Mapping human adult hippocampal neurogenesis with single-cell transcriptomics

Tosoni, Giorgia; Ayyildiz, Dilara; Bryois, Julien; Macnair, Will; Fitzsimons, Carlos P; Lucassen, Paul J; Salta, Evgenia

published in
Neuron
2023

DOI (link to publisher)
10.1016/j.neuron.2023.03.010

document version
Publisher's PDF, also known as Version of record

Neuron

Viewpoint

Mapping human adult hippocampal neurogenesis with single-cell transcriptomics: Reconciling controversy or fueling the debate?

Highlights

- Single-cell profiling of adult hippocampal neurogenesis can offer key insights
- Methodological and conceptual confounders can impact the resulting datasets
- Sample size and stratification, data processing, and marker selection are critical
- Efforts should focus on optimization and public sharing of protocols and pipelines

Authors

Giorgia Tosoni, Dilara Ayyildiz, Julien Bryois, Will Macnair, Carlos P. Fitzsimons, Paul J. Lucassen, Evgenia Salta

Correspondence

e.salta@nin.knaw.nl

In brief

Tosoni et al. probe the challenges related to the design, analysis, and interpretation of single-cell transcriptomic studies in the adult human hippocampal neurogenic niche and propose a series of critical points of attention via re-analysis and meta-analysis of previously published datasets.

Tosoni et al., 2023, Neuron *111*, 1–18
June 7, 2023
https://doi.org/10.1016/j.neuron.2023.03.010

CellPress

Neuron

CellPress
OPEN ACCESS

Viewpoint

Mapping human adult hippocampal neurogenesis with single-cell transcriptomics: Reconciling controversy or fueling the debate?

Giorgia Tosoni,[1] Dilara Ayyildiz,[1] Julien Bryois,[2] Will Macnair,[2] Carlos P. Fitzsimons,[3] Paul J. Lucassen,[3,4] and Evgenia Salta[1,5,*]

[1]Laboratory of Neurogenesis and Neurodegeneration, Netherlands Institute for Neuroscience, 1105 BA, Amsterdam, the Netherlands

[2]Roche Pharma Research and Early Development, Neuroscience and Rare Diseases, Roche Innovation Center, CH-4070, Basel, Switzerland

[3]Brain Plasticity group, Swammerdam Institute for Life Sciences, Faculty of Science, University of Amsterdam, 1098 XH, Amsterdam, the Netherlands

[4]Center for Urban Mental Health, University of Amsterdam, 1098 SM, Amsterdam, the Netherlands

[5]Lead contact

*Correspondence: e.salta@nin.knaw.nl

https://doi.org/10.1016/j.neuron.2023.03.010

SUMMARY

The notion of exploiting the regenerative potential of the human brain in physiological aging or neurological diseases represents a particularly attractive alternative to conventional strategies for enhancing or restoring brain function. However, a major first question to address is whether the human brain does possess the ability to regenerate. The existence of human adult hippocampal neurogenesis (AHN) has been at the center of a fierce scientific debate for many years. The advent of single-cell transcriptomic technologies was initially viewed as a panacea to resolving this controversy. However, recent single-cell RNA sequencing studies in the human hippocampus yielded conflicting results. Here, we critically discuss and re-analyze previously published AHN-related single-cell transcriptomic datasets. We argue that, although promising, the single-cell transcriptomic profiling of AHN in the human brain can be confounded by methodological, conceptual, and biological factors that need to be consistently addressed across studies and openly discussed within the scientific community.

INTRODUCTION

Harnessing the endogenous regenerative potential of the human brain during aging or in neurodegenerative and neuropsychiatric conditions represents an intriguing concept that has attracted considerable attention from the general public and sparked intense controversies within the field. However, whether adult hippocampal neurogenesis (AHN) persists in the human brain is a question that continues to be extensively debated.[1–8] Earlier studies employing 5-bromo-2'-deoxyuridine (BrdU) labeling of dividing cells or carbon-14 (^{14}C)-based retrospective birth dating of neurons in postmortem adult human dentate gyrus (DG) suggested the existence of adult-born neuronal populations in human hippocampus.[9,10] However, immunohistochemical studies based on the detection of presumably specific markers for dividing neuronal progenitors, neuroblasts (NBs), and immature neurons (ImNs) have reported a wide range of AHN rates in postmortem human brain.[11–16] Both conceptual and methodological confounders have likely contributed to these seemingly opposing observations, as recently discussed.[3,8,17] Hence, elucidating the extent of (functional) AHN and identifying putative neurogenic cell populations and states in human brain in aging or during neurodegeneration remains a challenge.

Recent advances in single-cell transcriptomic technologies have provided valuable insights into the molecular profiling of distinct cell types in postmortem human brain and in several brain diseases.[18–21] Similarly, moving beyond the identification of cell types and states based on only a limited set of marker genes, single-cell or single-nucleus RNA sequencing (scRNA-seq or snRNA-seq) represents a powerful strategy for leveraging combinatorial gene expression profiles to unravel the full complexity and dynamics of the presumed neurogenic niche in the adult human hippocampus.[22–24] However, recent single-cell transcriptomic studies in the adult human hippocampus have yielded conflicting results: two of them did identify cells with neural stem cell (NSC) or immature neuronal characteristics,[22,23] whereas a third one reported an inability to detect any neurogenic populations.[24] Hence, are these novel approaches—once again—failing to finally settle the controversy regarding the existence of AHN in humans? Will we eventually be able to overcome the conceptual and technical challenges and reconcile these—seemingly—opposing views and findings?

Please cite this article in press as: Tosoni et al., Mapping human adult hippocampal neurogenesis with single-cell transcriptomics: Reconciling controversy or fueling the debate?, Neuron (2023), https://doi.org/10.1016/j.neuron.2023.03.010

Here, we caution that the design, analysis, and interpretation of sc/snRNA-seq studies in the adult human hippocampal neurogenic niche (i.e., the spatially defined complex microenvironment supporting the transition from adult NSCs toward neural progenitors, NBs, ImNs, and, eventually, mature granule cells [GCs]) can be confounded by specific issues, which ask for conceptual, methodological, and computational adjustments. By reanalyzing previously published datasets, we probe a series of specific challenges that require particular attention and would greatly profit from an open discussion in the field. In the following sections, we discuss these important considerations in detail. We argue that these novel technologies, when appropriately applied, offer a unique opportunity to map AHN in the human brain, assess which putatively neurogenic populations are present, compare the neurogenic process across species, and explore which cell types and states may be possibly most amenable to therapeutic interventions in aging, and neurodegenerative and neuropsychiatric diseases. We propose that there may actually be more that we agree on than previously believed.

RESULTS

(Sample) size matters: Achieving adequate power

The number of cell types identified in sc/snRNA-seq studies critically depends on the protocol used,[25] the number of cells sequenced (the term "cells" will be used from here on to collectively refer to the single observations in sc/snRNA-seq studies, i.e., cells or nuclei), the sequencing depth,[26] and the clustering methods employed.[27] Being able to capture all cell types, including rare or transient cell states, and distinguish them from each other requires the appropriate *a priori* estimation of how many cells and which sequencing depth will be aimed for. Larger sample sizes increase the odds of capturing minute cellular subpopulations in a tissue. The relative decline in sequencing cost in recent years has now enabled the generation of gene expression profiles of many thousands of cells per experiment.[28] Given the expected low abundance of neurogenic cell populations in the adult human hippocampus, it would be reasonable to assume that achieving the appropriate power in human AHN studies is critical. To our knowledge, there are only 5 published studies to date surveying the transcriptomic profile of the adult human hippocampus at single-cell resolution (Figure 1A). These studies span from a few thousands[23,29] to hundreds of thousands of profiled cells.[22,24,30]

When trying to estimate the appropriate power for sc/snRNA-seq studies, it is important to factor in that the relationship between the probability of identifying rare cell types and the number of cells sequenced is not linear. The minimal number of sequenced cells required to observe a rare cell type at a certain frequency can be modeled with a negative binomial (Satija-how many cells [https://satijalab.org/howmanycells/] and Abrams et al.[31]) or multinomial distribution[32] (Figure 1B).

Using previous knowledge of cell proportions in adult human hippocampus from published immunohistochemical[4,14,15] and radiocarbon-based birth dating studies,[9] we estimated the expected fraction of postulated NBs/ImNs per total GCs in the human DG or per total DG cells (Figure 1C) (STAR Methods). We also calculated the expected size of these cellular populations per 30,000 GCs, which is the number of GCs sequenced in the studies with the highest statistical power published this far.[22,24] We then used available computational tools (https://satijalab.org/howmanycells) to estimate the sample sizes of sequenced single cells that would be required to detect the expected number of doublecortin $(DCX)^+$ cells, presumed to be NBs/ImNs (DCX-based definitions of cell identity were inferred from each of the studies, and may be confounded by limited cell type specificity of DCX[1,24,33]) per 30,000 GCs using sc/snRNA-seq. We found that in most cases, the calculated numbers of cells required for detecting such rare cell types by sc/snRNA-seq technology are much higher than the numbers of cells sequenced in previous reports (Figure 1C), which could, at least partially, explain why some of these studies failed to detect neurogenic signatures in their datasets.

Franjic et al. implemented a similar strategy (Table S2 in Franjic et al.[24]) to estimate the numbers of NBs expected in their dataset. However, those calculations did not account for sc/snRNA-seq technology constraints (i.e., technical and biological factors affecting the relationship between the number of cells sequenced and the probability of reliably identifying rare cell populations[32,34,35]), assuming linearity, thereby overestimating the theoretically expected numbers of neurogenic populations in their dataset. This led the authors to propose that if such high numbers of NBs/ImNs were present in their samples, they should have been able to detect them with their snRNA-seq approach. The inability to do so led them to argue against the existence of AHN in the human brain.

Wang et al. identified populations with radial glial-like cell (RGL, the term is used hereafter to collectively describe adult NSCs) and ImN profiles in their dataset, which only consisted of a total of ~22,000 cells.[23] However, Franjic et al. and Zhou et al. reached conflicting results, although they sequenced approximately 6 times more cells than Wang et al.[22,24] Of note, nuclei in the Wang et al. study[23] were sequenced considerably deeper (85K reads per nucleus), compared with the other available datasets (Figure 1A), putting forward the sequencing depth as a critical factor contributing to the identification of putatively neurogenic populations. Whether the cellular populations identified by Wang et al.[23] represent *bona fide* adult RGLs and ImNs or are rather subclusters of the bigger glial and neuronal populations with distinct transcriptional profiles would require further analysis (see "Finding the holy grail: On the quest for human cell type-specific neurogenic markers"). Evidently, solely relying on the binomial probabilistic model to infer required cell numbers may be too simplistic: other critical technical and biological factors not considered by this model (see "Sample and dasatet processing: The risk of throwing the baby out with the bathwater") may further affect the capability to distinguish rare cell types.

Another key variable related to the sample size that can impact the detection of AHN in the human brain is subject stratification. The rate of AHN may highly vary across individuals, as was recently also shown in non-human primates, where 22 of the 23 annotated neural progenitor cells (NPCs) were detected in only 1 of the 3 animals[24] (Figure S1) and in humans.[4,14,15,22] In addition, the proportions of neurogenic populations are expected to decline dramatically with aging and be influenced by genetic background, existing pathologies, medication, and

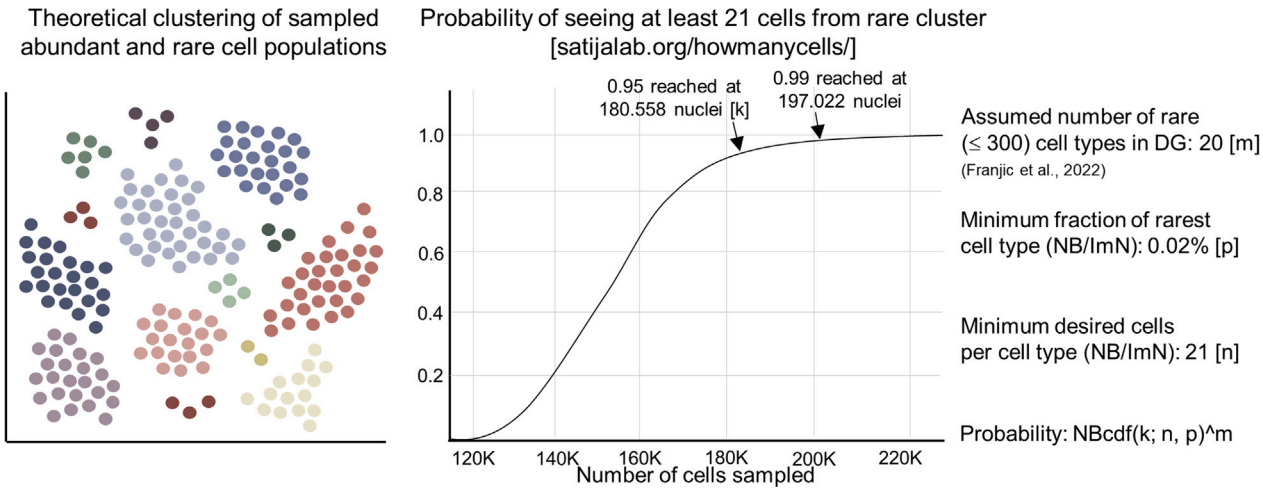
Please cite this article in press as: Tosoni et al., Mapping human adult hippocampal neurogenesis with single-cell transcriptomics: Reconciling controversy or fueling the debate?, Neuron (2023), https://doi.org/10.1016/j.neuron.2023.03.010

CellPress
OPEN ACCESS

A

Brain region	Nr. donors	Age (years)	Nr. total nuclei (after QC)	Nr. DG granule cells	Sequencing depth (reads/nucleus)	Clinical documentation	Avg RIN	Avg PMD (hours)	Reference
Hippocampus Cortex	5	40-65	14 137	1453	10K	'Non-diseased'	7.3	12.4	Habib et al. 2017
Hippocampus	4	67-92	22 119	Unknown	85K	'Without neurological disorders'	Unknown	10	Wang et al. 2022
Hippocampus	5	26-60	129 908	~9K	20K	Temporal lobe epilepsy	Unknown	NA (fresh tissue)	Ayhan et al. 2021
Dentate gyrus	6	44-79	139 187	32 067	20K	'Clinically unremarkable' (5x) & status epilepticus (1x)	Unknown	15.6	Franjic et al. 2022
Hippocampus	38	0,1-88	152 184	35 187	23K	Free from neurological disorders & AD	Unknown	17.1	Zhou et al. 2022

B



C

Species	Detection Method	Cell proportion Calculation	Cell proportion (of total GCs)	Cell proportion (of total DG cells)	Expected # cells per 30.000 GCs with immunolabeling	Total DG nuclei required to detect the same number of cells with snRNAseq	Reference
Human	DCX Immunolabeling	10.345 DCX⁺ cells (Boldrini et al., 2018) out of 15M GCs (Jonas et al., 2014). GC number is 23% of total DG cell number (Franjic et al., 2022)	0.07% DCX⁺ cells	0.02% DCX⁺ cells	21	180.000 (to detect 21 NBs/ImNs)	Boldrini et al. 2018
Human	DCX Immunolabeling	127.342 DCX⁺ cells (Tobin et al., 2018) out of 15M GCs (Jonas et al., 2014). GC number is 23% of total DG cell number (Franjic et al., 2022)	0.85% DCX⁺ cells	0.2% DCX⁺ cells	255	>500.000 (to detect 255 NBs/ImNs)	Tobin et al. 2019
Human	DCX Immunolabeling	32.000 DCX⁺ cells/mm³ out of 840.000 GCs/mm³ (Moreno-Jimenez et al., 2019). GC number is 23% of total DG cell number (Franjic et al., 2022)	3.8% DCX⁺ cells	0.87% DCX⁺ cells	1140	>500.000 (to detect 1140 NBs/ImNs)	Moreno-Jimenez et al. 2019
Human	^{14}C incorporation	700 newly generated GCs/day (Spalding et al., 2013). Estimation of total newborn GCs per 15M GCs (Franjic et al., 2022; Jonas et al., 2014). GC number is 23% of total DG cell number (Franjic et al., 2022)	0.84% ImNs	0.2% DCX⁺ cells	252	>500.000 (to detect 252 ImNs)	Spalding et al. 2013

Figure 1. Estimating the appropriate power for sc/snRNA-seq studies of human AHN

(A) Summary of the published snRNA-seq studies in adult human hippocampus.

(B) Probability estimation using How Many Cells|Satija Lab online software (https://satijalab.org/howmanycells/) assessing how many cells need to be sampled to detect at least *n* cells of each type. For a given cell type, the probability of seeing at least *n* cells in a sample of size *k* follows the cumulative distribution function of a negative binomial NBcdf (*k; n*, *p*), with *p* being the relative abundance.

(C) Table reporting the proportions of putative NBs/ImNs per total amount of DG cells, the expected numbers of NBs/ImNs per 30,000 GCs (number of GCs sequenced by the most powered published studies to date), and the number of cells that need to be sequenced to reach the same expected number of cells when using sc/snRNA-seq, according to different references. Avg, average; RIN, RNA integrity number; PMD, postmortem delay; NBs, neuroblasts; ImNs, immature neurons; GCs, granule cells.

Please cite this article in press as: Tosoni et al., Mapping human adult hippocampal neurogenesis with single-cell transcriptomics: Reconciling controversy or fueling the debate?, Neuron (2023), https://doi.org/10.1016/j.neuron.2023.03.010

lifestyle parameters in general.[36,37] Many individuals may have died from a serious illness or suffered chronic conditions, presenting for instance with higher inflammatory signatures, which could impact neurogenesis.[23,38–40] Alterations in the inflammatory indices have been proposed by Wang et al. as a putative factor explaining the contrasting results of their study to the findings of Franjic et al., where no neurogenic traces were detected in the adult human brain.[23,24] Following a similar approach to Wang et al. to assess inflammation, we observed a negative association between inflammation scores, which we computed in the three macaques that were analyzed by Franjic et al.[24] (STAR Methods), and the proportion of the total neurogenic populations identified in each of the three animals (Figure S1). Evidently, given the small sample size, these results should be interpreted with caution. Cell types other than the astrocytes analyzed here (e.g., microglia) could also impact inflammatory processes in the hippocampal niche. Importantly, inflammation alone may not be able to broadly explain the discrepancy between the recently published snRNA-seq studies in human postmortem DG: based on the available data, we cannot exclude that inflammation levels may differentially impact distinct neurogenic populations since we did not find the proportion of the RGLs themselves in the three macaques of Franjic et al. to be associated with the inflammation scores in these animals (Figure S1) (see also the section Finding the holy grail: On the quest for human cell type-specific neurogenic markers). Although all the above factors together prompt the importance of adequate sample size (in this case defined as the number of brain donors) when surveying the neurogenic process in the adult human hippocampus, most of the current studies (4 of 5, Figure 1A) have focused on only a few subjects, possibly jeopardizing a reliable identification of small (neurogenic) populations.

Sample and dataset processing: The risk of throwing the baby out with the bathwater

Cell type identification in sc/snRNA-seq is a multi-step process starting from sample collection, proceeding to cell/nuclei isolation, and terminating in the data analysis. The number of cells sequenced is not the only parameter affecting the successful identification of rare and uncharacterized cell types within a dataset.

Access to freshly resected human tissue is often limited; therefore, most single-cell transcriptomic studies (4 of 5, Figure 1A) were conducted postmortem on frozen tissue, from which nuclei, instead of cells, are isolated. Several comparative studies reported a similar performance of scRNA-seq and snRNA-seq in identifying specific cellular populations[41–45] and highlighted certain advantages of snRNA-seq over scRNA-seq, including reducing bias in cell population enrichment that is caused by cell dissociation and cell size per se, minimizing cellular stress-induced transcriptional artifacts, and allowing to capture nascent mRNA for temporal studies. Indeed, snRNA-seq studies have identified disease-specific cellular states in the human brain.[23,46,47] However, other studies reported differences in possible enrichment and detection bias for certain transcripts between snRNA-seq and scRNA-seq.[48–50] Such differences between whole-cell versus nucleus RNA-seq could be partially attributed to technical issues associated with sampling and recovery of distinct subsets of the transcriptome. This could potentially also explain the difficulty of snRNA-seq in detecting a significant number of transcripts of proliferation markers, such as *MKI67*, in amplifying neuronal progenitors, even in hippocampal samples from human infants, where many KI67-immunolabeled proliferating cells have been depicted.[15,22,23]

Sample quality and tissue preservation and processing are crucial parameters for immunohistochemical studies investigating AHN and have been therefore at the center of the ongoing debate.[2–4] When working with postmortem material, a key factor affecting the quality of the tissue is the postmortem delay (PMD), which is defined as the period between the time of death or the cessation of blood flow and the time the sampled tissue is frozen, stabilized, and/or preserved. Increasing PMD was shown to compromise the intensity of DCX immunolabeling and impact the number of DCX-positive cells in the adult DG of rats (at 1 h PMD, decreased dendritic labeling; at 12 h PMD, fewer DCX-positive cells[51]) and macaques (at 16 h PMD, reduction of DCX-immunolabeled cells[24]). Although such systematic time course analyses have not been performed in sc/snRNA-seq studies yet, Franjic et al. recently demonstrated that *DCX* expression in NPCs of adult pig DG declines sharply at 7 h of PMD in their snRNA-seq dataset.[24] Whether such effects are also observed for other AHN-related transcripts and cellular populations is currently not known. Although no causal links can be drawn at the moment, the shortest average PMD across all recent snRNA-seq studies in the adult human DG corresponds to the only one that reported the identification of adult RGLs.[23] Temperature and time play a fundamental role in RNA stability. Most of the changes at the transcript level occur as early as 30 min postmortem, whereas some can become apparent 24 or 48 h after death, depending on the tissue type, patient's health, and death process, which can affect the pH of the brain tissue.[52,53] Notably, after 24 h, a reduction in specific cell types, higher percentage of mitochondrial reads, and increased background ambient RNA (which can introduce noise in the resulting data) are observed in a tissue-specific manner.[54–56] Importantly, the nuclei (or cell) isolation protocol itself may additionally play a fundamental role in the results obtained.[57,58]

All things considered, obtaining consistent numbers of clinically unremarkable, well-matched, good-quality samples and, thereby, (adequate numbers of) reliable single-cell transcriptomes remains particularly difficult. Often, clinical information, reports on sample quality, and detailed protocols are missing or are not adequately described (for instance, missing information on RIN values, Figure 1A), thereby hampering the ability to reproduce and interpret the contrasting results obtained by different studies. The use of a conventional set of parameters to document clinical information could help optimize data comparison across different studies and, thereby, contribute to increasing reproducibility. Efforts to optimize (and publicly share!) protocols for nuclei isolation that aim at minimizing the loss of specific cell types from adult human hippocampal specimens are, therefore, invaluable (examples in Ayhan et al.[59]).

Regarding data processing, extra caution should be applied when exploring human AHN, as the ultimate aim in that case would be the identification of putatively rare and previously uncharacterized cellular populations. The cell types of interest

could on average exhibit low gene expressions (like RGLs residing in a quiescent state), which in some cases may not cross the cutoff points for thresholding low-quality nuclei: for instance, we find lower gene numbers in RGLs in a mouse scRNA-seq dataset,[60] a macaque snRNA-seq study,[24] and in NPCs in a macaque scRNA-seq dataset, relatively to the other cell clusters within each dataset (Figure 2), although the additional impact of dataset-intrinsic methodological or computational variables cannot be ruled out. Furthermore, given the continuity of the neurogenic process, some cells presenting with "doublet-like" features and expressing markers of multiple cell types may in fact represent transitioning states.[24,61] Therefore, standard sc/snRNA-seq data preprocessing pipelines may need to be adapted to avoid filtering out true neurogenic populations. Interestingly, certain cell clusters of potential relevance may be removed from a dataset due to unknown identity and/or the small number of cells.[30]

Another concern relates to the common practice of clustering and trajectory inference methods, which is based on selecting highly variable genes and performing dimension reduction using principal-component analysis (PCA) and t-distributed stochastic neighbor embedding/uniform manifold approximation and projection (tSNE/UMAP).[64] This strategy has proven effective in identifying major cellular phenotypes. However, the focus on selecting highly variable genes can impact the results of the subsequent clustering and trajectory analysis: genes that are not highly variable (and therefore filtered out) may be more useful for defining rare cell subpopulations.[65,66] The lack of markers for rare cell types would motivate the use of unsupervised clustering approaches. Method development[67,68] and comparative analysis[31,69–71] of tools for effective unsupervised clustering of cells are being already performed. However, the identification of such rare subpopulations, solely based on unsupervised clustering of a single dataset, remains particularly challenging[72] and often requires manual annotation, which, besides being time-consuming, introduces an additional level of bias. The use of more sophisticated supervised machine learning models could help solve this issue. Several tools aiming at identifying rare cell populations using machine learning are in development.[73–76] In a recent publication, Zhou et al. developed a machine learning analytical approach that was instrumental for the identification of ImNs from postmortem human brain in their snRNA-seq dataset.[22] More specifically, as a training set, the authors employed transcriptional signatures derived from ImN-like populations in early postnatal mouse hippocampus and infant human brain, which were positive for *DCX* and *PROX1*, a pan-GC marker, and negative for *CALB1*, which has been shown to be specific to mature neurons and largely absent in ImNs in the mouse hippocampus.[77,78] Subsequently, weighted panels of "prototype" molecular features for each species were imputed and tested in the mouse and human hippocampus across ages. ImN signatures could be identified by the trained model in adult mice and human hippocampus using a conservative similarity score cutoff. Since this supervised learning approach depends on the initial marker selection panel to classify "prototypical cell types," it may provide a snapshot of only a specific cellular subpopulation with particular ImN-like characteristics. Whether these cells represent *bona fide* adult-born ImNs or not remains to be addressed. Hence, further customization of data processing and analysis tools may be particularly critical for the mapping and characterization of neurogenic populations in postmortem human brain.

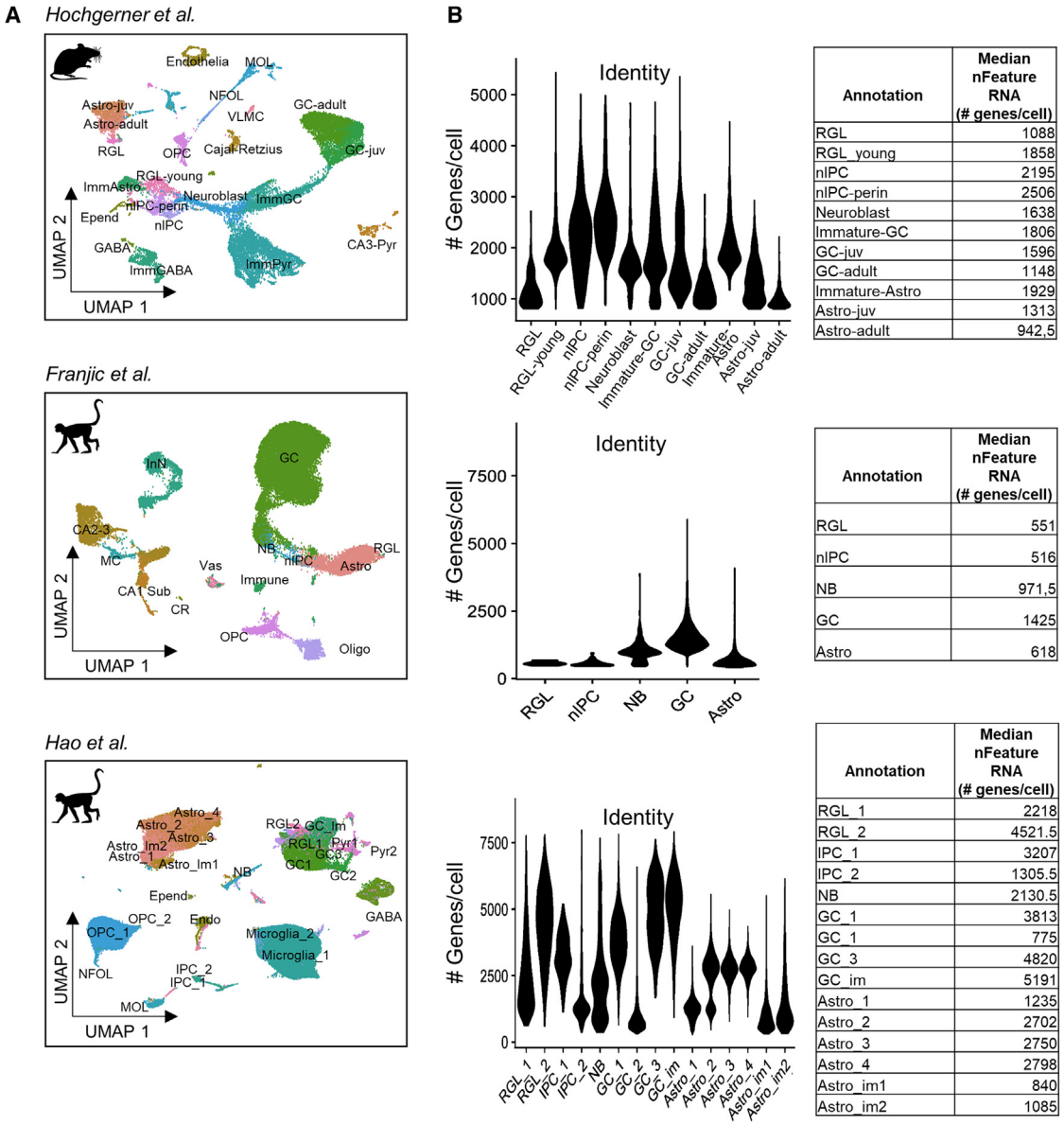
RGL cells and astrocytes: An inseparable duo

The neurogenic trajectory in the adult mammalian brain has been described as a molecular continuum of dynamic cell states.[23,24,60,79–81] Identifying *bona fide* RGLs, with the potential to become activated by environmental stimuli, thereby triggering the neurogenic cascade, would provide the ultimate proof for the existence of functional AHN in the human brain. In rodents, adult RGLs constitute a minor population within the complex tissue microenvironment, intermingled with their progeny at different developmental stages and with several other supporting cells that together form a heterogeneous cell layer.[79] Quiescent RGLs become activated and give rise to proliferating precursors, which then differentiate further into NBs and eventually ImNs.[82] Due to the scarcity of RGLs and their transcriptional similarity to astrocytes (indicative of their common ancestry[83]), adult RGL and astrocytic populations often cluster together when dimensionality reduction methods are applied (e.g., tSNE and UMAP) in sc/snRNA-seq datasets derived from various species.[23,24,79,81,84]

The highly overlapping expression of markers between RGLs and astrocytes[2,85] is a major roadblock to reliably separating the two populations with sc/snRNA-seq, even in species with unarguably robust levels of AHN, like rodents. When processing a relatively small dataset employing standard pipelines, neither principal component nor differential expression or pseudotemporal trajectory analyses could detect significant differences between quiescent RGLs and astrocytes within the mouse subventricular neurogenic niche, another neurogenic niche of the adult rodent brain.[84] In a scRNA-seq profiling study in mouse DG, in the absence of an apparently distinct astrocytic cluster, the authors collectively referred to a population with mixed RGL/astrocytic characteristics as neural stem cells (indicated as NSCs[81]). In a much larger scRNA-seq dataset in adult mouse DG, however, Hochgerner et al. did identify an RGL cluster that was closely related to, but clearly distinct from, astrocytes,[60] similar to other large-scale studies in the subventricular zone,[86,87] suggesting that when appropriately powered, scRNA-seq can differentiate astrocytic and RGL molecular signatures, at least in rodents.

To explore the molecular signature of the RGL population in the study of Artegiani et al.,[81] we probed the expression of their top putative RGL markers (*Hopx*, *Aldoc*, *Gfap*, *Id4*, *Sox9*, *Apoe*, *Sox2*, and *Slc1a3*, Table S1) in the Hochgerner et al.[60] dataset (Figure 3A). Both the signature score of the combined gene set (calculated using Seurat AddModuleScore function) (Figure 3B) and each gene separately (Figures 3D and 3E) were detected not only in the RGL cluster but were also markedly expressed in astrocytes and, to a lesser extent, in perinatal NPCs ([neural] intermediate progenitor cell [nIPCs]). Comparisons of the averaged gene expression between RGLs and astrocytes did not yield a significant difference between the two clusters (although a trend for higher expression in astrocytes was observed) (Figure 3C). In addition, when using the module score calculated on this gene set as a method to classify RGLs, the precision

Please cite this article in press as: Tosoni et al., Mapping human adult hippocampal neurogenesis with single-cell transcriptomics: Reconciling controversy or fueling the debate?, Neuron (2023), https://doi.org/10.1016/j.neuron.2023.03.010



Annotation	Median nFeature RNA (# genes/cell)
RGL	1088
RGL_young	1858
nIPC	2195
nIPC-perin	2506
Neuroblast	1638
Immature-GC	1806
GC-juv	1596
GC-adult	1148
Immature-Astro	1929
Astro-juv	1313
Astro-adult	942,5

Annotation	Median nFeature RNA (# genes/cell)
RGL	551
nIPC	516
NB	971,5
GC	1425
Astro	618

Annotation	Median nFeature RNA (# genes/cell)
RGL_1	2218
RGL_2	4521.5
IPC_1	3207
IPC_2	1305.5
NB	2130.5
GC_1	3813
GC_1	775
GC_3	4820
GC_im	5191
Astro_1	1235
Astro_2	2702
Astro_3	2750
Astro_4	2798
Astro_im1	840
Astro_im2	1085

Figure 2. Distinct cell types present different quality control parameters

(A) UMAP plots of data from Hochgerner et al.,[60] representing 24,185 cells from mouse DG (upper panel), Franjic et al.[24] representing 36,107 nuclei from macaque DG (middle panel), Hao et al.[62] representing 207,785 cells from macaque DG (lower panel). Unique molecular identifier (UMI) count matrices from published datasets (summarized in Table S1) were retrieved from the respective repositories and processed independently, using the same criteria described in the original publications, into the R package Seurat (4.1.1).[63] Cells/nuclei (dots) are labeled and colored by cluster membership labels from the original study: Astro, astrocytes; GCs, granule cells; RGL, radial glia-like cells; (n)IPCs, (neural) intermediate progenitor cells; NBs, neuroblasts; juv, juvenile; im, immature.

(B) Violin plots and tables indicating the median number of genes (median nFeature RNA) per cell in each of the annotated clusters.

(positive predictive value calculated based on the numbers of identified cells in the RGL and astrocytic cluster expressing these genes) and the computed F_1-score (harmonic mean of precision and recall, employed here as a measure of cell type identification accuracy) was higher in astrocytes than that in RGLs (Figure 3B; Table S2; STAR Methods). Hence, proof-of-principle analysis employing a well-defined mouse DG dataset, suggests that putatively rare RGL populations could be "masked" by co-clustering astrocytes, particularly in the case of smaller datasets.

Evidently, the difficulty to reliably identify a distinct RGL population as a clear transcriptional outgroup becomes even more of a challenge in species other than rodents, where AHN may have been evolutionarily restricted leading to smaller RGL numbers in the DG, as it is the case in non-human primates and humans.[23,24,88] This is further complicated by the high inter-individual variability in the latter—partially arising from genetic variation being largely absent in inbred model species—discussed earlier, and the current critical lack of human-specific markers,

Please cite this article in press as: Tosoni et al., Mapping human adult hippocampal neurogenesis with single-cell transcriptomics: Reconciling controversy or fueling the debate?, Neuron (2023), https://doi.org/10.1016/j.neuron.2023.03.010

CellPress
OPEN ACCESS

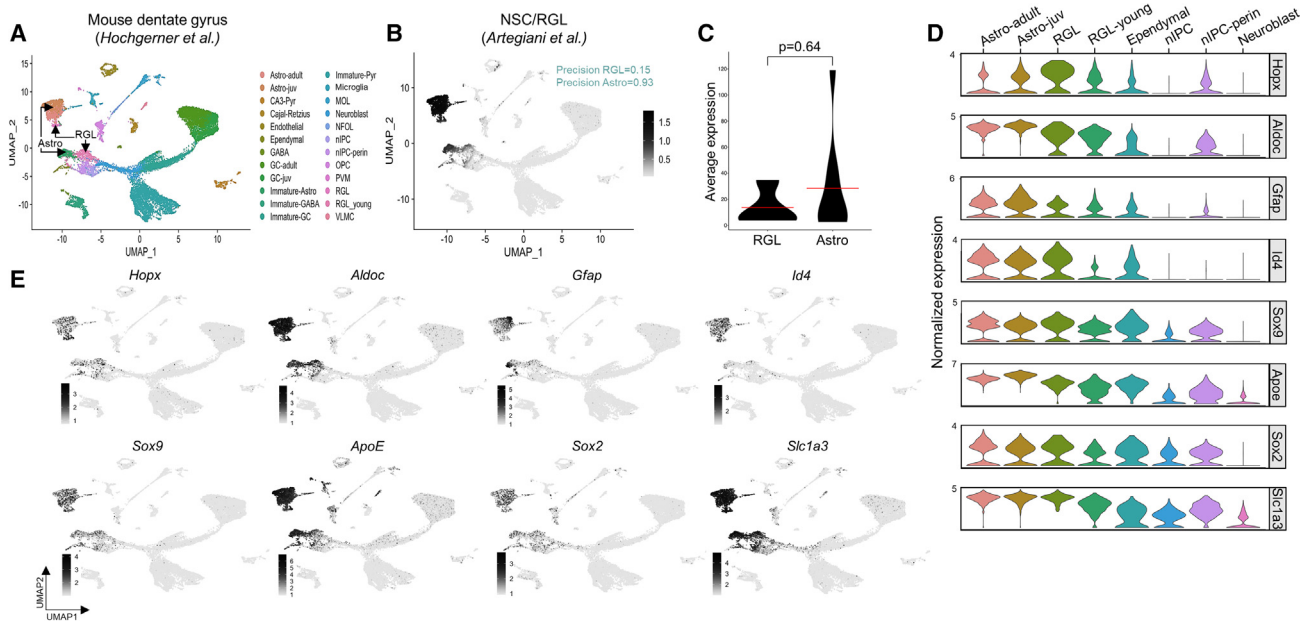


Figure 3. Differentiating RGLs from astrocytes in mouse DG

(A) UMAP plot of data from Hochgerner et al.[60] where Astro and RGLs separate as individual clusters (arrows).

(B) UMAP plot as in (A), colored by the Seurat module score calculated using RGL marker genes provided by Artegiani et al.[81] (Table S1). Blue captions indicate precision values of the RGL marker gene set of Artegiani et al.[81] in the RGL and astrocytic cluster of Hochgerner et al.[60] (imputed as described in STAR Methods and Table S2).

(C) Violin plot showing the distribution of the average expression of RGL marker genes[81] in clusters annotated as RGL and Astro in the Hochgerner dataset.[60] For statistical analysis, the Wilcoxon rank sum test was applied using the wilcox.test function implemented in R. Red line indicates mean expression; n = 8 genes.

(D) Stacked violin plots indicating the normalized expression of marker genes provided by Artegiani et al.[81] in the Hochgerner et al.[60] dataset. NSC, neural stem cell; RGL, radial glia-like cell; Astro, astrocyte; nIPCs, (neural) intermediate progenitor cells.

(E) UMAP plots as in (A), colored by the level of expression of single RGL marker genes provided by Artegiani et al.[81]

which is thoroughly addressed in the next sections. Marker-based fluorescence-activated cell sorting (FACS) of RGL populations or other enrichment strategies may increase the power of such approaches. However, to identify such markers, credible cluster annotation and RGL/astrocyte differentiation will need to first be achieved. Interestingly, simultaneous single-cell profiling of transcriptomic and epigenetic signatures in the mouse subventricular zone, recently identified differential DNA methylation profiles between RGLs and astrocytes, despite similar gene expressions.[89] These findings suggest that integrating methylome and transcriptome information could potentially be instructive to delineate cell identity of neurogenic populations also in the human brain.

Of mice and humans: Same same but different

Neurogenic cell types in the human brain have been so far largely defined on the basis of expression of (a few) rodent-inferred markers.[4,11,12,14,24] In the absence of more unbiased, genome-wide approaches to molecularly characterize these cells in humans, the use of (a limited number of) markers derived from rodents has been a central argument against the validity of the published histological studies in human postmortem brain.[1,2] Indeed, given the evolutionary differences between mice and humans (e.g., species longevity, brain size constraints, and cellular maturation period[90,91]), the molecular signatures of specific cell types/states involved in the neurogenic process in the two species may very likely have diverged as well. sc/snRNA-seq technologies are best suited to overcome these limitations, as they can in principle capture the full molecular and cellular dynamics of the neurogenic niche based on combined expression of known, as well as novel, markers.

Dcx has been traditionally used to label ImNs in rodents,[92,93] and DCX immunolabeling has been commonly considered a proxy for AHN also in humans.[11,14,16,94] However, different (non-neurogenic) cell types were shown to express DCX under certain conditions in postmortem human brain tissue (low correlation between mRNA and protein expression levels could also explain differences across technologies).[1,2,24,33] Recently, in a first attempt to profile the neurogenic trajectory in the adult human DG with snRNA-seq, Franjic et al. could not identify a defined population of putative *DCX*-positive NBs or ImNs.[24] The authors reported *DCX* expression in several neuronal populations, including GCs and interneurons. This led them to quite reasonably suggest that inferring AHN rates solely from *DCX* expression patterns would lead to misleading interpretations. However, rather surprisingly, the absence of a single cell cluster specifically expressing the—otherwise unreliable—*DCX* marker led the authors to argue against the existence of AHN in humans.

The interspecies variance in the transcriptomic profiles of niche-resident cellular populations was further confirmed by

three independent snRNA-seq studies.[22,23,30] In particular, Wang et al. reported a set of primate-specific RGL signatures in the DG on integrated cross-species analysis between mice and macaques.[23] Similarly, Zhou et al. identified age-dependent transcriptional alterations in human immature GCs regarding cellular functionality, intercellular interactions, and disease relevance, which were distinct from those in mice.[22] Using machine learning, the authors demonstrated that orthologous genes specific to human and mouse ImNs exhibit substantial species differences.

To address whether mouse-inferred neurogenic markers could be suitable for the human brain, we surveyed common and divergent gene expression profiles in distinct neurogenesis-related cell types (astrocytes, RGLs, NPCs, NBs, ImNs, and GCs) between the two species, by performing a systematic meta-analysis of (single-cell) transcriptomic datasets from mouse (hippocampus-only studies) and human brain (Figure 4A; STAR Methods; Table S3). To date, there are no transcriptomic studies characterizing the full neurogenic trajectory in the adult human hippocampus. Therefore, the human marker genes were derived either from developmental studies or datasets from other brain regions. For all cell types assessed, only a small overlap was observed between the two species (Figure 4B). Of note, no common marker genes were identified for RGLs, again hinting at the difficulty to effectively segregate this cell type from other cellular populations in the human brain. Next, we tested putative expression differences of conserved markers across the two species versus human-only markers using Seurat module scores in published snRNA-seq datasets (Figures 4D, S2A, and S2B). We employed macaque DG datasets,[24,62] since macaque studies are—to our knowledge—the only primate resources available to date that identified a clear and comprehensive neurogenic trajectory in adult DG, along with one of the available snRNA-seq datasets from human DG[24] (Figure 4C). In the two macaque datasets, the human-specific markers for NPCs, NBs, astrocytes, and GCs yielded similar or higher precision and F_1-scores (a measure of accuracy) compared with the mouse-human conserved counterparts (Figures 4D and 4E; Table S2). Interestingly, human-specific NPC markers showed considerably higher precision and accuracy when compared with the conserved ones only in the first of the two macaque datasets,[24] whereas NB markers (both human-specific and conserved) yielded higher precision and accuracy in the second macaque dataset.[62] In the human dataset,[24] we observed similar precision and F_1-scores between the human-specific markers and the conserved ones for astrocytes and GCs.

Our analysis indicates that astrocytes and GCs are generally easier to distinguish from other more immature or progenitor-like cell types within each dataset using marker-based approaches. Neurogenic cell types were less consistently identified across datasets than GCs and astrocytes. This observation is in line with previous reports demonstrating a limited level of premitotic diversity in neuronal progenitors, compared with the more "crisply" defined mature neurons.[95–97] Overlay between ImN and NB human markers (based on the recently reported transcriptional similarity between these two populations[22]) did not increase cell type identification precision or accuracy for the NB population (Figure S2C; Table S2). Our selection of human datasets also included studies not performed in the hippocampus or datasets from the embryonic human brain. This could—at least partially—explain why the human-specific gene sets did not perform equally well for all cell types. Still, human-specific markers in our meta-analysis yielded overall higher precision and accuracy (especially in NPCs) than genes identified in both humans and mice, suggesting that mouse-inferred markers may not be suitable for studying AHN in the human brain.

More evolutionary issues: Is species conservation overrated?

As more sc/snRNA-seq datasets from the mammalian DG become available, a more evolutionary "holistic" approach can now be used to screen for neurogenesis-related populations in the adult human brain. Instead of being limited to markers only derived from rodent studies, employing markers conserved across multiple species and cross-species data integration was recently put forward as a more powerful approach. In a recent work encompassing data from multiple species (mouse, pig, and macaque),[24] the authors implemented this strategy to screen for NPC and NB populations in their human dataset (Figures 5A and 5B). Only two cells (one NPC and one NB) were found to co-express respective NPC (*PROX1/TOP2A/CENPF/MKI67*) and NB (*PROX1/DCX/CALB2/NEUROD6/DPYSL3*) species-conserved markers and to cluster in the appropriate UMAP space, leading the authors to conclude that neurogenesis is technically absent in the adult human DG. Surprisingly, when we used, as positive control, the same conserved-across-multiple-species markers to screen for NPCs and NBs in the macaque dataset of the same study (where a well-defined neurogenic trajectory was outlined) (Figure 5C), the numbers of neurogenic cells identified were noticeably lower compared with the numbers of NPCs and NBs annotated by the authors (0 NPCs compared with 23 initially annotated; 3 NBs compared with 400 initially annotated) (Figures 5D and 5I). Again, when the same conserved-across-multiple-species markers were tested in another macaque dataset[62] (Figure 5E), a similar result was obtained (Figures 5F and 5I). These observations indicate that this holistic approach might not be suitable when the investigating processes that may have been under strong evolutionary pressure, possibly like AHN.[1] The implementation of appropriate positive computational controls using "proof-of-concept" datasets, (e.g., applying the same cross-species transfer learning in fetal hippocampal datasets[98]), becomes, evidently, a *sine qua non* for reliable benchmarking of cell type annotation.

A more cautious alternative to using markers that are conserved across multiple species would be to use marker genes derived from a single system that would resemble more the adult human situation, like, for instance, the non-human primate adult brain. To test this hypothesis, we assessed the expression of macaque NPC and NB markers from two independent snRNA-seq studies[24,62] in the aforementioned human dataset[24] (Figures 5G–5I; Table S1). When calculating the Seurat module score of NPC and NB macaque markers against the human dataset, we observed several nuclei showing an NPC/NB profile, mainly in excitatory neuronal clusters, possibly highlighting the existence of a transcriptionally diverse population of excitatory neurons (Figures 5G and 5H). When using the top 4–5 NPC and NB

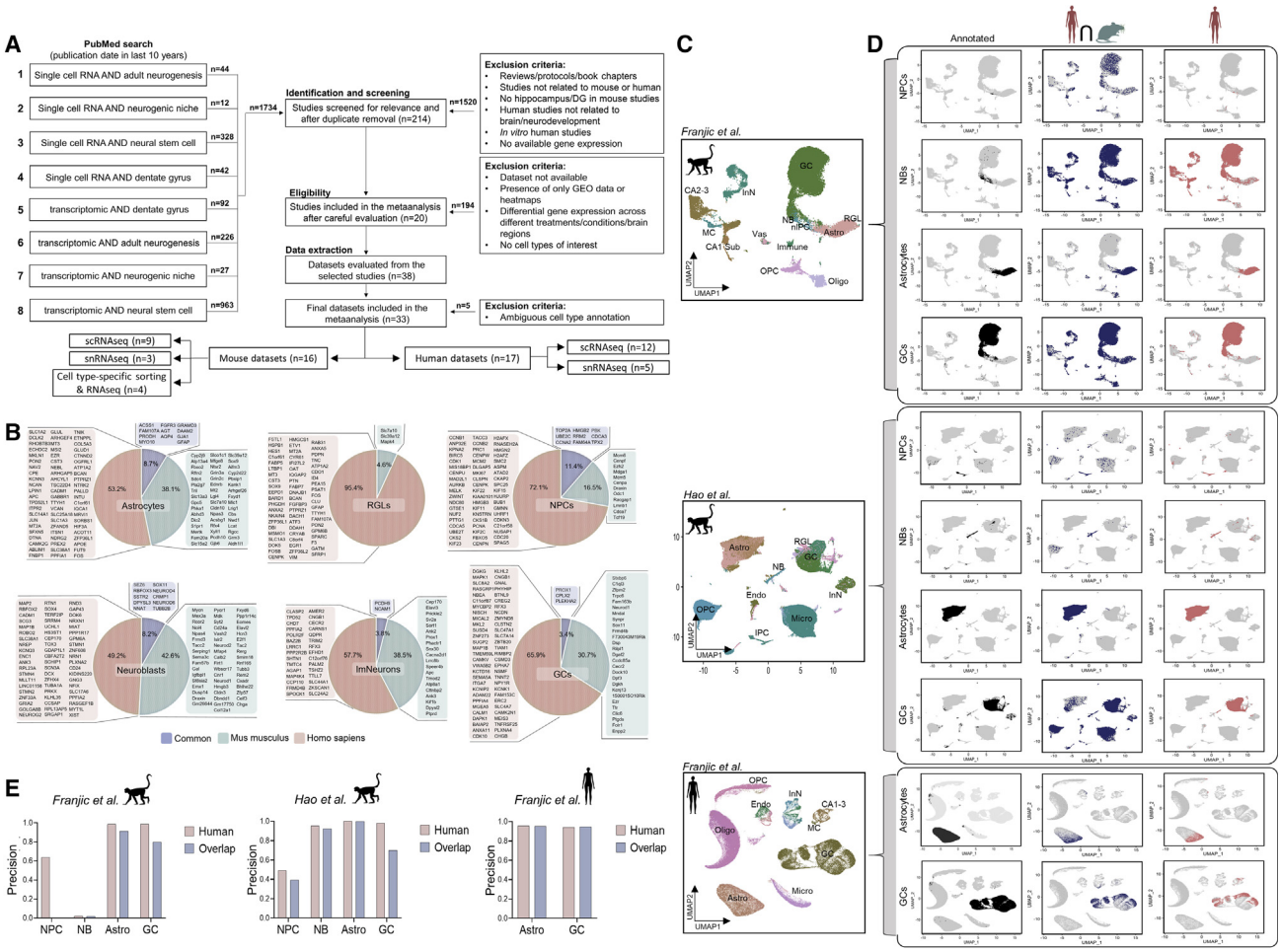


Figure 4. Meta-analysis of sc/snRNA-seq datasets to identify human-specific and cross-species conserved AHN markers

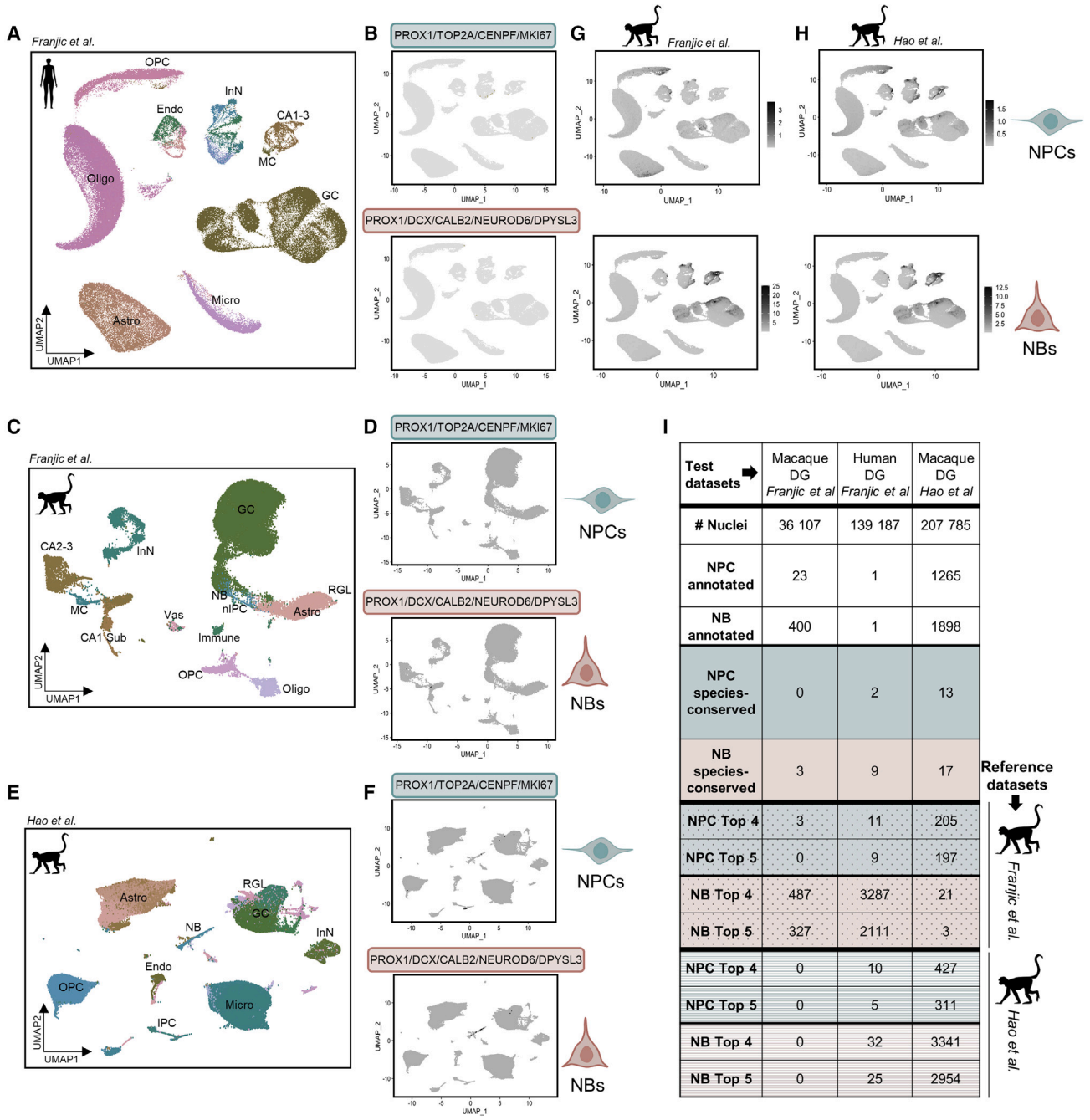
(A) Schematic of the meta-analysis strategy: a systematic literature search was conducted using PubMed (search terms indicated in the boxes on the left). Eight independent PubMed searches conducted in October 2021 yielded a total of 1,734 studies. These were subsequently screened on whether the studies fulfilled all the inclusion/exclusion criteria reported in the boxes on the right. A final selection of 33 datasets was included for further analysis (see Table S3 and STAR Methods).

(B) Human-specific (red), mouse-specific (green), and conserved across the two species (blue), markers for each cell type, obtained from the final selection of 33 datasets analyzed in (A).

(C) UMAP plots of data from Franjic et al.[24] representing 36,107 nuclei from macaque DG, Hao et al.[62] representing 207,785 cells from macaque DG, Franjic et al.[24] representing 139,187 nuclei from human DG. Cells/nuclei (dots) are labeled and colored by cluster membership labels from the original studies.

(D and E) Performance of the meta-analysis-identified markers in the macaque and human datasets per cell cluster: (D) UMAP plots as in (C), with highlighted cells representing cells with a module score higher than the calculated cutoff (implemented for binary identification of true and false positives, see STAR Methods and Table S2 for details) for the tested gene sets. See Figures S2A and S2B for the non-thresholded module scores. (E) Precision score of the different gene sets in the two macaque and human datasets calculated as described in STAR Methods and Table S2. Astro, astrocytes; Endo, endothelial cells; InN, inhibitory neuron; GC, granule cell; MC, mossy cell; Oligo, oligodendrocyte; OPC, oligodendrocyte precursor cell; NB, neuroblast, IPC, intermediate progenitor cell; Micro, microglia; Vas, vasculature cell; Human icon, human-specific markers (indicated in red in B–E); human∩mouse icon, intersection/overlap between human and mouse marker datasets (markers conserved across the two species, indicated in blue in B–E).

macaque markers to subset the nuclei from the human dataset (genes highlighted in yellow, Table S1), we again observed several nuclei showing an NPC/NB-like profile (Figure 5I). Among the nuclei expressing NPC markers, some were found co-clustering with oligodendrocyte progenitor cells (OPCs) (suggesting common progenitor-like transcriptional profiles, as previously reported[99,100]), whereas nuclei expressing NB markers were largely found within neuronal clusters (with most being GCs) (Figure S3). Of note, when assessing the top NB markers from the different macaque studies in the human dataset, the results greatly differ, with the Franjic et al.[24] top markers identifying an unexpectedly high number of nuclei (2,111). On the other hand, the NPC top markers showed higher congruence (top 5 NPC markers from Franjic et al.[24] identified 9 nuclei, top 5 NPC markers from Hao et al.[62] identified 5 nuclei, Figure 5I). When assessing the top macaque NPC/NB markers derived from one macaque study in the other macaque dataset, we noted important differences: for instance, the use of Hao et al.[62] top markers on the Franjic et al.[24] macaque



Test datasets	Macaque DG *Franjic et al*	Human DG *Franjic et al*	Macaque DG *Hao et al*
# Nuclei	36 107	139 187	207 785
NPC annotated	23	1	1265
NB annotated	400	1	1898
NPC species-conserved	0	2	13
NB species-conserved	3	9	17
NPC Top 4	3	11	205
NPC Top 5	0	9	197
NB Top 4	487	3287	21
NB Top 5	327	2111	3
NPC Top 4	0	10	427
NPC Top 5	0	5	311
NB Top 4	0	32	3341
NB Top 5	0	25	2954

Figure 5. Performance of shared-across-species versus macaque-only markers

(A) UMAP plot of data from Franjic et al.[24] representing 139,187 nuclei from human DG, colored by cluster membership labels derived from the original study.
(B) Same UMAP plot as in (A), where nuclei co-expressing at least 1 UMI of shared-across-species markers, as used by Franjic et al.,[24] are highlighted.
(C) UMAP plot of data from Franjic et al.[24] representing 36,107 nuclei from macaque DG, colored by cluster membership labels derived from the original study.
(D) Same UMAP plot as in (C), where nuclei co-expressing at least 1 UMI of shared-across-species markers, as used by Franjic et al.,[24] are highlighted.
(E) UMAP plot of data derived from Hao et al.[62] representing 207,785 cells from macaque DG, colored by cluster membership labels from the original study.
(F) Same UMAP plot as in (E), where cells co-expressing at least 1 UMI of shared-across-species markers, as used by Franjic et al.,[24] are highlighted.
(G and H) UMAP plots as in (A), colored by the Seurat module score calculated using NPC and NB macaque markers from Franjic et al.[24] (C) and Hao et al.[62] (E) computed using the FindMarkers() function in Seurat (Table S1).

(legend continued on next page)

dataset resulted in 0 NPC- or NB-like nuclei (Figure 5I). These differences across datasets from the same species highlight—once again—how different studies using distinct sampling and analysis pipelines might identify only partially overlapping marker genes (Figure S4). Evidently, as already mentioned before, the accuracy of code and pipeline reporting is crucial for dataset comparison and will critically interrelate with reproducibility.[101,102]

Of note, the studies comparing cellular populations in the brain of a series of mammalian species (e.g., human, macaque, marmoset, and mice) found that most cell types are homologous between species.[103–105] However, most cell type-specific genes were not conserved across different species, pointing toward the existence of recently evolved divergent genes that may have species-specific functions.[103] Hence, the significance of evolutionary conservation studies does not only lie in the identification of commonalities in core cellular functions but also (and perhaps even more so) in the characterization of non-conserved cell types or novel states that may be causally linked to species-specific traits and disorders.[103] Our observations suggest that such a perspective on evolutionary—molecular and perhaps also mechanistic—divergence may be of particular relevance when studying AHN in the human brain.

Finding the holy grail: On the quest for human cell type-specific neurogenic markers

Given the cross-species divergence, the use of rodent or non-human primate markers to infer the presence of neurogenesis in the human brain did not appear to perform optimally in our analyses. The advent of new human snRNA-seq datasets, in both fetal and in the past year also adult human DG, in which neurogenic signatures were found, re-ignited the debate over human AHN and at the same time provided unprecedented insights into the matter. We tested the expression of fetal hippocampal progenitor markers[98] and markers from the two recent studies that reported adult RGLs and ImNs[23] or only ImNs[22] (Table S1) in the Franjic et al.[24] dataset, in which no neurogenic populations were identified. For all the markers tested, we observed a subgroup of nuclei in the GC cluster displaying more immature profiles (green arrows, Figure 6A, upper panel). Interestingly, when testing fetal progenitor and adult RGL markers, an additional group of nuclei residing in the astrocytic cluster was depicted (red arrows, Figure 6A, upper panel). When using fetal progenitor markers, nuclei in the OPC cluster also showed expression for these markers, suggesting again possible transcriptome-level similarities between OPCs and NPCs. In addition, although all the fetal/adult human markers tested here seem to identify some nuclei in the appropriate UMAP space, these nuclei were not confined to one single cluster, as they were also found within other excitatory (ExN) and inhibitory neuron (InN) populations (Figure 6A, upper panel). With respect to the Zhou et al. markers, this observation could be explained by the fact that we did not integrate computed gene weights in our analysis, as the authors did.

Our re-analysis confirmed the machine learning-based observation by Zhou et al.[22] on the existence of ImN-like cells in the Franjic et al.[24] dataset. However, the re-analysis of the Franjic et al. dataset by Wang et al.[23] did not identify any neurogenic signatures, supporting the conclusion of the authors of the original study. To explore this controversy a bit further, we zoomed in and re-clustered the astrocytic population of Franjic et al.[24] (Figure 6A, lower panel). We were able to identify the two main subclusters annotated in the original study: one expressing high levels of *AQP4* and *GFAP* and the second one marked by high expression of *AQP4* and *CHRDL1*. Interestingly, when plotting on these nuclei the gene sets originally used by Wang et al.[23] as markers to differentiate between astrocytic and RGL populations in their dataset (Table S1), we noticed that the *AQP4/GFAP* subcluster in Franjic et al. overlapped with the astrocytic signature of Wang et al., whereas the *AQP4/CHRDL1* subpopulation of Franjic et al. showed high expression of the NSC marker genes derived from the Wang et al. dataset (Figures 6A and S4).

Next, we followed a similar approach to zoom in on the astrocytic/RGL population identified in the Wang et al. dataset (Figure 6B). Rather surprisingly, when plotting the expression of the marker gene sets corresponding to the *AQP4/GFAP* and *AQP4/CHRDL1* astrocytic subpopulations derived from the Franjic et al. dataset (Table S1), we noticed that they marked the astrocytic and RGL populations of Wang et al., respectively (Figure 6B), confirming the similar transcriptomic profile observed between Franjic *AQP4/GFAP* astrocytes and Wang astrocytes and that between Franjic *AQP4/CHRLD1* astrocytes and Wang RGLs (Figures 6A and S4). Although CHRDL1, a primarily astrocyte-secreted factor essential for stabilizing synapses, is widely used as a marker of protoplasmic astrocytes in the cortex,[106,107] a role in RGLs has not been reported yet. Our findings, apart from underlying once more the transcriptional affinity between astrocytes and RGLs, also suggest that populations with similar transcriptomic profiles (but disparate annotation) can be found in the Wang et al. and the Franjic et al. datasets. Whether (subsets of) these represent an astrocytic subpopulation with RGL-like features or *bona fide* RGLs will require further investigation.

Overall, the discrepancies across datasets (Figure S4) may indicate that distinct studies, using different sample stratification, sampling and lysis protocols, and computational pipelines, either identify diverse subpopulations and cell states involved in the neurogenic process or, alternatively, detect the same populations but pick up different marker genes.[108,109]

DISCUSSION

Moving forward: Seeing the forest

Here, we have taken a critical approach toward the promises and challenges that novel sc/snRNA-seq technologies bring to the study of human AHN. Profiling the presumed neurogenic niche in the adult human brain entails a whole series of hurdles, ranging

(I) Summary table reporting the number of nuclei/cells identified in the 3 different datasets (Franjic et al., human[24]; Franjic et al., macaque[24]; Hao et al., macaque[62]), when using the Franjic et al. shared-across-species markers ("NPC species-conserved" in green and NB species-conserved in pink) versus when using the top 4–5 markers from macaque-only studies ("NPC Top 4/5" in green and NB Top 4/5 in pink; dotted boxes, tested in Franjic et al.[24]; striped boxes, tested in Hao et al.[62] Top4/5 marker genes are indicated in Table S1). Astro, astrocytes; Endo, endothelia; InN, inhibitory neuron; GC, granule cell; MC, mossy cell; Oligo, oligodendrocyte; OPC, oligodendrocyte precursor cell; NB, neuroblast, IPC, intermediate progenitor cell; Micro, microglia; Vas, vasculature cell.

Please cite this article in press as: Tosoni et al., Mapping human adult hippocampal neurogenesis with single-cell transcriptomics: Reconciling controversy or fueling the debate?, Neuron (2023), https://doi.org/10.1016/j.neuron.2023.03.010

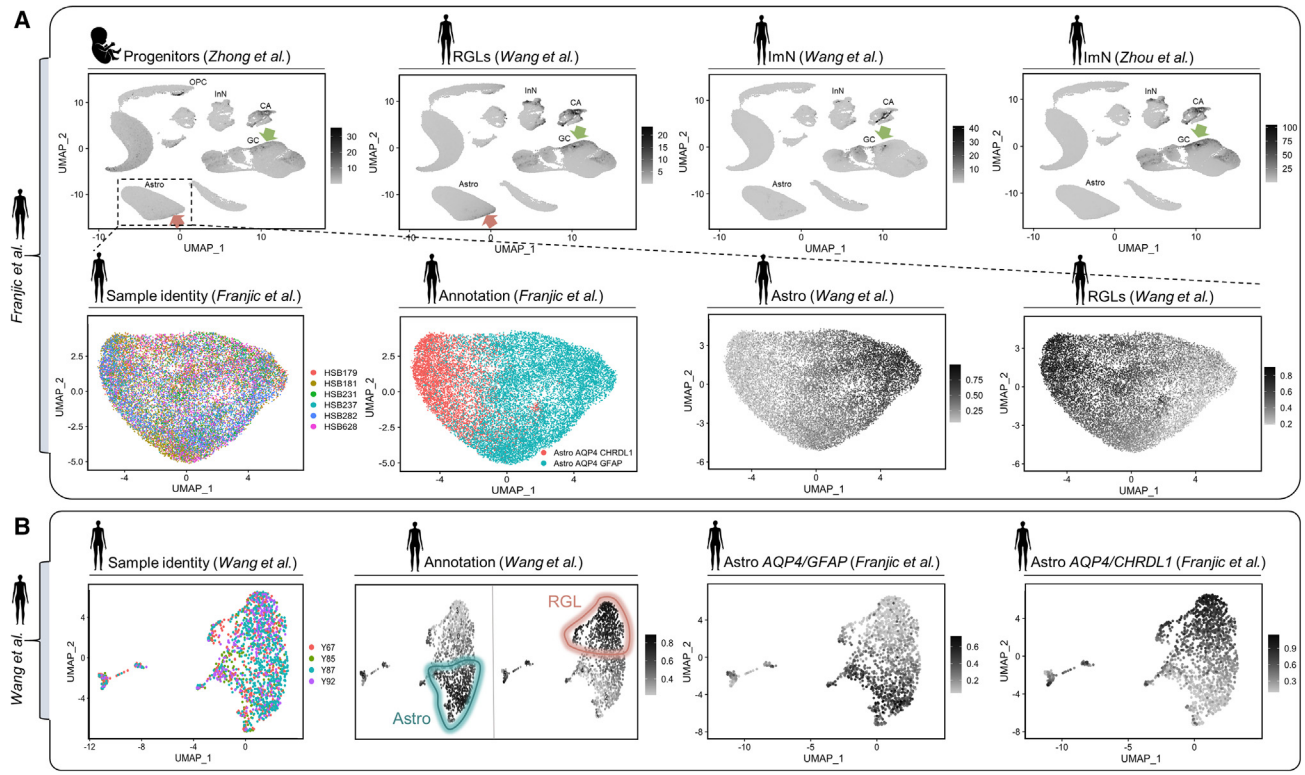


Figure 6. Performance of human-specific markers

(A) Upper panel: UMAP plots as in Figure 5A annotated according to the original study (Franjic et al.[24]), colored by the Seurat module score, calculated using fetal human progenitors markers from Zhong et al.,[98] adult human NSC markers from Wang et al.,[23] adult human ImN markers from Wang et al.,[23] and adult human ImN markers from Zhou et al.,[22] respectively. Lower panel: UMAP plot of subset astrocyte clusters from upper panel colored by sample identity, astrocytic subtype annotation derived from the original study, and module score computed using astrocytic and RGLs markers derived from Wang et al.[23]

(B) UMAP plot of subset astrocyte clusters from Wang et al.[23] colored by sample identity, annotation derived by module score plotting using the gene lists provided by the authors, module score using the computed markers derived from the astrocytic subtypes of Franjic et al.[24] (see Table S1 for marker genes used. Astro, astrocytes; Endo, endothelia; InN, inhibitory neuron; GC, granule cell; MC, mossy cell; Oligo, oligodendrocyte; OPC, oligodendrocyte precursor cell; NB, neuroblast, IPC, intermediate progenitor cell; Micro, microglia; RGL, radial glia-like; Vas, vasculature cell.

from—primarily—working with archived postmortem material to not having *a priori* knowledge of the cellular profiles we are looking for. We employed a comprehensive re-analysis of previously published sc/snRNA-seq datasets in mammalian adult DG, including non-human primates and humans. We argue that controversy as to whether human AHN exists often arises from methodological, conceptual, and even biological confounders that need to be consistently addressed across future sc/snRNA-seq studies. We believe that the fundamental principles the field agrees on might essentially outweigh what may separate human AHN believers from sceptics.

As thoroughly discussed here, sample processing, experimental design, and computational analysis may all impact the reliable identification of rare cellular populations, as is the case for the different neurogenic cell types in the adult hippocampus. Inter-individual variability, as already documented in non-human primates[24] and likely even more so in humans,[4,14,15,22] may further complicate the profiling of AHN in humans, asking for large-scale studies that are not always feasible due to the scarcity of high-quality postmortem brain tissue. In addition, co-existing pathologies that may directly interfere with the neurogenic process and/or induce neuroinflammation are common in the adult (particularly the aged and diseased) human brain and can negatively (and possibly differentially) affect the presence of neurogenic populations.[23]

Another major common concern that we and others[2–4] extensively addressed is whether markers previously identified in rodents are reliable enough to also label neurogenic populations in the adult human brain. A critical observation and potential confounder herein is that the robust AHN observed in mice is usually derived from studies in young (4–8 weeks old) animals, whereas in humans, AHN is often addressed in adult or aging populations.[110,111] The developmental time window of newborn neurons and the rate of their generation, which may be interrelated to cellular functionality, niche interactions, and relevance to disease, differ across species.[3,22,111] Therefore, assuming that such biologically significant differences required adaptations in gene expression programs in any putative neurogenic population is not a big leap of faith. Transcriptomic analyses in human and non-human primate cerebral organoids identified evolutionary and molecular constraints to orthologous gene expression across species, pointing toward the existence of human-specific

neurogenic traits.[112,113] These observations were recently confirmed by snRNA-seq studies in adult macaque and human DG.[22–24] Wang et al. reported a macaque-specific RGL signature not found in mice, whereas Zhou et al. observed only a minor overlap (15.5%) of orthologous genes shared between mouse and human ImNs. Hence, relying on rodent (or other species-conserved) markers to delineate the presumed neurogenic trajectory in human DG comes close to using a screwdriver to pound a nail: Non-specific markers will lead to poor cell type identification and vice versa. However, despite a broad consensus on this issue, it is still sometimes difficult to avoid circular argumentation[2,24]: markers X, Y, and Z are arguably not specific; however, the fact that they do not label a distinct cell cluster is used to suggest that there is no AHN in humans.

Independently of which neurogenic populations or cell states persist in the adult human brain, it is generally postulated that there may be an overall lower rate of AHN in the adult human DG, as reported in the studies suggesting decreasing numbers of RGLs,[14] proliferating precursors,[11] and ImNs[4,11,22] across the lifespan. These observations were initially supported by immunohistochemical studies[14–16] and more recently by similar findings in (sn)RNA-seq datasets.[22,23] Wang et al.[23] found that the proportion of active (proliferating) RGLs decreases in the adult-aged human brain, as also suggested by a sharp decline of the RGL mitotic score from 2 years of age onward. Similarly, Zhou et al.[22] reported that the proportion of ImNs among all GCs drastically declines with age (51.8% in the prenatal stage, 9.4% in infancy, 3.1%–7.5% from 4 years old on).

Evidently, to reliably identify such rare, complex, and previously uncharacterized populations and cell states and considering all the pitfalls discussed here, particularly sensitive and scalable novel methods are required.[103] One of the main issues that both supervised and unsupervised methods of cell clustering and annotation share is the requirement for some level of *a priori* knowledge to accurately assign cell identity.[114] In addition, many of these approaches are designed to identify distinct cellular populations, overlooking transitioning states with putatively mixed transcriptional features.[114] In fact, conventional clustering and finer partitioning pipelines were insufficient to separate ImNs that were found intermingled with mature GCs in one of the recent snRNA-seq datasets in human DG.[22] In this case, the authors followed a machine learning-based analytical approach that enabled the segregation of an ImN population. More specifically, the machine learning classifier employed generated a weighted list of more than 400 genes of positive or negative impact to precisely define immature GCs, going beyond any simplistic marker-based approach.[8]

Similarly, a stratified cross-validation machine learning approach was recently used in another study, allowing the identification of a particularly rare (120 cells of a total population of 40,000) disease-specific neuronal cell type in postmortem brain tissue of Parkinson's disease patients.[47] As longitudinal lineage tracing approaches are precluded in humans, novel computational tools that can determine lineage relationships based on automatic cell annotation from sc/snRNA-seq datasets will be of particular importance to identify and order cells along the neurogenic trajectory.[73,115–117] These alternative approaches will also be critical to differentiate *bona fide* progenitors and ImNs from developmentally born populations that may express similar transcriptional profiles. Importantly, for any new methodology, it is imperative to implement proper positive controls (e.g., validating machine learning models using labeled 'ground truth' datasets) for method validation. For example, in the study of Zhou et al.,[22] the same machine learning-based approach that enabled the identification of ImNs in the adult human brain was applied to a scRNA-seq dataset from mouse DG,[60] where it proved to be able to identify the same cell types on unsupervised clustering.

An emerging powerful new approach is the *ex vivo* fate mapping of newborn cells in primary human tissue that was recently also implemented for the tracing of newborn neurons in the surgically resected adult brain.[22,116] Moreover, although a challenging undertaking, data integration across studies would be instrumental to increase the sample size and achieve adequate power. Here, batch effects across datasets derived from different technologies and processing pipelines can introduce additional variation, which may confound the biological variance of interest and the downstream analysis. Of note, computational methods for batch correction are often unable to uncover rare cell types, further complicating matters.[118,119] In addition, overlaying with epigenetic[23] and spatial information to generate a more 'complete' atlas of the adult human neurogenic niche would be particularly useful, as it was recently shown in the embryonic and adult mouse brain.[89,120] Ultimately, functional validation of any newly identified cellular populations and states will be key to probing putative causal links to hippocampal functions, like memory and cognition.

As sc/snRNA-seq approaches start shedding light on the cellular and molecular complexities of the adult human DG, effort should be put into reconciling evidence from different studies to create a common taxonomy of cell types,[103,111] specify signatures of human progenitor populations, and even re-define AHN per se. Recent snRNA-seq evidence suggests that a substantial number of neurons with immature neuronal characteristics exist in the adult human hippocampus.[22,23] Although precise birth dating of these populations has not been possible yet, their presence (and even the presence of any putative developmentally born but still ImNs, as it has been suggested for rodents and non-human primates[121]), possibly due to protracted maturation,[3,22] could potentially provide a continuous reservoir of cells that may enhance hippocampal plasticity "on demand" by eventually maturing and integrating into the adult hippocampal network under specific conditions.[2,3,8,22] So would, in that case, the inability to detect amplifying progenitors imply the absence of AHN in the human brain? As Sorrells et al. argued in their recent commentary,[2] "adult neurogenesis refers to the birth of a neuron in the adult brain." According to this definition, the identification of ImNs in the adult human hippocampus supports the possibility that AHN exists in the human brain. However, it may be essential to move away from semantics: independently of how we will—decide to—label them, understanding what these cellular populations can do in the healthy and diseased adult human brain will be far more important.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.neuron.2023.03.010.

ACKNOWLEDGMENTS

We wish to thank Valentina La Monica for the initial collection of datasets; we are grateful to Hongjun Song, Yi Zhou, Yijing Su, Amber Penning, Oliver Polzer, and Hannah Walgrave for providing critical feedback on the manuscript. This work was supported by funding from Alzheimer Nederland to G.T. and E.S. (WE.03-2020-04) and used the Dutch national e-infrastructure with the support of the Dutch Research Council (Nederlandse Organisatie voor Wetenschappelijk Onderzoek, NWO) and the SURF Cooperative (EINF-1405 and EINF-4880). C.P.F. receives funding from Alzheimer Nederland (WE.03-2020-07), and P.J.L. from Alzheimer Nederland (WE.03-2018-01), ZonMW Memorabel (MODEM consortium), and the Center for Urban Mental Health.

AUTHOR CONTRIBUTIONS

Conceptualization, G.T. and E.S; methodology, G.T., D.A., J.B., W.M., and E.S.; software and formal analysis, G.T. and D.A.; data interpretation, G.T., D.A., J.B., W.M., C.P.F., P.J.L., and E.S.; writing, review & editing, G.T., D.A., J.B., W.M., C.P.F., P.J.L., and E.S.; supervision, E.S.; funding acquisition, E.S.

DECLARATION OF INTERESTS

J.B. and W.M. are employees of F. Hoffmann-La Roche AG.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: October 11, 2022
Revised: February 6, 2023
Accepted: March 8, 2023
Published: April 3, 2023

REFERENCES

1. Duque, A., Arellano, J.I., and Rakic, P. (2022). An assessment of the existence of adult neurogenesis in humans and value of its rodent models for neuropsychiatric diseases. Mol. Psychiatry *27*, 377–382. https://doi.org/10.1038/S41380-021-01314-8.
2. Sorrells, S.F., Paredes, M.F., Zhang, Z., Kang, G., Pastor-Alonso, O., Biagiotti, S., Page, C.E., Sandoval, K., Knox, A., Connolly, A., et al. (2021). Positive controls in adults and children support that very few, if any, new neurons are born in the adult human hippocampus. J. Neurosci. *41*, 2554–2565. https://doi.org/10.1523/JNEUROSCI.0676-20.2020.
3. Lucassen, P.J., Fitzsimons, C.P., Salta, E., and Maletic-Savatic, M. (2020). Adult neurogenesis, human after all (again): classic, optimized, and future approaches. Behav. Brain Res. *381*, 112458. https://doi.org/10.1016/j.bbr.2019.112458.
4. Moreno-Jiménez, E.P., Terreros-Roncal, J., Flor-García, M., Rábano, A., and Llorens-Martín, M. (2021). Evidences for adult hippocampal neurogenesis in humans. J. Neurosci. *41*, 2541–2553. https://doi.org/10.1523/JNEUROSCI.0675-20.2020.
5. Paredes, M.F., Sorrells, S.F., Cebrian-Silla, A., Sandoval, K., Qi, D., Kelley, K.W., James, D., Mayer, S., Chang, J., Auguste, K.I., et al. (2018). Does adult neurogenesis persist in the human hippocampus? Cell Stem Cell *23*, 780–781. https://doi.org/10.1016/j.stem.2018.11.006.
6. Liu, H.K. (2022). Human adult hippocampal neurogenesis is back, again? Cell Res. *32*, 793–794. https://doi.org/10.1038/s41422-022-00698-8.
7. Terreros-Roncal, J., Flor-García, M., Moreno-Jiménez, E.P., Rodríguez-Moreno, C.B., Márquez-Valadez, B., Gallardo-Caballero, M., Rábano, A., and Llorens-Martín, M. (2022). Methods to study adult hippocampal neurogenesis in humans and across the phylogeny. Published online October 18, 2022. Hippocampus. https://doi.org/10.1002/HIPO.23474.
8. Kempermann, G., Gage, F.H., Aigner, L., Song, H., Curtis, M.A., Thuret, S., Kuhn, H.G., Jessberger, S., Frankland, P.W., Cameron, H.A., et al. (2018). Human adult neurogenesis: evidence and remaining questions. Cell Stem Cell *23*, 25–30. https://doi.org/10.1016/J.STEM.2018.04.004.
9. Spalding, K.L., Bergmann, O., Alkass, K., Bernard, S., Salehpour, M., Huttner, H.B., Boström, E., Westerlund, I., Vial, C., Buchholz, B.A., et al. (2013). Dynamics of hippocampal neurogenesis in adult humans. Cell *153*, 1219–1227. https://doi.org/10.1016/J.CELL.2013.05.002.
10. Eriksson, P.S., Perfilieva, E., Björk-Eriksson, T., Alborn, A.M., Nordborg, C., Peterson, D.A., and Gage, F.H. (1998). Neurogenesis in the adult human hippocampus. Nat. Med. *4*, 1313–1317.
11. Sorrells, S.F., Paredes, M.F., Cebrian-Silla, A., Sandoval, K., Qi, D., Kelley, K.W., James, D., Mayer, S., Chang, J., Auguste, K.I., et al. (2018). Human hippocampal neurogenesis drops sharply in children to undetectable levels in adults. Nature *555*, 377–381. https://doi.org/10.1038/NATURE25975.
12. Cipriani, S., Ferrer, I., Aronica, E., Kovacs, G.G., Verney, C., Nardelli, J., Khung, S., Delezoide, A.L., Milenkovic, I., Rasika, S., et al. (2018). Hippocampal radial glial subtypes and their neurogenic potential in human fetuses and healthy and Alzheimer's disease adults. Cereb. Cortex *28*, 2458–2478. https://doi.org/10.1093/CERCOR/BHY096.
13. Dennis, C.V.v., Suh, L.S., Rodriguez, M.L., Kril, J.J., and Sutherland, G.T. (2016). Human adult neurogenesis across the ages: an immunohistochemical study. Neuropathol. Appl. Neurobiol. *42*, 621–638. https://doi.org/10.1111/NAN.12337.
14. Boldrini, M., Fulmore, C.A., Tartt, A.N., Simeon, L.R., Pavlova, I., Poposka, V., Rosoklija, G.B., Stankov, A., Arango, V., Dwork, A.J., et al. (2018). Human hippocampal neurogenesis persists throughout aging. Cell Stem Cell *22*. 589–599.e5. https://doi.org/10.1016/J.STEM.2018.03.015.
15. Tobin, M.K., Musaraca, K., Disouky, A., Shetti, A., Bheri, A., Honer, W.G., Kim, N., Dawe, R.J., Bennett, D.A., Arfanakis, K., and Lazarov, O. (2019). Human hippocampal neurogenesis persists in aged adults and Alzheimer's disease patients. Cell Stem Cell *24*. 974–982.e3. https://doi.org/10.1016/J.STEM.2019.05.003.
16. Moreno-Jiménez, E.P., Flor-García, M., Terreros-Roncal, J., Rábano, A., Cafini, F., Pallas-Bazarra, N., Ávila, J., and Llorens-Martín, M. (2019). Adult hippocampal neurogenesis is abundant in neurologically healthy subjects and drops sharply in patients with Alzheimer's disease. Nat. Med. *25*, 554–560. https://doi.org/10.1038/S41591-019-0375-9.
17. Lucassen, P.J., Toni, N., Kempermann, G., Frisen, J., Gage, F.H., and Swaab, D.F. (2019). Limits to human neurogenesis—really? Mol. Psychiatry *25*, 2207–2209. https://doi.org/10.1038/s41380-018-0337-5.
18. Kamath, T., Abdulraouf, A., Burris, S.J., Langlieb, J., Gazestani, V., Nadaf, N.M., Balderrama, K., Vanderburg, C., and Macosko, E.Z. (2022). Single-cell genomic profiling of human dopamine neurons identifies a population that selectively degenerates in Parkinson's disease. Nat. Neurosci. *25*, 588–595. https://doi.org/10.1038/S41593-022-01061-1.

19. Mathys, H., Davila-Velderrain, J., Peng, Z., Gao, F., Mohammadi, S., Young, J.Z., Menon, M., He, L., Abdurrob, F., Jiang, X., et al. (2019). Single-cell transcriptomic analysis of Alzheimer's disease. Nature *570*, 332–337. https://doi.org/10.1038/S41586-019-1195-2.

20. Tran, M.N., Maynard, K.R., Spangler, A., Huuki, L.A., Montgomery, K.D., Sadashivaiah, V., Tippani, M., Barry, B.K., Hancock, D.B., Hicks, S.C., et al. (2021). Single-nucleus transcriptome analysis reveals cell-type-specific molecular signatures across reward circuitry in the human brain. Neuron *109*, 3088–3103.e5. https://doi.org/10.1016/J.NEURON.2021.09.001.

21. Garcia, F.J., Sun, N., Lee, H., Godlewski, B., Mathys, H., Galani, K., Zhou, B., Jiang, X., Ng, A.P., Mantero, J., et al. (2022). Single-cell dissection of the human brain vasculature. Nature *603*, 893–899. https://doi.org/10.1038/S41586-022-04521-7.

22. Zhou, Y., Su, Y., Li, S., Kennedy, B.C., Zhang, D.Y., Bond, A.M., