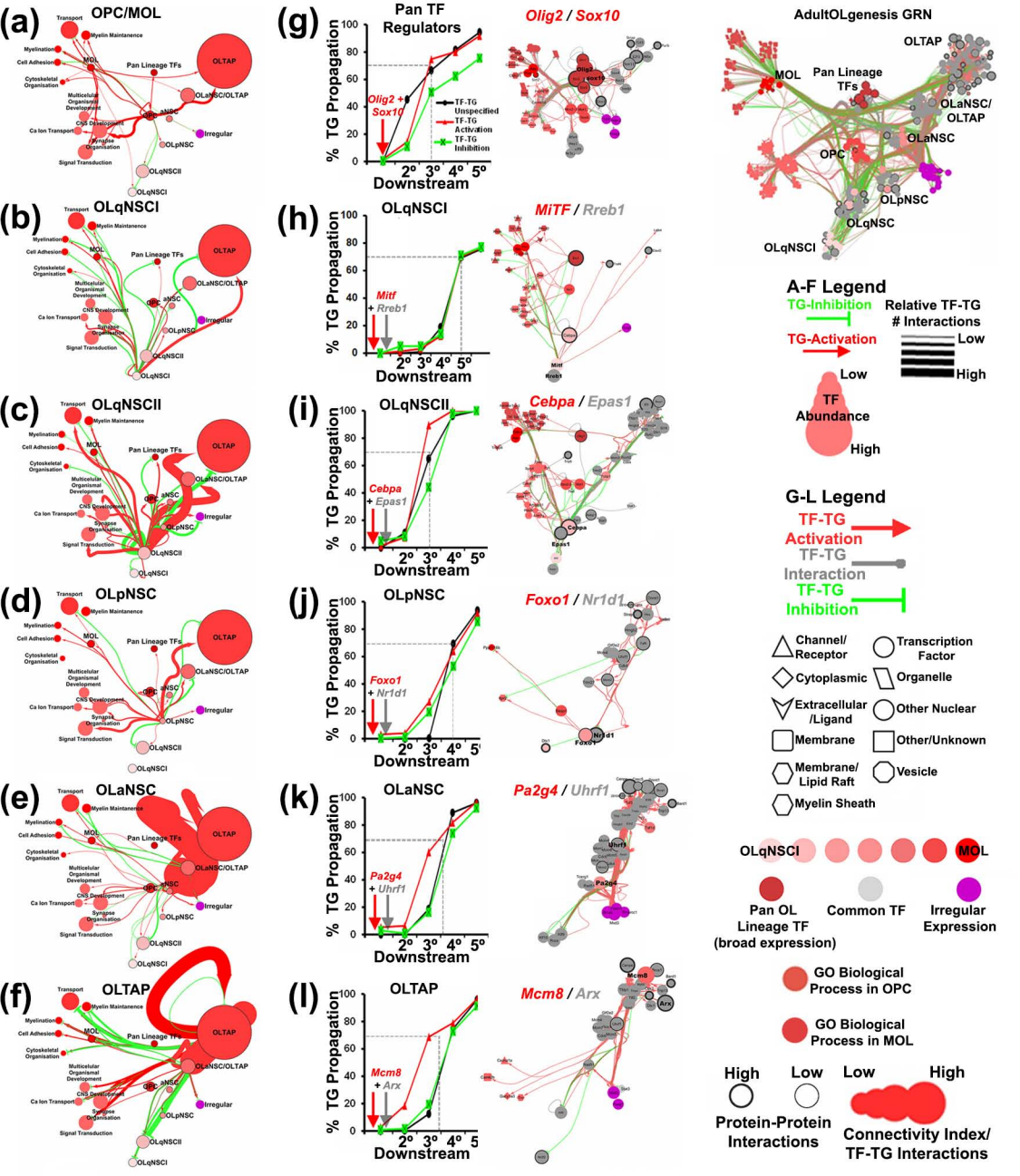
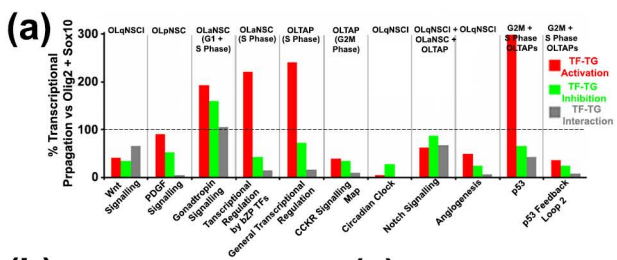
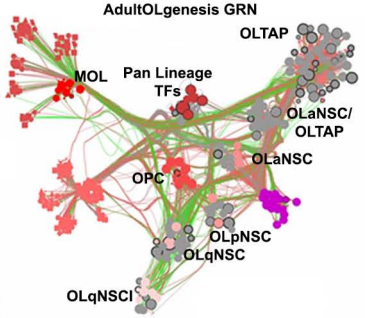
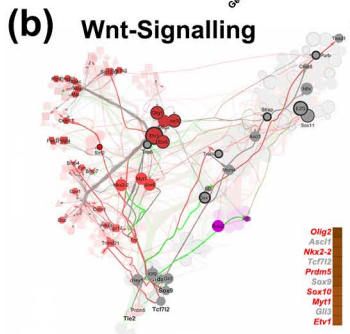
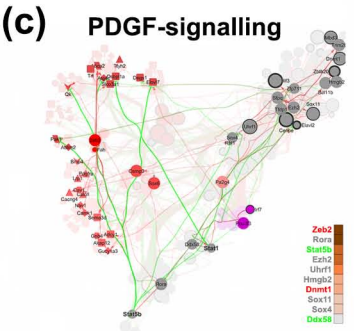


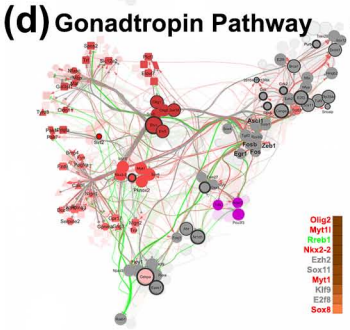
Figure 6: Mapping the impact of Signaling/PANTHER pathways onto GRNs.

