Received: 25 July 2022 | Revised: 17 October 2022 | Accepted: 19 November 2022

DOI: 10.1002/hipo.23482

RESEARCH ARTICLE

WILEY

Early fate bias in neuroepithelial progenitors of hippocampal neurogenesis

Leonie von Berlin[1] | Jakub Orzechowski Westholm[2] | Michael Ratz[1,3] | Jonas Frisén[1]

[1]Department of Cell and Molecular Biology, Karolinska Institute, Stockholm, Sweden

[2]Department of Biochemistry and Biophysics, National Bioinformatics Infrastructure Sweden, Science for Life Laboratory, Stockholm University, Solna, Sweden

[3]KTH Royal Institute of Technology, Department of Gene Technology, Stockholm, Sweden

Correspondence
Jonas Frisén and Michael Ratz, Department of Cell and Molecular Biology, Karolinska Institute, Stockholm, Sweden.
Email: jonas.frisen@ki.se and michael.ratz@ki.se

Funding information
DFG Research Fellowship, Grant/Award Number: RA 2889/1-1; European Research Council, Grant/Award Number: 695096; Karolinska Institute funding for doctoral education, Grant/Award Number: 2019-01007; Knut och Alice Wallenbergs Stiftelse, Grant/Award Number: 2018.0063; Stiftelsen för Strategisk Forskning, Grant/Award Number: SB16-0014; Swedish Cancer Foundation, Grant/Award Number: 19 0452 Pj 01 H; Swedish Research Council, Grant/Award Number: D0761801

Abstract

Hippocampal adult neural stem cells emerge from progeny of the neuroepithelial lineage during murine brain development. Hippocampus development is increasingly well understood. However, the clonal relationships between early neuroepithelial stem cells and postnatal neurogenic cells remain unclear, especially at the single-cell level. Here we report fate bias and gene expression programs in thousands of clonally related cells in the juvenile hippocampus based on single-cell RNA-seq of barcoded clones. We find evidence for early fate restriction of neuroepithelial stem cells to either neurogenic progenitor cells of the dentate gyrus region or oligodendrogenic, non-neurogenic fate supplying cells for other hippocampal regions including gray matter areas and the Cornu ammonis region 1/3. Our study provides new insights into the phenomenon of early fate restriction guiding the development of postnatal hippocampal neurogenesis.

KEYWORDS
clonal relationships, fate choice, fate specification, neuroepithelial precursors, neuroepithelial stem cells

1 | INTRODUCTION

The dentate gyrus (DG) of the murine hippocampus (HC) is one of two canonical areas maintaining neurogenesis during adulthood in the mouse (Denoth-Lippuner & Jessberger, 2021). The structure arises from progenitors in the dentate neuroepithelium, a small area in the ventricular zone of the medial pallidum, around embryonic day 10.5 (E10.5). These neuroepithelial progenitor cells proliferate and, at around E15.5, begin migrating medially to form the dentate migratory stream and eventually the primitive DG (Berg et al., 2019). Perinatally the progenitor cell clones populate the DG with granule neurons and give rise to the first radial glia-like adult neural stem cells (RGDG) (Bayer & Altman, 1974; Berg et al., 2019; Namba et al., 2005). Their cytogenic activity declines quickly with the entrance into a long-term quiescence state mostly around postnatal (P) day three. However, DG development including the formation of the subgranular zone is only concluded

Michael Ratz and Jonas Frisén are equal last authors.

around 2 weeks after birth. How progenitor cell populations contribute to the different hippocampal areas remains to be determined. However, it has recently been shown that Hopx$^+$ embryonic progenitors of the dentate neuroepithelium produce cells for no other region than the dentate gyrus, while the adjacent ammonic neuroepithelium was found to contribute mainly to Cornu ammonis (CA) neurons (Berg et al., 2019).

Neuroepithelial stem cells and their progeny theoretically have the potential to commit to several fates within the neural lineage. However, it is widely acknowledged that RGDGs in the HC do not produce oligodendrocyte lineage cells as has been shown by several fate mapping studies (Ahn & Joyner, 2005; Bonaguidi et al., 2011; Pilz et al., 2018; Suh et al., 2007). Jessberger et al. could show that overexpression of the transcription factor Ascl1 served as a switch from neuronal fate to the exclusive production of oligodendrocytes (Jessberger et al., 2008). However, how early in development this fate restriction is established remains unclear.

Astrocytes are thought to be produced by adult neurogenic stem cells (Bonaguidi et al., 2011) even though in much lower numbers than neurons. On the other hand, in Hopx$^+$ progenitor cell clones no astrocytes have been found during embryonic development and only sparsely at postnatal days (Berg et al., 2019).

A new model of continuous lineage specification and granule neuron production for embryonic Hopx$^+$ progenitors has been proposed (Berg et al., 2019). This is opposed to the previously suggested "sequential" model where unknown progenitor cells produce cell types of different lineages in a stepwise manner (Kriegstein & Alvarez-Buylla, 2009) and to the "set aside" model where a part of the embryonic progenitor cells is set aside for later adult neurogenesis (Fuentealba et al., 2015; Furutachi et al., 2015). Whether the continuous model is a valid model for the RGDGs and whether different models are true for different progenitor populations has not been validated further yet.

In this study, we take advantage of our previously published single-cell RNA-sequencing (scRNAseq) dataset containing clonal information to understand early fate restriction of DG neurogenesis progenitors (Ratz et al., 2022). We provide statistical evidence, unbiased and in hundreds of clones, that progenitors of postnatal neurogenesis in the hippocampus are strongly biased toward production of neurons for the DG and gray matter astrocytes but not to CA1/3 neurons. The oligodendrocyte lineage in the HC, on the other hand, was found to descend from a different progenitor population as early as E9.5 and did not show restriction to the hippocampal region. White matter astrocytes were more regularly found in these clones, giving the notion of a cell production for white matter areas. In summary, our results point to clearly separated regional and fate specifications of postnatal DG neurogenesis and oligodendrocyte progenitors in early embryonic development.

2 | MATERIAL AND METHODS

2.1 | Mice

Two female and three male CD-1 mice from Charles River Germany between P11-14 (1× P11 female; 1× P11 male; 1× P12 male control without injection; 1× P12 female; 1× P14 male) have been used for the study. Animals were kept in standard housing conditions (ambient temperature of 20°C–22°C and humidity of 40%–60%) with a 12/12 h light/dark cycle and with food and water ad libitum. All experimental procedures on animals were permitted by Norra Djurförsöksetiska Nämnd.

2.2 | Production of lentiviral library

The plasmid and lentiviral libraries have been prepared as described previously (Ratz et al., 2022). In short, lentiviral plasmids containing nuclei-located reporter transgene H2B-EGFP, 30 nucleotide long, random barcode called "cloneID" downstream of human EF1a promoter (LV-EF1a-H2B-EGFP-30N) have been assembled by replacement of PGK1 promoter from LV-GFP50 plasmid and with Gibson assembly method. The plasmids were used to generate lentivirus particles with a titer of >10^9 transducing units/ml by either the VirusTech core facility at Karolinska Institute or by GEG-Tech (Paris, France). The resulting lentiviral library contained around >10^6 uniformly represented and diverse cloneIDs per microliter.

2.3 | Ultrasound-guided in utero microinjection

We aimed to trace progeny of ventricular zone stem cells of the early developing nervous system with a modified version of a previously published procedure (Beronja et al., 2010). To do so, timed pregnancies were set up overnight and plug-positive females identified the next morning considered as embryonic (E) day 0.5. Pregnancy was confirmed with ultrasound at E8.5 of gestation. One day later, pregnant females were anesthetized with isoflurane, uterine horns were exposed and about 4–8 embryos per female injected with 0.6 µl lentivirus (≈0.95 × 10^6 unique cloneIDs) into the forebrain ventricles. To increase survival ratios, each surgery was limited to 30 min.

2.4 | Tissue collection and cell isolation

Mice were terminated with an isoflurane overdose and perfused transcardially with ice-cold artificial cerebrospinal fluid (aCSF: 87 mM NaCl, 2.5 mM KCl, 1.25 mM NaH_2PO_4, 26 mM $NaHCO_3$, 75 mM sucrose, 20 mM glucose, 2 mM $CaCl_2$, 2 mM $MgSO_4$. Equilibrated in 95% O_2/5% CO_2). Brains were collected in ice-cold aCSF and 1 mm coronal sections prepared by inserting the brain in an acrylic brain matrix for mouse (World Precision Instruments). Hippocampus, cortex, and striatum were isolated from each section under a stereomicroscope at constantly cool temperatures. Tissue was cut in pieces and further dissociated by 20–30 min enzymatic digestion with Papain (Worthington Biochemical) at 37°C and subsequent trituration with fire-polished Pasteur pipettes. Filtering of dissociated tissue through sterile 30-µm aCSF-equilibrated Filcon strainer (BD Biosciences) into 15-ml centrifuge tube containing 9 ml of aCSF and 0.5% BSA yielded a

cellular suspension. After mixing, cells were pelleted in a centrifuge at 300*g* and 4°C for 5 min and the supernatant discarded. Next, cells were resuspended in 1 ml of aCSF containing reconstituted ovomucoid protease inhibitor with BSA. To create a discontinuous density gradient, 2 ml of undiluted albumin inhibitor solution were carefully overlaid with 1 ml of cell suspension, and subsequent centrifugation at 100*g* and 4°C for 6 min. The supernatant was carefully discarded, the pellet resuspended in 1 ml of aCSF containing 0.5% BSA and transferred to a round-bottom tube (BD Biosciences) for flow cytometry.

Single $EGFP^+$ cells were collected into one DNA LoBind tubes (Eppendorf) per sample containing aCSF with 0.5% BSA, separated from EGFP- cells by a BD Influx equipped with a 140-µm nozzle and a cooling unit with a sample temperature of 4°C. All samples were pelleted in a centrifuge at 300*g* for 5 min at 4°C and the supernatants were carefully removed. Finally, the cells were resuspended in a minimal volume of aCSF and their concentrations were determined using a Bürker chamber. All steps except for enzymatic digestion were performed on ice or 4°C.

Except for the FACS step, uninjected control brains underwent the same procedure.

2.5 | Single-cell RNA-sequencing

Two versions of 10× Genomics Chromium Single Cell Kit were used: Brain 1–2 were processed with version 2 and brain 3–5 with version 3. $EGFP^+$ cell suspensions were prepared as described above, counted, and resuspended in aCSF before being added to 10× Chromium RT mix. Suspensions from control brains were prepared as described above, diluted in aCSF to concentrations between 800 and 1000 cells/µl and added to 10× Chromium RT mix. Following steps including cDNA synthesis with 12 polymerase chain reaction (PCR) cycles, library preparation, and sequencing were performed according to the manufacturer's instructions.

2.6 | CloneID enrichment from cDNA

As described in Ratz et al., a nested PCR strategy was applied to enrich cloneIDs from full-length cDNA. Each amplicon library was sequenced on a MiSeq or NovaSeq 6000. Cell Ranger version 3.0.1 count was employed for data processing of amplicon libraries and the TREX pipeline to extract cloneIDs (see below).

2.7 | Single-cell RNA-sequencing data processing

The sequencing data from three regions (hc, cx, str) of four barcoded and one control brain were analyzed with Cell Ranger and with the R-package Seurat version 3 (Stuart et al., 2019).

First, we downsampled the number of cells of the control brain to 9000 (cx), 8000 (hc), and 7000 cells (str) to adjust for the lower number of cells in each barcoded brain and region. Running Cell Ranger count yielded gene expression matrices for each region which were merged using merge() in Seurat version 3 (Stuart et al., 2019). Genes that were expressed in less than ≈0.1% of all cells and cells expressing less or more than 500–10,000 genes were removed. The data were log-normalized with a scale factor of 10,000 using the NormalizeData ()function, followed by linear transformation/scaling. Doublets were removed based on expression of mutually exclusive markers for various cell types (Igf2, Pf4, Hexb, Rsph1, Pdgfra, Bmp4, Mog, Clic6, Rgs5, Cldn5, Reln, Igfbpl1, Slc32a1, Slc17a7, and Aldoc). To adjust for cell cycle states, a cell cycle score was calculated for each cell and differences between G2M and S phase scores were regressed out. Highly variable features were selected using FindVariableFeatures(), followed by principal component analysis and the selection of the first 10–30 significant principal components for graph-based clustering (shared nearest neighbor graph calculation and clustering using Louvain). Major cell classes were assigned to each cluster (astroependymal, immune, neurons, oligodendrocytes, and vascular) based on differentially expressed genes and canonical markers. Then each cell class was further subclustered and each cluster extensively annotated based on canonical marker genes from published data and from www.mousebrain.org (Zeisel et al., 2015). Cell type nomenclature was also adopted from (Zeisel et al., 2015). At each step, we removed (1) clusters classified with ambiguous labels and (2) outlier cells on the edges of clusters in uniform manifold approximation and projection (UMAP) space to further eliminate doublets. All cells were merged into a single file together with metadata and annotations.

For most parts of this study, we worked with a reduced dataset based on the merged, final dataset described above. For that, cells found in the hippocampus and hippocampal clones were extracted and cell types not descending from the neural lineage (microglia, perivascular macrophages, vascular leptomeningeal cells, vascular endothelial cells, and vascular smooth muscle cells) were removed. The neuronal cell class was split into two classes: neurons and neuronal progenitors. UMAP representations were re-calculated to include only the extracted cells.

The filtered cellIDs were exported and used as input for cloneID extraction and clone calling.

2.8 | CloneID extraction and clone calling

Raw 10× Genomics Chromium v2 or v3 sequencing data were preprocessed with Cell Ranger version 3.0.1. For read mapping a custom reference was created consisting of the GRCm38 (mm10) genome and an additional sequence representing the H2B-EGFP-N transgene, in which the cloneID region was marked with 30 "N" wildcard characters. The resulting BAM file of aligned sequencing reads was then processed with TREX, our custom Python tool for cloneID extraction and clone calling. TREX was set to only use reads from previously filtered cells (see above) that align to the H2B-EGFP-N transgene. CloneIDs are recovered from those alignments that cover or sufficiently stretch into the masked cloneID region. If soft clipping is encountered at one of the bases adjacent to the region, the alignment is assumed to

continue ungapped into the region. All cloneIDs with identical unique molecular identifiers (UMIs) that come from the same cell (have the same cellID) are collapsed to a consensus sequence. They are assumed to originate from the same mRNA molecule. To error-correct cloneIDs, they were single-linkage clustered using a Hamming distance of at most five as linking criterion. In each cluster, all cloneIDs are replaced with the cloneID occurring most frequently in that cluster. From the resulting final cellID–cloneID combinations, those that are supported by only one UMI and one read are discarded. Also removed are cloneIDs that are supported by only one UMI and have a high frequency in another cell. We assume that those cloneIDs are contaminations. The cleaned data are transformed into a count matrix showing UMI counts for each cloneID in each cell. This matrix was used to sort cells into clones of cells with similar cloneID combinations. To prevent erroneous cloneIDs to contribute to formation of clones, clone calling was done by determining the Jaccard similarity between each pair of cloneID-expressing cells using the R package proxy (Meyer, 2019). The Jaccard score is the number of overlapping barcodes divided by the total number of unique barcodes in a pair of cells. A Jaccard score of 0.7 was determined as a cutoff for related cells (Biddy et al., 2018). A lower score resulted in incorrect assignment of unrelated cells into clones and a higher score did not lead to further splits. Clones were defined as groups of two or more related cells.

2.9 | Statistical analysis

Average clone sizes and cell counts in clones were compared using two-sided Mann–Whitney *U*/Wilcoxon rank sum test for two non-normally distributed, independent groups. Statistical tests were performed using R. Results of statistical tests are always denoted as mean ± *SD* if not specified otherwise. Combined violine and boxplots were created with vioplot R package (Adler & Kelly, 2020).

2.10 | Calculation of clonal coupling score

Clonal coupling scores (Wagner et al., 2018) were calculated considering all clones from all four barcoded brains containing at least three cells per clone. Using the BiRewire R package (Iorio et al., 2016) the clone–cell type associations were shuffled randomly, while preserving the number of cell types related to each clone and the number of clones related to each cell type, to create 1000 randomized datasets. We compared the observed co-appearances of cell types to the randomized datasets to obtain empirical *p*-values and *z*-scores for each pair of cell types indicating how often we expect to see the observed clonal association. The Pearson correlations of *z*-scores between each pair of cell types were calculated and the correlation coefficients were transformed using Fisher *z*-transformation.

The resulting scores were hierarchically clustered based on Euclidian distance with the "complete" clustering method and plotted as heat map using the R package pheatmap.

2.11 | Immunostaining and counting of EGFP+ cells

P11–P14 mice were terminated with an isoflurane overdose and transcardially perfused with cold PBS and fresh 4% formaldehyde (FA). Isolated brains were fixed in 4% FA overnight at 4°C and then moved to 30% sucrose overnight at 4°C. For sectioning, they were embedded in O.C.T. (Sakura) and cut into 50-µm thick sections with VT1000S vibratome (Leica). Sections were incubated with blocking/permeabilization buffer (5% donkey serum and 0.3% Triton X-100 in DPBS) and stained with antibodies against EGFP (chicken, 1:2000, Aves Labs, AB_2307313) and NeuN (rabbit, 1:500, Atlas Antibodies, AB_10602305) at 4°C overnight. Subsequently, DPBS was used to wash the sections which were then incubated with fluorophore-conjugated donkey secondary antibodies (all 1:500, Jackson ImmunoResearch) against the respective species (anti-chicken Alexa Fluor 488, 703-545-155, AB_2340375; anti-rabbit Alexa Fluor 647, 711-605-152, AB_2492288) in blocking buffer at room temperature for 1 h, followed by washing and mounting with ProLong Diamond Antifade Mountant (ThermoFisher, P36961). Confocal images were captured with a laser scanning confocal microscope (LSM700, Carl Zeiss) using a Plan-Apochromat ×10/0.45 or ×20/0.8 objective. Image processing and analysis was performed using Fiji software and Omero.

For EGFP+ counting in the hippocampus, three 50-µm thick sections were analyzed for each of three brains. The area of the hippocampus was determined with Fiji software and the EGFP+ cell count for each hippocampal volume (50 * hippocampal area) determined. The total volume of each hippocampus was determined using consecutive sections and the total EGFP+ cell count for one hippocampus extrapolated from the count of each section. An average number was calculated from the extrapolations of all three sections for each hippocampus.

3 | RESULTS

3.1 | A single-cell and clonal dataset to explore the development of hippocampal postnatal neurogenesis

To study the clonal relationships between neural stem cells of the ventricular zone and the origin of postnatal neurogenesis in the hippocampus during mouse development, we took advantage of a recently published dataset (Ratz et al., 2022). Previously, we developed a method for clonal TRacing and EXpression profiling in the mouse brain using scRNAseq (short: TREX) (Figure 1a). We cloned a library of random, 30 nucleotide-long barcodes (which we refer to as "cloneIDs" from here on) downstream of a reporter gene encoding nuclear localized enhanced green fluorescent protein (EGFP) and the human translation elongation factor 1-alpha (Ef1a) promoter, which is strongly and ubiquitously expressed in the central nervous system. We then packaged this barcode library into a lentiviral vector and performed ultrasound-guided microinjection into the forebrain ventricles of E9.5

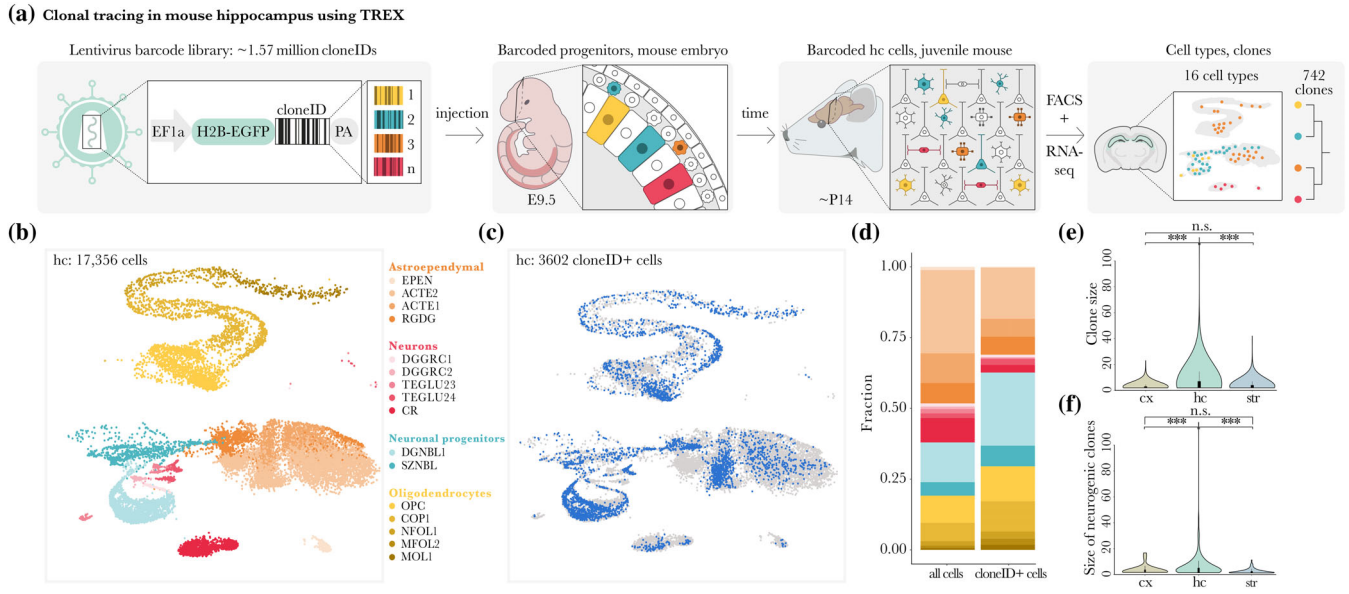

FIGURE 1 A single cell and clonal dataset to explore the development of hippocampal postnatal neurogenesis. (a) Experimental approach for clonal labeling of brain progenitor cells adapted from Ratz et al. (2022). A lentiviral library consisting of an enhanced green fluorescent protein (EGFP) reporter and a diverse barcode (cloneID) library were injected into the forebrain ventricles of mouse embryos at embryonic day 9.5 where they integrated into the genome of the lining progenitor cells. EGFP$^+$ progeny cells isolated at postnatal day 11–14 from hippocampus (as well as striatum and cortex) were analyzed with scRNAseq to reveal both clonal and transcriptional identity. Analysis was focused on hippocampus-derived cells and clones containing at least 10% hippocampal cells. (b) Visualization of hippocampal cells of neural lineage in UMAP space. After removal of non-neural and non-hippocampal cells, 17,356 cells from five brains remained which we grouped into 16 cell types of four classes: Astroependymal (orange), neurons (red), neuronal progenitors (blue) and oligodendrocyte lineage cells (yellow). Cell type nomenclature was taken from www.mousebrain.org (Zeisel et al., 2018). (c) The same UMAP as in (b) highlighting all cloneID$^+$ cells (3602 or 21%) in blue. (d) Stacked bar plot displaying fractions of cells belonging to each cell type among all cells of the dataset (left) and among cloneID$^+$ cells (right). (e,f) Combined violine and box plots showing the distributions of clone sizes of all clones (e) and neurogenic clones (f) in cortex (cx), hippocampus (hc), and striatum (str). Neurogenic clones are defined as clones with at least two cells with neurogenic potential. (e) hc versus cx: $p < 2.2e{-}16$, Wilcoxon rank sum test (WRST), 5.37 ± 3 versus 2.99 ± 1.83, mean ± SD, $n = 559$ and $n = 627$, respectively. hc versus str: $p = 4.28e{-}15$, WRST, 5.37 ± 3 versus 3.3 ± 2.67, mean ± SD, $n = 559$ and $n = 640$, respectively. cx versus str: $p = .09$, WRST, 2.99 ± 1.83 versus 3.3 ± 2.67, mean ± SD, $n = 627$ and $n = 640$, respectively. (f) hc versus cx: $p = 7.36e{-}4$, WRST, 7.71 ± 10.3 versus 3.96 ± 3.77, mean ± SD, $n = 208$ and $n = 23$, respectively. hc versus str: $p = 6.59e{-}15$, WRST, 7.71 ± 10.3 versus 3.46 ± 1.97, mean ± SD, $n = 208$ and $n = 228$, respectively. cx versus str: $p = .54$, WRST, 3.46 ± 1.97 versus 3.96 ± 3.77, mean ± SD, $n = 208$ and $n = 23$, respectively. ns, not significant. ***$p < .001$.

mouse embryos. Since the lentiviral vector integrates the cloneID-sequence into the genome, we achieved stable clonal labeling of neuroepithelial stem cells of the ventricular wall and their progeny. We could confirm the integration of the vector by immunostaining hippocampi of P11 mice for EGFP (Figure S1). By counting EGFP+ cells in several sections, we extrapolated an average total number of 95,451 EGFP+ cells per hippocampal volume (Table S1). Next, from P11-14 mice we isolated EGFP$^+$ cells from HC, cortex (CX), and striatum (STR, including the subventricular zone) as well as EGFP$^-$ cells from non-injected brains by fluorescent activated cell sorting. We performed scRNAseq for cellular profiling and reconstruction of clones based on cloneID-mRNA sequence (Figure 1a). The obtained dataset contains 62,388 cells from three regions and of 40 cell types.

In this study, we focused on cells isolated from the HC. Since we aimed to decipher the origins of progeny of neural stem cells, we removed all cell types not descending from the neural lineage such as microglia and vascular cells. Our final dataset comprises 17,356 hippocampal cells from five brains (three male, two female) grouped into 16 relevant cell types of four classes (Figure 1b). We annotated the cell types using the nomenclature and marker profiles from a published cell atlas of the mouse brain (Zeisel et al., 2018). In class one we grouped all astroependymal cells, such as ependymal cells (EPEN), white matter astrocytes (ACTE1), gray matter astrocytes (ACTE2), and dentate gyrus radial glia-like cells (RGDG). Immature and mature granule neurons (DGGRC1 and DGGRC2, respectively) as well as excitatory neurons of the hippocampal CA1/3 regions (TEGLU24 and TEGLU23, respectively) and Cajal–Retzius cells (CR) were classified as neurons. In contrast, we considered neuronal intermediate progenitor cells (SZNBL) of the DG and dentate gyrus neuroblasts (DGNBL1) as neural progenitors. Finally, we joined all stages of the oligodendrocyte lineage to the class "oligodendrocytes" (including OPC, COP1, NFOL1, MFOL2, and MOL1). Therefore, our dataset contains all main neural cell types of the HC.

We defined a clone as a group of at least two cells that share the same cloneID. Guided by this definition we used a customized algorithm to extract cloneIDs and to group hippocampal cloneID$^+$ cells (3602 or 21%) into 742 distinct clones (Figure 1c). Cells of such clones were usually found in the HC only (559 or 75% of clones) but one quarter spread across more than this region. We have not included these 367 cells from STR or CX in the UMAP in Figure 1b.

We asked whether all main neural cell types of the HC are sufficiently represented among the cloneID$^+$ cells. We found that the distribution of cell types among cloneID$^+$ cells differs from the overall cell type distribution (Figure 1d). The ratio of astroependymal cells shrinks from 41% to 24% mostly due to reduction of white matter astrocyte counts among cloneID$^+$ cells. Similarly, neurons are underrepresented and their proportion is reduced from 14% to 6%, which is mainly caused by a loss of CR cells. On the other hand, ratios of neuronal progenitors increase due to a high number of cloneID$^+$ cells among DG neuroblasts (from 19% to 33%). Oligodendrocyte ratios increase in all five cell differentiation stages (from 19% to 30% in total). However, all cell types are represented among cloneID$^+$ cells (Figure 1d). The dataset is therefore suitable for analysis of clones in the HC.

To determine whether we captured postnatal neurogenesis sufficiently to draw meaningful conclusions from clonal compositions, we compared clone sizes and counts across regions. We focused on clones consisting of cells from HC only, CX only or STR only, respectively (Figure 1e), excluding clones that are covering two or more brain regions. We found that the average size of hippocampal clones ($n = 559$) was significantly larger than the average size of cortical clones ($n = 627$; $p < 2.2e{-}16$, Wilcoxon rank sum test [WRST], 5.37 ± 3 vs. 2.99 ± 1.83, henceforth mean ± SD if not denoted otherwise) and the average size of striatal clones ($n = 640$; $p = 4.28e{-}15$, WRST, 5.37 ± 3 vs. 3.3 ± 2.67). This difference is mostly due to the size of neurogenic clones (Figure 1f). We defined neurogenic clones as those clones with at least two cells with neurogenic potential, that is, RGDGs, neuronal intermediate progenitors, and DG neuroblasts. We defined this threshold to reduce the likelihood of extracting neurogenic clones based on wrongly assigned cells (0.2% false-discovery rate; Ratz et al., 2022). Neurogenic clones in the HC ($n = 208$) are on average significantly larger than neurogenic clones in STR ($n = 228$; $p = 6.59e{-}15$, WRST, 7.71 ± 10.3 vs. 3.46 ± 1.97) and neurogenic clones in CX ($n = 23$; $p = 7.36e{-}4$, WRST, 7.71 ± 10.3 vs. 3.96 ± 3.77). Hippocampal neurogenic clones can become as big as 118 cells while striatal or cortical clones reach a maximum of 12 or 17 cells, respectively. At the same time, the total number of neurogenic clones is lower in HC than in STR (208 vs. 228). When plotting the average clone composition across different clone sizes it became apparent that large clones (21 cells and more) are dominated by neurogenic cells and some astrocytes while clones under 20 cells show a more even distribution including more oligodendrocytes (Figure S2a). We also compared clone sizes between replicates to ensure the dominance of neurogenic cells in large clones is not a singular occurrence. We found similar relationships between the size of neurogenic clones and other types of clones in all four replicates (Figure S2b–f and Table S2).

In conclusion, our dataset captures the neurogenic potential of the hippocampus and allows to trace back this potential to neuroepithelial stem cells of early embryonic development.

3.2 | Early fate bias of postnatal neurogenesis progenitors

Our two time points of choice, early embryonic development for injection and 2 weeks after birth (P11–P14) for collection of cloneIDs, enabled us to connect neurogenic cells with their neuroepithelial progenitors. We sought to understand the potential of these progenitors at the single-cell level regarding the production of neurogenic and non-neurogenic cell types and how these two groups are linked.

First, we performed clonal coupling analysis on all cell types. For that, we assessed the likelihood for each pair of cell types to share a cloneID and compared this clonal coupling score with randomized data. We identified cell type combinations likely sharing a common ancestor cell at the injection time point and cell types that are likely not descending from the same lineage. We clustered the pairwise coupling scores hierarchically and plotted the results in a heat map (Figure 2a). We found two blocks of highly related cell types that form two separate clusters. On the one hand, oligodendrocyte lineage cells form one distinct block of related cells, on the other hand, cell types of the neurogenic lineage (RGDG, SZNBL, DGNBL1 in Figure 2a) and mature granule neurons (DGGRC1/2 in Figure 2a) clustered together. These two blocks of cell types show negative coupling scores between each other, indicating that neuroepithelial stem cells at E9.5 tend to produce either neurogenic cells and DG neurons or oligodendrocyte lineage cells but not both.

For some cell types an unexpectedly low coupling score was calculated. For example, OPCs and MOL1 cells do not occur in the same clones often. Most likely, OPC-rich clones did not have time to mature to MOL1s at the time point of collection. Generally, mature oligodendrocytes are rare in the dataset due to difficulties to isolate them in a viable state. Finally, biological explanations cannot be excluded either. OPCs are produced in three waves and the OPCs present in this dataset might stem from another wave than the MOL1s (Kessaris et al., 2006).

To understand the fate bias of neuroepithelial stem cells in more detail, we thoroughly investigated the clonal structures. Similar to above defined neurogenic clones, we denoted clones with at least two cells of the oligodendrocyte lineage as "oligodendrogenic." We extracted all oligodendrogenic clones among hippocampal clones (187 out of 559 clones, 1107 cells), plotted them in UMAP space, and determined their cell type ratios (Figure 2b). Interestingly, extracting oligodendrogenic clones reduced the ratio of neuronal progenitors from 33% to 5%, of mature neurons from 6% to 3%, and RGDGs from 6% to 2% while astrocyte ratios remained almost unchanged (24%–25%). Oligodendrocyte ratios increased from 30% to 66%. In contrast, selecting for clones that contain at least two neurogenic cells (208 clones, 1603 cells) reduced the oligodendrocyte lineage ratio from 30% to 6% (Figure 2c). The fraction of RGDGs and neuronal progenitors almost doubled (6%–11% and 33%–64%, respectively). Mature neurons decreased from 6% to 3% due to a loss of CR cells which seem not to be related to neurogenic or oligodendrogenic clones. Astrocyte numbers decreased from 24% to 16%.

In summary, by selecting for oligodendrogenic clones we reduced the ratio of neuronal cells drastically. Vice versa, extracting neurogenic clones resulted in a clear shrinkage of oligodendrocyte lineage cell counts. These results support the observation made with lineage coupling analysis (Figure 2a) that the neurogenic and oligodendrogenic lineages are largely separated.

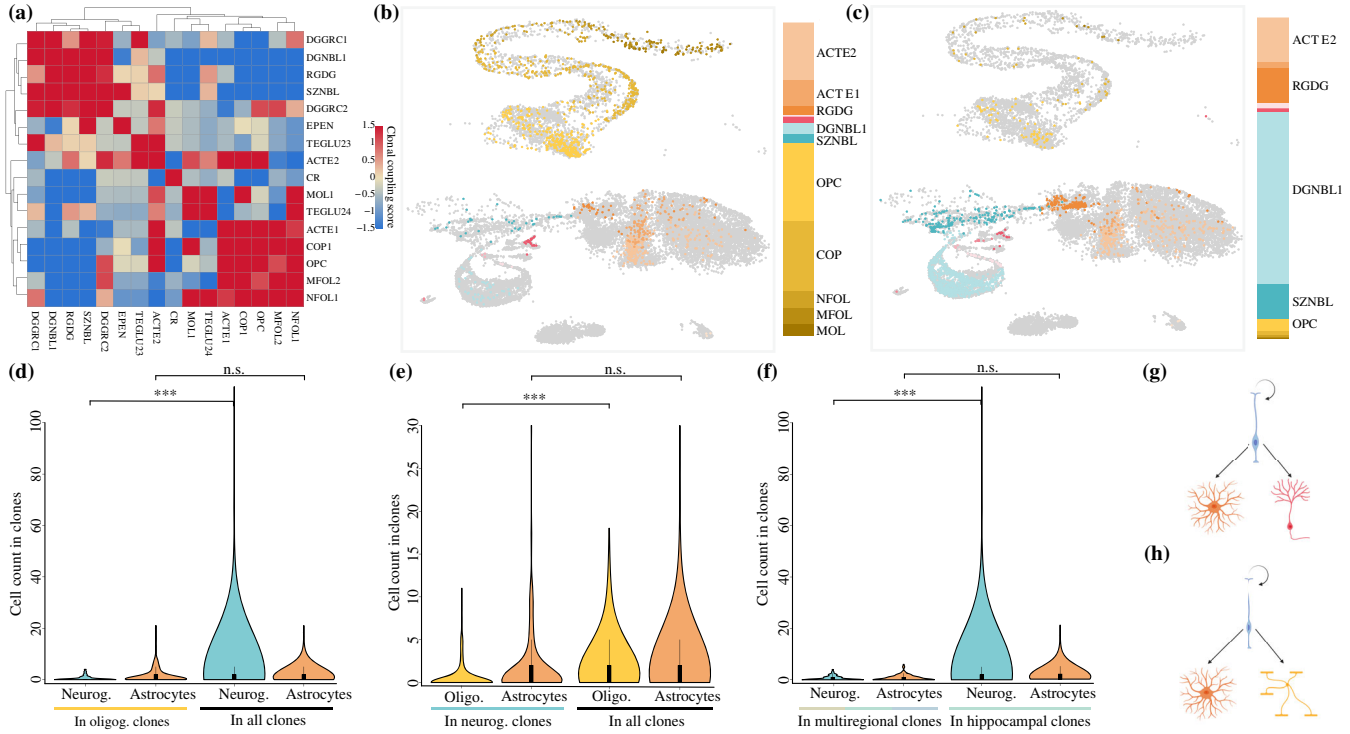

FIGURE 2 Early fate bias of postnatal neurogenesis progenitors. (a) Clustered heat map showing correlation between clonal coupling scores for each pair of cell types. Scores were calculated as the total number of shared cloneIDs between two cell types in comparison to randomized data. High values indicate a likely linkage of the two cell types by a common ancestor while low values indicate that the two cell types are likely not related to each other. (b,c) UMAP plots highlighting all cells in oligodendrogenic (b) and neurogenic (c) clones (left-hand panel) and the respective cell type distributions in such clones (right-hand panel). Oligodendrogenic/neurogenic clones are clones with at least two cells from the oligodendrocyte/neurogenic lineage. Colors are cell-type specific and cell type abbreviations are taken from www.mousebrain.org. (d–f) Combined violine and box plots showing the distributions of cell counts in clones. ns, not significant. ***$p < .001$. Neurog., neurogenic; Oligog., oligodendrogenic. (d) Comparison of cell count distributions of neurogenic cells (blue) and astrocytes (orange) in oligodendrogenic clones (yellow bar) and all hippocampal clones (black bar). Neurogenic in oligodendrogenic versus all: $p = 3.535e{-}14$, Wilcoxon rank sum test (WRST), 0.41 ± 0.89 versus 2.31 ± 6.39, mean ± SD, $n = 187$ and $n = 559$, respectively. Astrocyte in oligodendrogenic versus all: $p = .9964$, WRST, 1.46 ± 2.64 versus 1.25 ± 2.18, mean ± SD, $n = 187$ and $n = 559$, respectively. (e) Comparison of cell count distributions of oligo. (=oligodendrocyte lineage cells, yellow) and astrocytes (orange) in neurogenic clones (blue bar) and all hippocampal clones (black bar). Oligo. in neurogenic versus all: $p = 1.953e{-}13$, WRST, 0.46 ± 1.32 versus 1.44 ± 2.31, mean ± SD, $n = 208$ and $n = 559$, respectively. Astrocytes in neurogenic versus all: $p = .06128$, WRST, 1.21 ± 2.44 versus 1.25 ± 2.18, mean ± SD, $n = 208$ and $n = 559$, respectively. (d) Comparison of cell count distributions of neurogenic cells (blue) and astrocytes (orange) in multiregional clones (tricolored bar, 10%–80% hippocampal cells) and in hippocampal clones (green bar, 100% hippocampal cells). Neurogenic multi-regional versus hippocampal: $p = 1.65e{-}08$, WRST, 0.51 ± 0.96 versus 2.31 ± 6.39, mean ± SD, $n = 159$ and $n = 559$, respectively. Astrocytes multi-regional versus hippocampal: $p = .08$, WRST, 0.74 ± 1.1 versus 1.25 ± 2.18, mean ± SD, $n = 159$ and $n = 559$, respectively. (g,h) Summarizing scheme of the results. After rounds of self-renewing divisions, neuroepithelial stem cells are fate-biased to either produce astrocytes or adult neurogenic cells differentiating to dentate gyrus granule neurons for the hippocampus (g) or oligodendrocytes and white matter astrocytes for multiple regions (h). Created with BioRender.com.

Next, we aimed to test the statistical relevance of our findings. We first focused on oligodendrogenic clones and examined the number of neurogenic cells in these clones (Figure 2d). The average number of neurogenic cells in oligodendrogenic clones is significantly lower than compared with all hippocampal clones ($p = 3.535e{-}14$, WRST, 0.41 ± 0.89 vs. 2.31 ± 6.39). This is not true for astrocytes in oligodendrogenic clones versus all clones ($p = .9964$, WRST, 1.46 ± 2.64 vs. 1.25 ± 2.18). Similarly, when comparing the distribution of oligodendrocyte lineage cells in neurogenic clones with the distribution in all hippocampal clones we found a significant smaller average ($p = 1.953e{-}13$, WRST, 0.46 ± 1.32 vs. 1.44 ± 2.31; Figure 2e). Again, astrocytes do not show this difference ($p = .06128$, WRST, 1.21 ± 2.44 vs. 1.25 ± 2.18). These statistical results support our notion that neuroepithelial stem cells display early fate bias to either produce progenitors of postnatal neurogenesis (Figure 2g) or oligodendrocyte lineage cells (Figure 2h) while astrocytes are produced in both cases. It is possible that we observed these results due to small oligodendrogenic clone size, making it less likely for oligodendrocytes to occur together with neurogenic cells in the same clone. To exclude this explanation for the fate split, we extracted clones with a minimum clone size of 5 and found statistical support for the same fate split between neurogenic and oligodendrogenic clones as presented in Figure 2d,e for all clones (Figure S3a,b). Vice versa, excluding clones with more than 5 cells did not lead to a less significant separation of

neurogenic and oligodendrogenic fate (Figure S3c,d). Furthermore, we could show that clones of 16–20 cells consist of more than 30% oligodendrocytes on average and only in clones larger than 20 cells the ratio of oligodendrocytes decreases significantly (Figure S2a).

Despite the strong and significant differentiation bias between neurogenic and oligodendrogenic fate, we observed neurogenic clones with oligodendrocyte lineage cells and vice versa: 9% of neurogenic clones contain at least two oligodendrocyte lineage cell and 10% of oligodendrogenic clones contain at least two neurogenic cells.

While we detected hints of a fate split in striatum and cortex, results are harder to interpret in these two regions. Extraction of neurogenic or oligodendrogenic clones also led to significant shrinkage of the oligodendrocyte/neurogenic population in striatum and cortex (Figure S4). However, astrocyte numbers underwent a similar size reduction after selection which indicates that diversity of cell types in general was reduced and not neurogenic cells/oligodendrocytes specifically. The appearance of new-born neuronal cells in striatum and cortex was unexpected. Further investigation demonstrated that neurogenic cells in the striatum and cortex are to a great part either olfactory bulb neuroblasts (41% and 60%) or subgranular zone radial-glia like cells (16% and 0%). Thus, they either were directly collected from the subventricular zone or ended up being collected in striatum/cortex while migrating along the rostral migratory stream due to imperfect dissection (Obernier & Alvarez-Buylla, 2019). Since migration of neurogenic cells causes a spatially biased separation of clones and neurogenic clones are rare, the results of Figure S4 should be interpreted with caution. More investigation is needed to confirm a fate split in these two brain regions.

Finally, we asked whether neuroepithelial progenitors destined to produce progeny for postnatal neurogenesis are further restricted in their potential. Previously, we excluded all multiregional clones and focused on hippocampal cells only. We picked clones (n = 183 or 33%) with only 10%–80% hippocampal cells while the rest stem from STR or CX and compared their neurogenic cell counts to pure hippocampal clones (Figure 2f). Interestingly, we observed a significantly lower average count of neurogenic cells in multi-regional than in hippocampal clones (p = 1.65e−08, WRST, 0.51 ± 0.96 vs. 2.31 ± 6.39). The same is true when selecting for neurogenic clones only (p = 7.09e−4, WRST, 2.38 ± 0.67 vs. 5.8 ± 9.5). In contrast, when comparing these counts for astrocytes in multiregional and hippocampal clones we could not find significant differences (p = .08, WRST, 0.74 ± 1.1 vs. 1.25 ± 2.18). Oligodendrocytes, on the other hand, dominate multiregional clones across all clone sizes up to 20 cells (Figure S5a) and generally appear in significantly higher numbers in multiregional clones than in hippocampal ones (Figure S5b, p = 1.8e−08, WRST, 2.03 ± 2.21 vs. 1.44 ± 2.31).

These results suggest that progenitors of postnatal neurogenesis are not only fate-biased as early as day 9.5 of embryonic development but are also strongly limited to produce cells for the hippocampus. The same is not true for neuroepithelial stem cells biased to an oligodendrogenic fate as they supply cells to multiple regions.

3.3 | Early regional bias of postnatal neurogenesis progenitors

Our findings led us to the question whether fate restriction of neuroepithelial stem cells expands beyond oligodendrocyte lineage and neurogenic cells. We observed the occurrence of other cell types available in the dataset in neurogenic and non-neurogenic clones of the hippocampus. Isolated ependymal cells lining the ventricle walls near the HC appear in neurogenic clones in 38% of cases (3 out of 8; Figure 3). Thus, the ependymal and neurogenic lineage does not seem to be strictly separated. Ependymal cells also take part in oligodendrogenic clones (38%) suggesting that they emerge before lineage bifurcation. In contrast, no CR cell was found in a neurogenic or oligodendrogenic clone clearly suggesting a divergence of the CR and neurogenic lineages before E9.5 (0 out of 70; Figure 3). These results coincide well with the fact that ependymal cells have a high clonal coupling score with neuronal intermediate progenitor cells while Cajal–Retzius cells display low clonal coupling with any of the present cell types (Figure 2a).

So far, we have considered astrocytes as a single group. We have seen that they can be produced both by progenitors biased toward neurogenesis or oligodendrogenesis. However, when inspecting gray and white matter astrocytes separately, interesting differences emerged. Almost half or 44% (223 out of 509) of gray matter astrocytes were found in neurogenic clones and 36% in oligodendrogenic ones (Figure 3). This matches our previous finding that astrocytes can be contributed by both neurogenic and oligodendrogenic progenitors. In contrast, with 75% (163 out of 192) the greatest part of white matter astrocytes of the dataset was produced by non-neurogenic progenitors (Figure 3), 48% of cells by oligodendrogenic ones. This

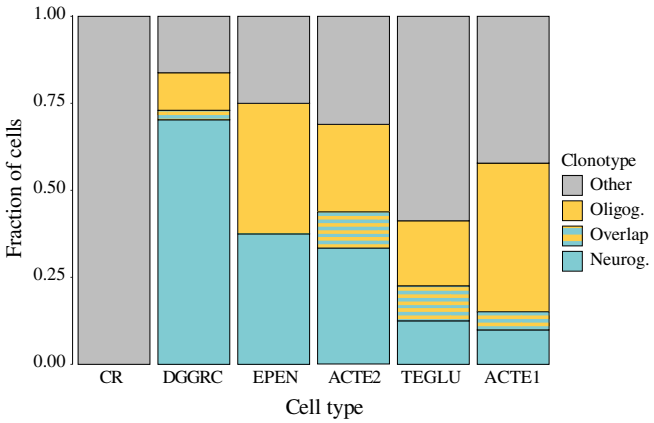

FIGURE 3 Early regional bias of postnatal neurogenesis progenitors. Stacked bar plot showing the fraction of cells found in neurogenic (blue), oligodendrogenic (yellow), both neurogenic and oligodendrogenic (blue and yellow stripes), and other (gray) clones. Each single bar shows a different cell type. ACTE1/2, white/gray matter astrocytes; CR, Cajal–Retzius cells; DGGRC, mature and immature dentate gyrus granule neurons; EPEN, ependymal cells; Neurog., neurogenic; Oligog., oligodendrogenic; TEGLU, excitatory neurons of the hippocampal CA1/3 regions.

finding indicates a bias of certain non-neurogenic progenitors to produce cells for the white matter (oligodendrocytes and white matter astrocytes) while neurogenic clones seem to be more specialized on the formation of gray matter. Hierarchical clustering of clonal coupling scores also placed white matter astrocytes close to the block of oligodendrocyte cells while gray matter astrocytes belong to neither of the blocks (Figure 2a).

Not surprisingly, neurogenic progenitors and their progeny are responsible for the production of at least 73% of granule neurons (27 out of 37; Figure 2e) and with that the establishment of the DG region. Our data do not show whether these neurons have been produced during developmental waves or are a result of more recent neurogenesis.

In contrast, neurons of the CA1 and CA3 regions are mainly not part of neurogenic clones (23% or 18 out of 80 clones; Figure 3) or oligodendrogenic ones (29%). Additionally, these CA1/3 neurons are not clonally coupled to any cell type in the neurogenic block in Figure 2a. While fate restriction is not absolute, it becomes clear that neurogenic progenitors do not necessarily play a role in the construction of the CA1 and CA3 regions (Iyer & Tole, 2020). These results further suggest a bias of neurogenic progenitors to not only the hippocampus but also to the dentate gyrus region within the hippocampus, while non-neurogenic clones tend to provide cells for other regions, including white matter areas.

4 | DISCUSSION

Here we took advantage of a previously published scRNAseq dataset that utilized genetic barcoding of precursor cells to study the developmental relationship between neuroepithelial progenitors and postnatal neurogenic clones in the hippocampus. We showed that a clear and significant bias of progenitors toward a neurogenic or oligodendrogenic fate exists as early as E9.5. Progenitors of neurogenesis in the HC tend to produce cells for the HC such as granule neurons for the DG but not cells for STR, CX or the hippocampal CA1/3 regions. This is opposed to neuroepithelial progenitors biased to an oligodendrogenic fate which generate cells of the oligodendrocyte lineage that spread across several regions and include white matter astrocytes too. Only a small part of ependymal cells was found in neurogenic clones and no Cajal–Retzius cells at all.

It is widely acknowledged that RGDGs in the DG produce granule neurons and astrocytes but not oligodendrocytes (Ahn & Joyner, 2005; Bonaguidi et al., 2011; Pilz et al., 2018; Suh et al., 2007). Oligodendrocytes of the region are instead generated by oligodendrocyte progenitor cells (OPCs). The question whether these OPCs share a lineage with RGDGs is not fully solved yet. When do the neurogenic and oligodendrogenic lineages bifurcate? Oligodendrocytes have not been shown to part of $Hopx^{+}$ progenitor clones in the hippocampus upon Cre-inducible labeling at E10.5 and inspection at P8 and P30 (Berg et al., 2019).

Our data stem from the juvenile mouse brain at around P11–P14 which coincides roughly with the completion of hippocampal structures but does not include adult processes. However, already 2 weeks after birth we captured sufficient neurogenic activity to draw statistical conclusions (Figure 1e,f, clones = 208 clones, average clone size = 7.7 cells/clone). Since we injected cloneIDs at E9.5, our data cover most of embryonic and postnatal development. We could find clear indications of separated lineages at this early time point, providing evidence for largely fate-specified adult neurogenesis progenitors in early embryonic development.

We further provided supporting evidence for previous observations (Berg et al., 2019) with a large number of clones and an unbiased labeling approach, that is, not relying on labeling of a specific progenitor population. We found the same regional restriction of progenitors of neurogenesis which supply cells mainly for the hippocampus and not for other regions. Furthermore, we saw that most CA1/3 neurons were produced in clones that are non-neurogenic postnatally. This is consistent with a previous study which found that the dentate neuroepithelium produces cells for the DG and the CA neuroepithelium for the CA1/3 regions of the hippocampus (Berg et al., 2019). Moreover, oligodendrogenic clones were shown to focus on the production of white and not gray matter astrocytes (Berg et al., 2019). These results give the impression of not only a fate restriction toward a certain lineage (e.g., neurogenic vs. oligodendrogenic) but also toward the requirements of a certain region. It is plausible that oligodendrogenic clones not only provide oligodendrocytes but also white matter astrocytes to build white matter regions such as the surrounding of the HC.

Very low numbers of astrocytes have been reported to be generated in postnatal adult neurogenic clones and none in embryogenesis (Berg et al., 2019). We also found many neurogenic clones not containing any astrocytes (62%) but on average 12% of cells in neurogenic clones were astrocytes. From our data it cannot be determined whether these astrocytes stem from the embryonic or postnatal phase. However, the numbers are probably higher than found by Berg et al. (and probably lower than real numbers in the tissue due to barcode dropout and undersampling of astrocytes, Figure 1d) and this might be due to the inclusion of $Hopx^{-}$ progenitors in the dataset. Our results also favor a continuous model of development for DG neurogenesis progenitors with a continuous specification to produce dentate granule neurons. However, since we regularly found astrocytes together with neurogenic cells (including RGDGs) in clones, we propose an extension of this model to include astrocytes. Whether those represent progeny of a "set aside" population or are produced in parallel with granule neurons remains to be studied.

Ependymal cells are thought to be born between E13.5 and E15.5 from radial glia cells. They have been found to be directly related to stem cells of the ventricular zone (Ortiz-Álvarez et al., 2019). We only reported low numbers of ependymal cells in our dataset, but the low linkage of these cells to neurogenic clones (38% of cells) of the HC might be explained by the fact that they mainly stem from another population of progenitor cells not contributing to the DG.

CRs cells, on the other hand, have not been found in neurogenic clones at all. On the one hand, the reason for this decoupled status can be the early birth of these cells. CR cells are the first neurons

populating the neocortex (Causeret et al., 2021). On the other hand, the peak of CR neurogenesis was determined to be at E10.5–E12.5, after injection. Thus, while CRs cells are thought to be of great importance for the HC and DG, they must originate from different progenitor cells than the postnatally neurogenic clones (Causeret et al., 2021).

Our approach is different from classical lineage tracing approaches in the sense that it does not require a priori knowledge of marker genes and allows unbiased labeling of progenitor cells. We collected 18,570 clonally labeled cells across three different anatomical regions which required the use of only four injected mice. Collecting cloneIDs together with information of the transcriptome of each cell also makes identification of progeny more reliable and detailed. For example, we could distinguish between white and gray matter astrocytes and assess different origins of these two subtypes.

The main challenge of our approach is cell loss and/or barcode dropout. Average clone sizes with our method are much lower (4 ± 0.1 cells per clone (mean ± *SEM*, $n = 2276$ clones) in comparison to traditional fate mapping that found around 200 cells per clone (Llorca et al., 2019). Therefore, the isolated clones in our dataset are most likely incomplete. However, a high variability of clone sizes has been found in both our study and other fate mapping paper (Llorca et al., 2019) hinting at an undetermined nature of clone expansion rates. To overcome our limitations, we sampled high numbers of clones, applied statistical approaches for common observations and expressed our results in fractions of total found cells. For this we assume that, within a cell type, cells end up in our dataset at random. Rare populations and clonal linkages are harder to detect and support with statistical testing. For our analyses, we also must consider that cell type ratios are slightly skewed due to the higher capture of $EGFP^+$ cells and cloneIDs in certain groups. For example, neuroblasts in the DG are overrepresented in $cloneID^+$ data because they divide more than other cell types. We have performed statistical tests using normalized cell ratios for the oligodendrocyte, astrocyte, and neuron class and found negligible differences to the results presented in this study.

The question of why fate biases are not absolute can be discussed in this light. The fact that most but not all dentate granule neurons are found in neurogenic clones can be explained by the dropout of neurogenic cells such that clones that are neurogenic in the tissue do not appear so in the dataset. The existence of oligodendrocytes in neurogenic clones (9% of clones), however, is not explainable with dropout. In our previous publication, we have determined a low error rate of 0.2% for falsely associating a specific cloneID to a single cell (Ratz et al., 2022). This error rate is not sufficient to explain the 2.6% of oligodendrocyte lineage cells that were found in neurogenic clones. Another explanation could be that when separating CX, STR, and HC for cell isolation, cells from nearby regions contaminated the HC sample. However, this argument is weakened when considering that most neurogenic/oligodendrogenic mixed clones contain neurogenic cells specific for the DG (such as RGDGs or DG neuroblasts) or even mature granule neurons. Therefore, we consider the interesting possibility that a small, unidentified subpopulation of neuroepithelial progenitors produces both oligodendrocytes and neurons for the HC, most likely diverging at early time points.

Our work contributes to the understanding of regional and fate specification of adult neurogenesis progenitors. In the future, we hope that the improvement and utilization of spatial transcriptomic approaches will allow a more detailed linkage of gene expression with fate decisions and precise locations of neuroepithelial progenitor cells and their progeny. Especially, it will be exciting to explore how the location, environment and intracellular constitution contribute to fate specification and neurogenesis throughout adulthood. Knowledge about these developmental processes will have important implications in the understanding of neurodevelopmental disorders. Future work that examines the underlying mechanisms of fate decisions toward neurogenesis will be eagerly awaited.

AUTHOR CONTRIBUTIONS

Leonie von Berlin performed all data analysis and created the figures. Jakub Orzechowski Westholm provided the pipeline for calculating clonal coupling scores. Michael Ratz and Jonas Frisén supervised the study. Leonie von Berlin, Michael Ratz, and Jonas Frisén wrote the manuscript, with input from all authors.

ACKNOWLEDGMENTS

We would like to thank Martyna Lukoseviciute for feedback and support on earlier versions of this manuscript.

FUNDING INFORMATION

Jakub Orzechowski Westholm is financially supported by the Knut and Alice Wallenberg Foundation as part of the National Bioinformatics Infrastructure Sweden at SciLifeLab. This study was supported by grants from the Swedish Research Council (D0761801), the Swedish Cancer Society (190452 Pj 01 H), the Swedish Foundation for Strategic Research (SB16-0014), Knut och Alice Wallenbergs Stiftelse (2018.0063) and the European Research Council (695096) to Jonas Frisén, a DFG Research Fellowship (RA 2889/1-1) to Michael Ratz and Karolinska Institute funding for doctoral education (KID Grant: 2019-01007) to Leonie von Berlin.

CONFLICT OF INTEREST

Jonas Frisén is consultant to 10× Genomics. The remaining authors declare no competing interests.

DATA AVAILABILITY STATEMENT

The underlying scRNAseq raw data are stored in the Gene Expression Omnibus, accessible with code GSE153424. Processed single cell data can be downloaded as RDS file with the following link: https://kise-my.sharepoint.com/:f:/g/personal/michael_ratz_ki_se/EndBZ9VI_rRHmHzZxrAwSZQBeE9e4RNmktbuCcHir1a5qQ?e=Ge2Fqm, and password 8RMG.xbzH?3v9Ef4. TREX source code is available under the MIT license from https://github.com/frisen-lab/TREX

ORCID

Leonie von Berlin https://orcid.org/0000-0002-7790-0395
Jakub Orzechowski Westholm https://orcid.org/0000-0002-6849-6220

Michael Ratz https://orcid.org/0000-0002-9795-8033
Jonas Frisén https://orcid.org/0000-0001-5819-458X

REFERENCES

Adler, D., & Kelly, S. T. (2020). vioplot: violin plot. https://github.com/TomKellyGenetics/vioplot

Ahn, S., & Joyner, A. L. (2005). In vivo analysis of quiescent adult neural stem cells responding to sonic hedgehog. *Nature*, *437*, 894–897.

Bayer, S. A., & Altman, J. (1974). Hippocampal development in the rat: Cytogenesis and morphogenesis examined with autoradiography and low-level X-irradiation. *Journal of Comparative Neurology*, *158*, 55–79.

Berg, D. A., Su, Y., Jimenez-Cyrus, D., Patel, A., Huang, N., Morizet, D., Lee, S., Shah, R., Ringeling, F. R., Jain, R., Epstein, J. A., Wu, Q.-F., Canzar, S., Ming, G.-L., Song, H., & Bond, A. M. (2019). A common embryonic origin of stem cells drives developmental and adult neurogenesis. *Cell*, *177*, 654–668.

Beronja, S., Livshits, G., Williams, S., & Fuchs, E. (2010). Rapid functional dissection of genetic networks via tissue-specific transduction and RNAi in mouse embryos. *Nature Medicine*, *16*, 821–827.

Biddy, B. A., Kong, W., Kamimoto, K., Guo, C., Waye, S. E., Sun, T., & Morris, S. A. (2018). Single-cell mapping of lineage and identity in direct reprogramming. *Nature*, *564*, 219–224.

Bonaguidi, M. A., Wheeler, M. A., Shapiro, J. S., Stadel, R. P., Sun, G. J., Ming, G., & Song, H. (2011). In vivo clonal analysis reveals self-renewing and multipotent adult neural stem cell characteristics. *Cell*, *145*, 1142–1155.

Causeret, F., Moreau, M. X., Pierani, A., & Blanquie, O. (2021). The multiple facets of Cajal-Retzius neurons. *Development*, *148(11):dev199409*.

Denoth-Lippuner, A., & Jessberger, S. (2021). Formation and integration of new neurons in the adult hippocampus. *Nature Reviews. Neuroscience*, *22*, 223–236.

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*, 1644–1655.

Furutachi, S., Miya, H., Watanabe, T., Kawai, H., Yamasaki, N., Harada, Y., Imayoshi, I., Nelson, M., Nakayama, K. I., Hirabayashi, Y., & Gotoh, Y. (2015). Slowly dividing neural progenitors are an embryonic origin of adult neural stem cells. *Nature Neuroscience*, *18*, 657–665.

Iorio, F., Bernardo-Faura, M., Gobbi, A., Cokelaer, T., Jurman, G., & Saez-Rodriguez, J. (2016). Efficient randomization of biological networks while preserving functional characterization of individual nodes. *BMC Bioinformatics*, *17*, 542.

Iyer, A., & Tole, S. (2020). Neuronal diversity and reciprocal connectivity between the vertebrate hippocampus and septum. *Wiley Interdisciplinary Reviews: Developmental Biology*, *9*, e370.

Jessberger, S., Toni, N., Clemenson, G. D., Jr., Ray, J., & Gage, F. H. (2008). Directed differentiation of hippocampal stem/progenitor cells in the adult brain. *Nature Neuroscience*, *11*, 888–893.

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. *Nature Neuroscience*, *9*, 173–179.

Kriegstein, A., & Alvarez-Buylla, A. (2009). The glial nature of embryonic and adult neural stem cells. *Annual Review of Neuroscience*, *32*, 149–184.

Llorca, A., Ciceri, G., Beattie, R., Wong, F. K., Diana, G., Serafeimidou-Pouliou, E., Fernández-Otero, M., Streicher, C., Arnold, S. J., Meyer, M., Hippenmeyer, S., Maravall, M., & Marin, O. (2019). A stochastic framework of neurogenesis underlies the assembly of neocortical cytoarchitecture. *eLife*, *8*, e51381.

Meyer, B. D. (2019). Proxy: Distance and similarity measures. R package version 0.4-26. https://cran.r-project.org/web/packages/proxy/proxy.pdf

Namba, T., Mochizuki, H., Onodera, M., Mizuno, Y., Namiki, H., & Seki, T. (2005). The fate of neural progenitor cells expressing astrocytic and radial glial markers in the postnatal rat dentate gyrus. *The European Journal of Neuroscience*, *22*, 1928–1941.

Obernier, K., & Alvarez-Buylla, A. (2019). Neural stem cells: Origin, heterogeneity and regulation in the adult mammalian brain. *Development*, *146*, dev156059.

Ortiz-Álvarez, G., Daclin, M., Shihavuddin, A., Lansade, P., Fortoul, A., Faucourt, M., Clavreul, S., Lalioti, M.-E., Taraviras, S., Hippenmeyer, S., Livet, J., Meunier, A., Genovesio, A., & Spassky, N. (2019). Adult neural stem cells and multiciliated ependymal cells share a common lineage regulated by the geminin family members. *Neuron*, *102*, 159–172.

Pilz, G.-A., Bottes, S., Betizeau, M., Jörg, D. J., Carta, S., Simons, B. D., Helmchen, F., & Jessberger, S. (2018). Live imaging of neurogenesis in the adult mouse hippocampus. *Science*, *359*, 658–662.

Ratz, M., Von Berlin, L., Larsson, L., Martin, M., Westholm, J. O., Manno, G. L., Lundeberg, J., & Frisén, J. (2022). Clonal relations in the mouse brain revealed by single-cell and spatial transcriptomics. *Nature Neuroscience*, *25*(3), 285–294.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., Hao, Y., Stoeckius, M., Smibert, P., & Satija, R. (2019). Comprehensive integration of single-cell data. *Cell*, *177*, 1888–1902.

Suh, H., Consiglio, A., Ray, J., Sawai, T., D'Amour, K. A., & Gage, F. H. (2007). In vivo fate analysis reveals the multipotent and self-renewal capacities of Sox2 + neural stem cells in the adult hippocampus. *Cell Stem Cell*, *1*, 515–528.

Wagner, D. E., Weinreb, C., Collins, Z. M., Briggs, J. A., Megason, S. G., & Klein, A. M. (2018). Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science*, *360*, 981–987.

Zeisel, A., Hochgerner, H., Lönnerberg, P., Johnsson, A., Memic, F., van der Zwan, J., Häring, M., Braun, E., Borm, L. E., Manno, G. L., Codeluppi, S., Furlan, A., Lee, K., Skene, N., Harris, K. D., Hjerling-Leffler, J., Arenas, E., Ernfors, P., Marklund, U., & Linnarsson, S. (2018). Molecular architecture of the mouse nervous system. *Cell*, *174*, 999–1014.

Zeisel, A., Muñoz-Manchado, A. B., Codeluppi, S., Lönnerberg, P., Manno, G. L., Juréus, A., Marques, S., Munguba, H., He, L., Betsholtz, C., Rolny, C., Castelo-Branco, G., Hjerling-Leffler, J., & Linnarsson, S. (2015). Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science*, *347*, 1138–1142.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: von Berlin, L., Westholm, J. O., Ratz, M., & Frisén, J. (2023). Early fate bias in neuroepithelial progenitors of hippocampal neurogenesis. *Hippocampus*, *33*(4), 391–401. https://doi.org/10.1002/hipo.23482