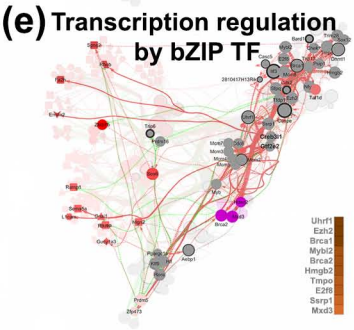


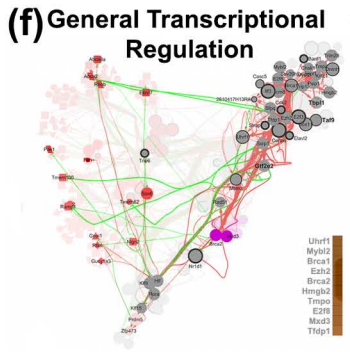


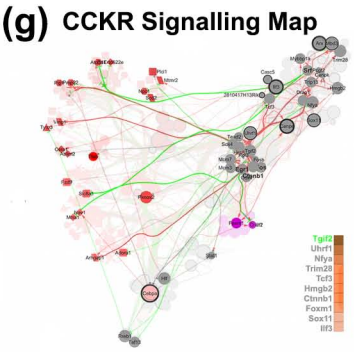


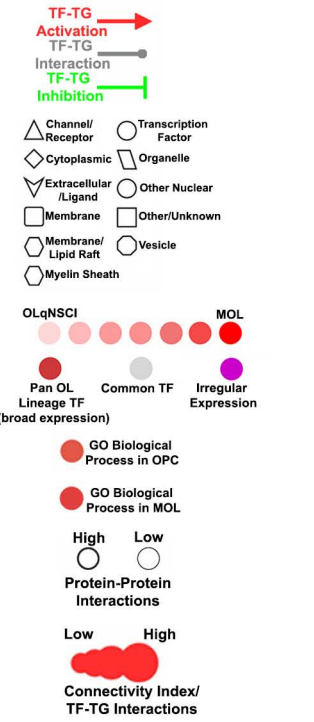












Supporting Information

Figure S1:
Summary of the datasets processed and overview of the transcription factor expression comparisons with postnatal bulk and adult single analysis. (a) Datasets re-processed illustrating the transgenic reporter and immunofluorescence strategy for isolating cells derived from the adult SVZ. Processed cells are shown as a tSNE plot. (b) Heatmap demonstrating that transcription factors expressed in single cells of OLqNSCI, OLqNSCII and OLpNSC (single cell early stage oligodendrocytes) group closely with postnatal dorsal bulk datasets. Lists of differentially expressed transcription factor-encoding genes from the indicated analyses were rendered as lists of binary values to allow grouping. (c) Clustergram illustrating transcriptomic differences and similarities of datasets generated from bulk versus single cell sequencing. Abbreviations: aNSC: activated neural stem cell; dTAPs: dorsal TAPs; dSVZ: dorsal subventricular zone; N: neuronal; NB: neuroblast; NSC: neural stem cell; oligodendrocyte: oligodendrocyte; OL: oligodendroglial; OPC: oligodendrocyte precursor cell; PANTHER: protein annotation through evolutionary relationship; PCA: principal component analysis; qNSC: quiescent neural stem cells; TAP: transiently amplifying progenitor.

Figure S2:
Stage and cell specific marker expression across single cells in tSNE plots. (a) Constant *Gapdh* expression levels across cell types studied. (b-g) Selected markers of stage-specific markers, including pan-early NSC populations (*Hes5* qNSC-pNSC) and neuroblasts (*Dcx*). (i-l) Landmark markers of oligodendrocytes. Abbreviations: mOL: mature oligodendrocyte; NSC: neural stem cell; OPC: oligodendrocyte precursor cell; pNSC: primed neural stem cell; qNSCI: quiescent neural stem cell (subtype I); qNSCII: quiescent neural stem cell (subtype II);TAP: transiently amplifying progenitor.

Figure S3:
Stage-specific comparisons of oligodendrocyte lineage versus neuronal lineage for PANTHER Protein Class and Pathways. (a-e) Cluster enriched genes for the oligodendroglial (red throughout) and neuronal linage (blue throughout) for each stages analysed were compared for Panther Protein Class and Pathways. The overlapping genes (grey throughout) in each stage were included in the analysis. Ontology terms enriched in the oligodendrocyte lineage, neuronal lineage and common to both lineages are highlighted in red, blue and black text, respectively. Abbreviations: NSC: neural stem cell; pNSC: primed neural stem cell; qNSCI: quiescent neural stem cell (subtype I); qNSCII: quiescent neural stem cell (subtype II);TAP: transiently amplifying progenitor; TF: transcription factor.

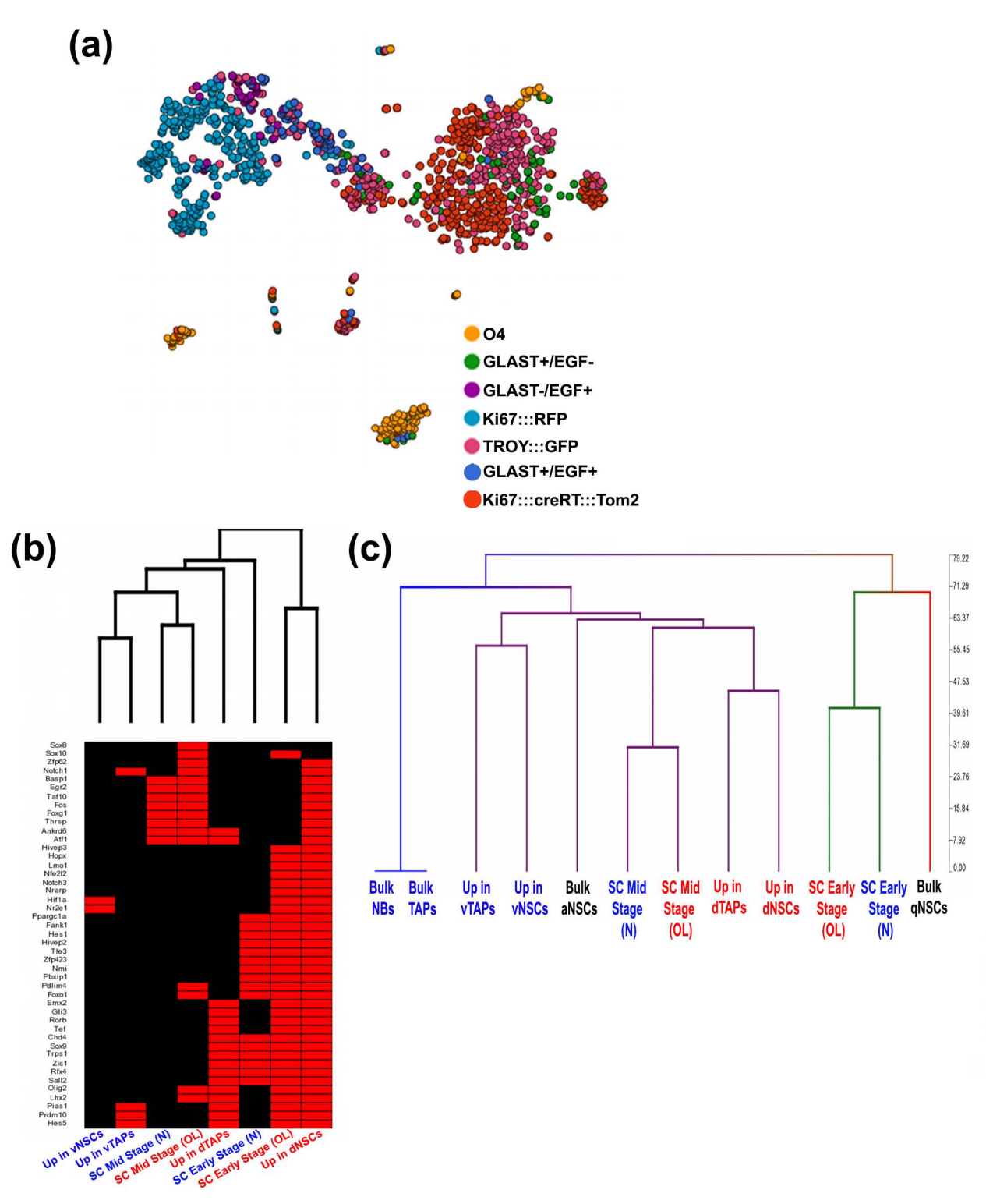
Figure S4:
Temporal PANTHER Pathway enrichment in the oligodendrocyte lineage. A time course of selected pathways enriched in the oligodendrocyte lineage from the PANTHER analysis shows enrichment of mechanisms during the course of oligodendrogenesis. Red, grey and blue colours in Venn diagrams signify, oligodendroglial, intersect (common), and neuronal lineage cells, respectively. A few pathways such as Angiogenesis and Notch signalling for example show enrichment in defined stages. Terms have been shorted/abbreviated to fit.

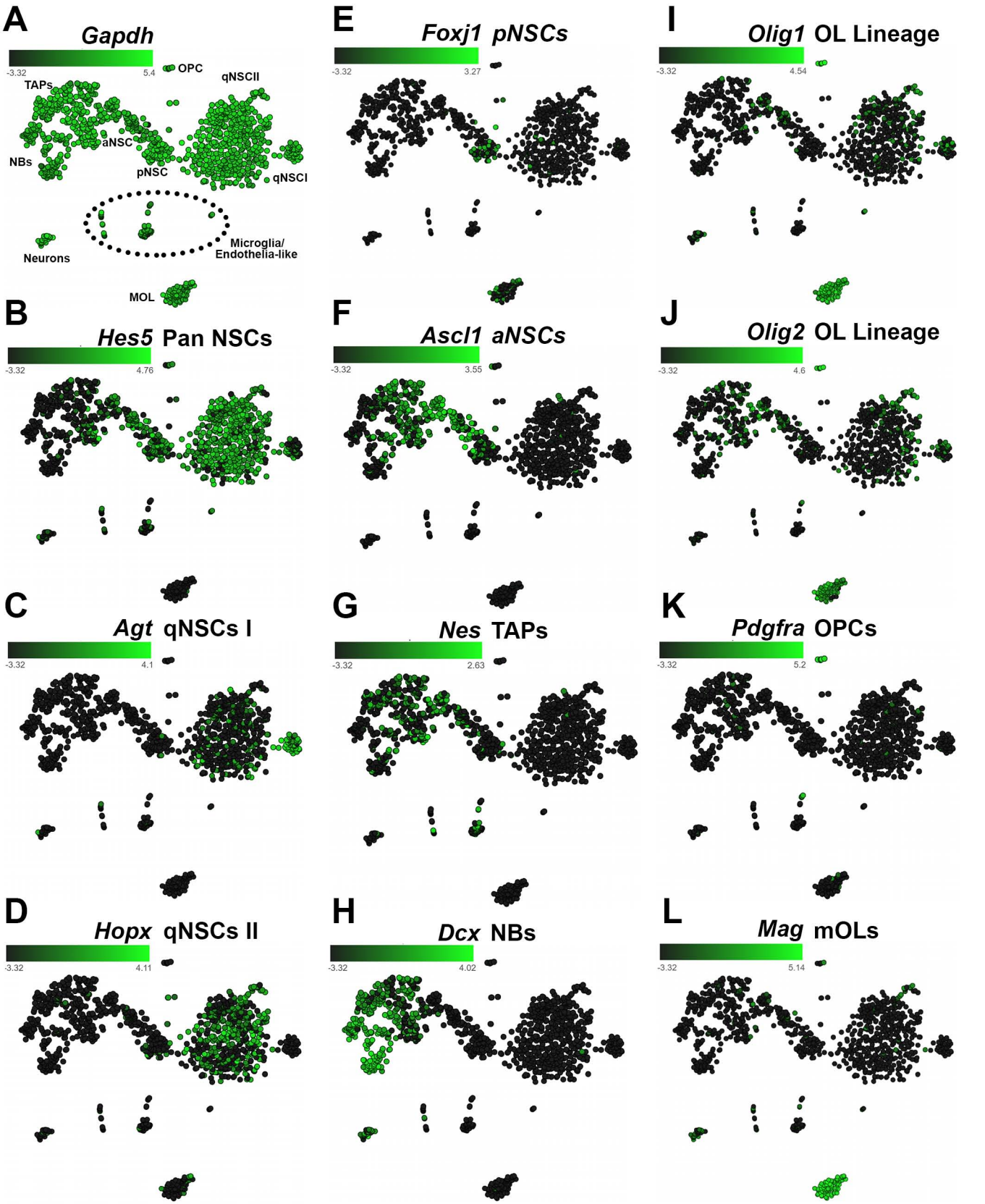
Figure S5:
Temporal Transcriptional Regulation in the oligodendrocyte lineage. Protein class terms abundant in the 3 stages in Figure S4 are expanded further as a heatmap for representing modes of transcription factors regulating the oligodendrocyte lineage. TF: transcription factor.

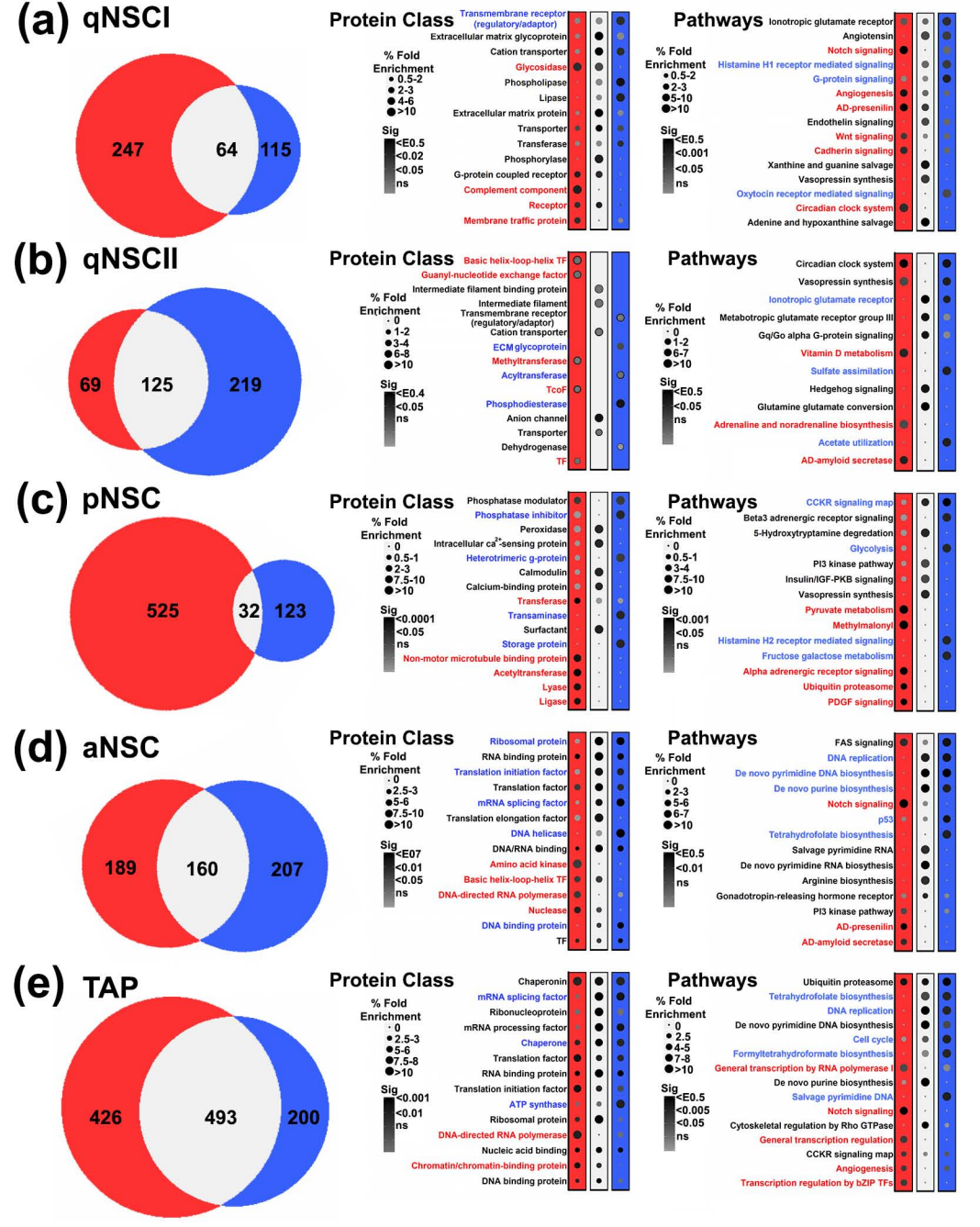
Figure S6:
Transcription factor expression in oligodendroglia points to shifts in gene regulatory network states as major drivers of differentiation. (a) Transcription factors expressed in the oligodendrocyte lineage (light red from OLqNSCI to dark red for mature oligodendrocytes) and expressed in both lineages, i.e. commonly expressed transcription factors between the 2 lineages (light grey from qNSCI to dark grey for TAPs) are ranked according to their heat diffusion impact for downstream genetic interactions, physical interactions with other transcription factors and combined. (b-d) Using gene lists for the defined stages for transcription factor expression, those transcription factors expressed in both lineages from qNSCI to TAP stage (b), enriched in the neuronal lineage from qNSCI to the NB stage e and oligodendroglial lineage from OLqNSCI to OPC stage (d). The size of each transcription factor node in the network is relative to its transcription factor-TG (target gene) connectivity index. The numbers of activating (red) or inhibiting (green) interactions from each network are shown as a percentage. (e) The top 25 transcription factors with the highest connectivity index are displayed as heatmaps for the numbers of regulatory interactions with TGs. Heatmaps show darker to lighter colour relative to the numbers of genes regulated by each transcription factor. (f) Examples of higher-, medium- and low-ranking transcription factors. See Figure 4d for description of the legend. Abbreviations: MOL: mature oligodendrocyte; NB: neuroblast; NSC: neural stem cell; OL: oligodendrocyte; OLaNSC: oligodendroglial activated neural stem cell; OLpNSC: oligodendroglial primed neural stem cell; OLqNSCI: oligodendroglial quiescent neural stem cell (subtype I); OLqNSCI: oligodendroglial quiescent neural stem cell (subtype II); OLTAP: oligodendroglial transiently amplifying progenitor; OPC: oligodendrocyte precursor cell; pNSC: primed neural stem cell; qNSCI: quiescent neural stem cell (subtype I); TAP: transiently amplifying progenitor; transcription factor; TG: target gene.

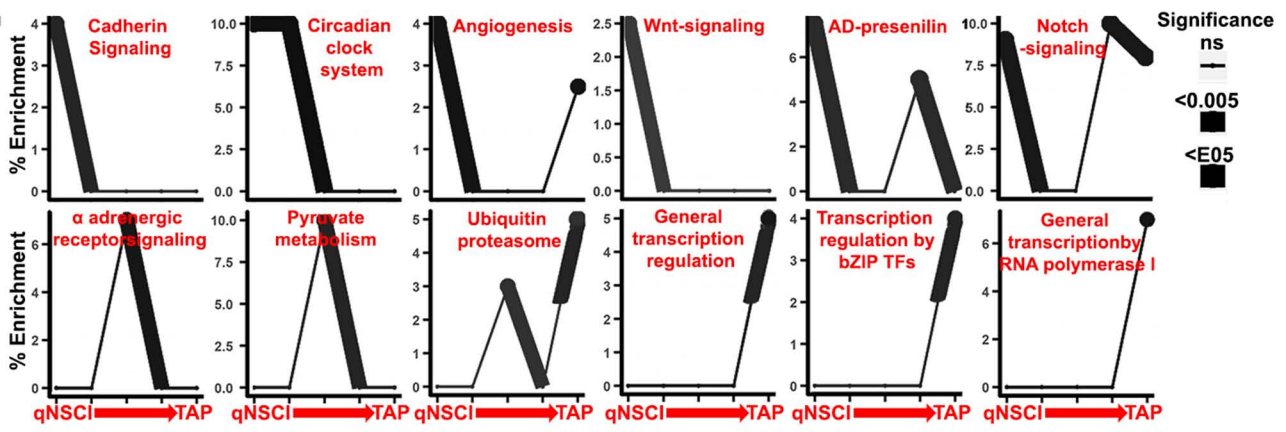
Figure S7:
Diffusion ranking for transcription factor functional relationships. (a) Transcription factors expressed in the oligodendrocyte lineage from OLqNSC to MOL stage and those that are common to both oligodendroglial and neuronal lineages were first converted to human gene symbols for maximising the data gathered (information on human gene orthologs are more numerous). (b,c) GeneMania was used to gather functional information on genetic and physical interactions (protein-protein interactions) for the genes sampled. All genes were initially assembled on a network without prior arrangement and the heat diffusion algorithm applied separately for genetic and physical interactions. Node sizes in the networks reflect diffusion rankings for the two combined parameters. Network organized using the CoSE algorithm for force directing nodes within the defined stages. (d) Grading of transcription factors within each stage. Larger to smaller node sizes are relative to the diffusion ranking scores of combined genetic and physical diffusion. Border thickness of each transcription factor directly proportional to protein-protein

interaction number. Transcription factors labels are shown for those prioritised and additional transcription factors shown are expressed in the oligodendrocyte lineage with significant expression, whilst the remaining are derived from more stringent criteria. Abbreviations: MOL: mature oligodendrocyte; NSC: neural stem cell; OL: oligodendrocyte: OLaNSC: oligodendroglial activated neural stem cell; OLpNSC: oligodendroglial primed neural stem cell; OLqNSCI: oligodendroglial quiescent neural stem cell (subtype I); OLqNSCI: oligodendroglial quiescent neural stem cell (subtype II); OLTAP: oligodendroglial transiently amplifying progenitor; OPC: oligodendrocyte precursor cell; pNSC: primed neural stem cell; qNSCI: quiescent neural stem cell (subtype I); TAP: transiently amplifying progenitor; transcription factor.

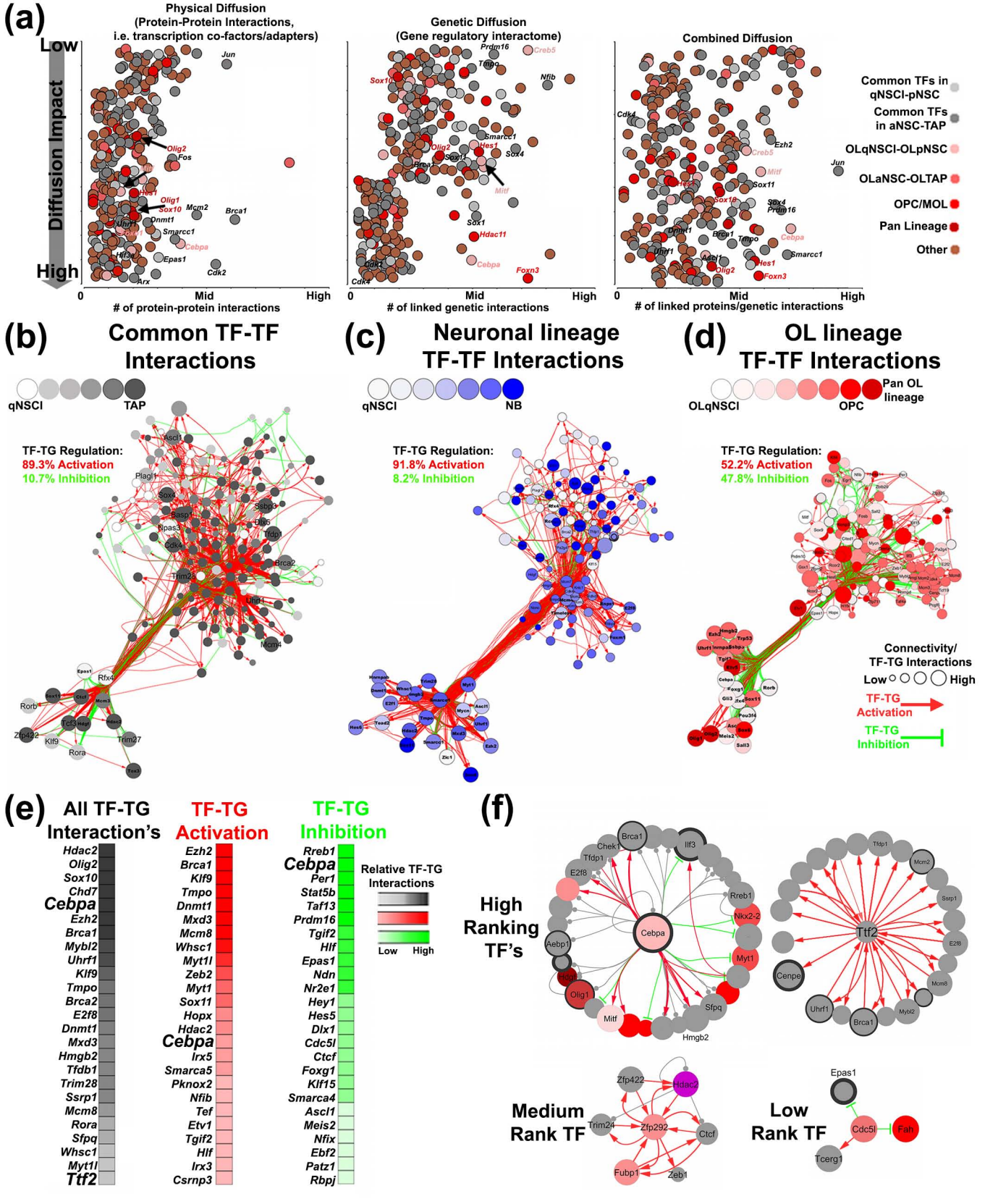

(a)
O4
GLAST+/EGF-
GLAST-/EGF+
Ki67:::RFP
TROY:::GFP
GLAST+/EGF+
Ki67:::creRT:::Tom2
(b)
Sox8
Sox10
Zfp62
Notch1
Basp1
Egr2
Taf10
Fos
Foxg1
Thrsp
Ankrd6
Atf1
Hivep3
Hopx
Lmo1
Nfe2l2
Notch3
Nrarp
Hif1a
Nr2e1
Ppargc1a
Fank1
Hes1
Hivep2
Tle3
Zfp423
Nmi
Pbxip1
Pdlim4
Foxo1
Emx2
Gli3
Rorb
Tef
Chd4
Sox9
Trps1
Zic1
Rfx4
Sall2
Olig2
Lhx2
Pias1
Prdm10
Hes5
Up in vNSCs
Up in vTAPs
SC Mid Stage (N)
SC Mid Stage (OL)
Up in dTAPs
SC Early Stage (N)
SC Early Stage (OL)
Up in dNSCs
(c)
79.22
71.29
63.37
55.45
47.53
39.61
31.69
23.76
15.84
7.92
0.00
Bulk NBs
Bulk TAPs
Up in vTAPs
Up in vNSCs
Bulk aNSCs
SC Mid Stage (N)
SC Mid Stage (OL)
Up in dTAPs
Up in dNSCs
SC Early Stage (OL)
SC Early Stage (N)
Bulk qNSCs

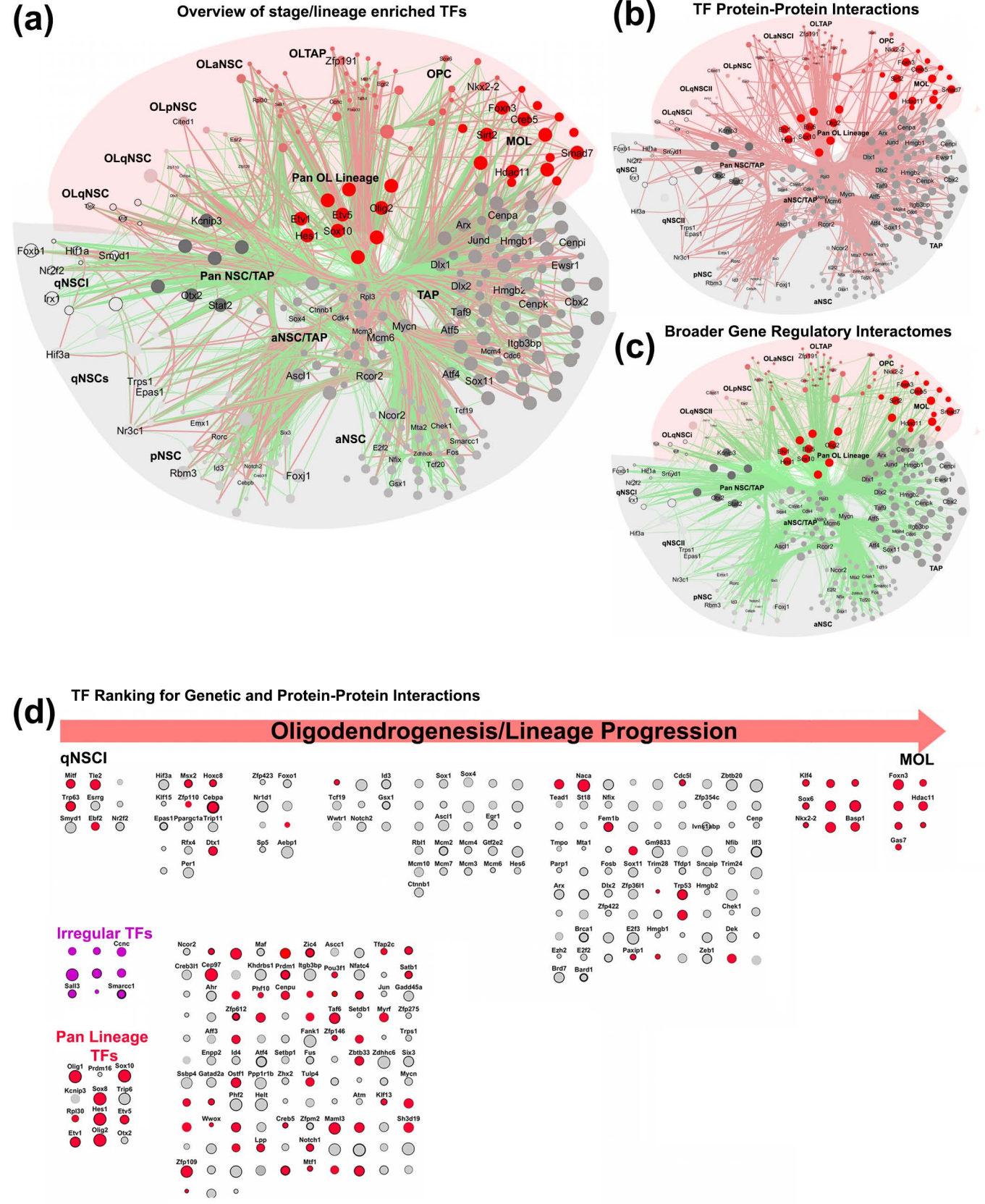

A
Gapdh
-3.32
5.4
OPC
qNSCII
TAPs
aNSC
NBs
pNSC
qNSCI
Microglia/
Endothelia-like
Neurons
MOL
B
Hes5 Pan NSCs
-3.32
4.76
C
Agt qNSCs I
-3.32
4.1
D
Hopx qNSCs II
-3.32
4.11
E
Foxj1 pNSCs
-3.32
3.27
F
Ascl1 aNSCs
-3.32
3.55
G
Nes TAPs
-3.32
2.63
H
Dcx NBs
-3.32
4.02
I
Olig1 OL Lineage
-3.32
4.54
J
Olig2 OL Lineage
-3.32
4.6
K
Pdgfra OPCs
-3.32
5.2
L
Mag mOLs
-3.32
5.14


(a) qNSCI
247
64
115
Protein Class
% Fold Enrichment
0.5-2
2-3
4-6
>10
Sig
<E0.5
<0.02
<0.05
ns
Transmembrane receptor (regulatory/adaptor)
Extracellular matrix glycoprotein
Cation transporter
Glycosidase
Phospholipase
Lipase
Extracellular matrix protein
Transporter
Transferase
Phosphorylase
G-protein coupled receptor
Complement component
Receptor
Membrane traffic protein
Pathways
% Fold Enrichment
0.5-2
2-3
5-10
>10
Sig
<E0.5
<0.001
<0.05
ns
Ionotropic glutamate receptor
Angiotensin
Notch signaling
Histamine H1 receptor mediated signaling
G-protein signaling
Angiogenesis
AD-presenilin
Endothelin signaling
Wnt signaling
Cadherin signaling
Xanthine and guanine salvage
Vasopressin synthesis
Oxytocin receptor mediated signaling
Circadian clock system
Adenine and hypoxanthine salvage
(b) qNSCII
69
125
219
Protein Class
% Fold Enrichment
0
1-2
3-4
6-8
>10
Sig
<E0.4
<0.05
ns
Basic helix-loop-helix TF
Guanyl-nucleotide exchange factor
Intermediate filament binding protein
Intermediate filament
Transmembrane receptor (regulatory/adaptor)
Cation transporter
ECM glycoprotein
Methyltransferase
Acyltransferase
TcoF
Phosphodiesterase
Anion channel
Transporter
Dehydrogenase
TF
Pathways
% Fold Enrichment
0
1-2
6-7
>10
Sig
<E0.5
<0.05
ns
Circadian clock system
Vasopressin synthesis
Ionotropic glutamate receptor
Metabotropic glutamate receptor group III
Gq/Go alpha G-protein signaling
Vitamin D metabolism
Sulfate assimilation
Hedgehog signaling
Glutamine glutamate conversion
Adrenaline and noradrenaline biosynthesis
Acetate utilization
AD-amyloid secretase
(c) pNSC
525
32
123
Protein Class
% Fold Enrichment
0
0.5-1
2-3
7.5-10
>10
Sig
<0.0001
<0.05
ns
Phosphatase modulator
Phosphatase inhibitor
Peroxidase
Intracellular ca2+-sensing protein
Heterotrimeric g-protein
Calmodulin
Calcium-binding protein
Transferase
Transaminase
Surfactant
Storage protein
Non-motor microtubule binding protein
Acetyltransferase
Lyase
Ligase
Pathways
% Fold Enrichment
0
0.5-1
3-4
7.5-10
>10
Sig
<0.001
<0.05
ns
CCKR signaling map
Beta3 adrenergic receptor signaling
5-Hydroxytryptamine degredation
Glycolysis
PI3 kinase pathway
Insulin/IGF-PKB signaling
Vasopressin synthesis
Pyruvate metabolism
Methylmalonyl
Histamine H2 receptor mediated signaling
Fructose galactose metabolism
Alpha adrenergic receptor signaling
Ubiquitin proteasome
PDGF signaling
(d) aNSC
189
160
207
Protein Class
% Fold Enrichment
0
2.5-3
5-6
7.5-10
>10
Sig
<E07
<0.01
<0.05
ns
Ribosomal protein
RNA binding protein
Translation initiation factor
Translation factor
mRNA splicing factor
Translation elongation factor
DNA helicase
DNA/RNA binding
Amino acid kinase
Basic helix-loop-helix TF
DNA-directed RNA polymerase
Nuclease
DNA binding protein
TF
Pathways
% Fold Enrichment
0
2-3
5-6
6-7
>10
Sig
<E0.5
<0.01
ns
FAS signaling
DNA replication
De novo pyrimidine DNA biosynthesis
De novo purine biosynthesis
Notch signaling
p53
Tetrahydrofolate biosynthesis
Salvage pyrimidine RNA
De novo pyrimidine RNA biosynthesis
Arginine biosynthesis
Gonadotropin-releasing hormone receptor
PI3 kinase pathway
AD-presenilin
AD-amyloid secretase
(e) TAP
426
493
200
Protein Class
% Fold Enrichment
0
2.5-3
5-6
7.5-8
>10
Sig
<0.001
<0.01
ns
Chaperonin
mRNA splicing factor
Ribonucleoprotein
mRNA processing factor
Chaperone
Translation factor
RNA binding protein
Translation initiation factor
ATP synthase
Ribosomal protein
DNA-directed RNA polymerase
Nucleic acid binding
Chromatin/chromatin-binding protein
DNA binding protein
Pathways
% Fold Enrichment
0
2.5
4-5
7-8
>10
Sig
<E0.5
<0.005
<0.05
ns
Ubiquitin proteasome
Tetrahydrofolate biosynthesis
DNA replication
De novo pyrimidine DNA biosynthesis
Cell cycle
Formyltetrahydroformate biosynthesis
General transcription by RNA polymerase I
De novo purine biosynthesis
Salvage pyrimidine DNA
Notch signaling
Cytoskeletal regulation by Rho GTPase
General transcription regulation
CCKR signaling map
Angiogenesis
Transcription regulation by bZIP TFs


Cadherin Signaling
Circadian clock system
Angiogenesis
Wnt-signaling
AD-presenilin
Notch -signaling
α adrenergic receptorsignaling
Pyruvate metabolism
Ubiquitin proteasome
General transcription regulation
Transcription regulation by bZIP TFs
General transcriptionby RNA polymerase I
% Enrichment
qNSC
TAP
Significance
ns
<0.005
<E05

Expanded TF Protein Class


(a)
Physical Diffusion
(Protein-Protein Interactions,
i.e. transcription co-factors/adapters)
Genetic Diffusion
(Gene regulatory interactome)
Combined Diffusion
Low
Diffusion Impact
High
0
Mid
High
# of protein-protein interactions
# of linked genetic interactions
# of linked proteins/genetic interactions
Common TFs in qNSCI-pNSC
Common TFs in aNSC-TAP
OLqNSCI-OLpNSC
OLaNSC-OLTAP
OPC/MOL
Pan Lineage
Other
(b)
Common TF-TF Interactions
qNSCI
TAP
TF-TG Regulation:
89.3% Activation
10.7% Inhibition
(c)
Neuronal lineage TF-TF Interactions
qNSCI
NB
TF-TG Regulation:
91.8% Activation
8.2% Inhibition
(d)
OL lineage TF-TF Interactions
OLqNSCI
OPC
Pan OL lineage
TF-TG Regulation:
52.2% Activation
47.8% Inhibition
Connectivity/
TF-TG Interactions
Low
High
TF-TG Activation
TF-TG Inhibition
(e)
All TF-TG Interaction's
Hdac2
Olig2
Sox10
Chd7
Cebpa
Ezh2
Brca1
Mybl2
Uhrf1
Klf9
Tmpo
Brca2
E2f8
Dnmt1
Mxd3
Hmgb2
Tfdb1
Trim28
Ssrp1
Mcm8
Rora
Sfpq
Whsc1
Myt1l
Ttf2
TF-TG Activation
Ezh2
Brca1
Klf9
Tmpo
Dnmt1
Mxd3
Mcm8
Whsc1
Myt1l
Zeb2
Myt1
Sox11
Hopx
Hdac2
Cebpa
Irx5
Smarca5
Pknox2
Nfib
Tef
Etv1
Tgif2
Hlf
Irx3
Csrnp3
TF-TG Inhibition
Rreb1
Cebpa
Per1
Stat5b
Taf13
Prdm16
Tgif2
Hlf
Epas1
Ndn
Nr2e1
Hey1
Hes5
Dlx1
Cdc5l
Ctcf
Foxg1
Klf15
Smarca4
Ascl1
Meis2
Nfix
Ebf2
Patz1
Rbpj
Relative TF-TG Interactions
Low
High
(f)
High Ranking TF's
Medium Rank TF
Low Rank TF


(a)
Overview of stage/lineage enriched TFs
OLaNSC
OLTAP
Zfp191
OPC
Nkx2-2
Foxn3
Creb5
Srf2
MOL
Smad7
Hdac11
OLpNSC
Cited1
OLqNSC
Pan OL Lineage
Olig2
Etv1
Etv5
Hes1
Sox10
Kcnip3
Arx
Cenpa
Jund
Hmgb1
Cenpi
Ewsr1
Dlx1
Dlx2
Hmgb2
Cenpk
Cbx2
Taf9
Atf5
Itgb3bp
Mcm4
Cdc6
Foxb1
Hif1a
Smyd1
Nr2f2
qNSCI
Irx1
Pan NSC/TAP
Otx2
Stat2
TAP
Rpl3
Ctnnb1
Sox4
Cdk4
Mycn
Mcm3
Mcm6
aNSC/TAP
Hif3a
qNSCs
Trps1
Epas1
Ascl1
Rcor2
Atf4
Sox11
Tcf19
Ncor2
Mta2
Chek1
Nr3c1
Emx1
Rorc
Six3
aNSC
E2f2
Smarcc1
pNSC
Nfix
Zdhhc8
Fos
Tcf20
Rbm3
Id3
Notch2
Cebpb
Foxj1
Gsx1
(b)
TF Protein-Protein Interactions
(c)
Broader Gene Regulatory Interactomes
(d)
TF Ranking for Genetic and Protein-Protein Interactions
Oligodendrogenesis/Lineage Progression
qNSCI
MOL
Irregular TFs
Pan Lineage TFs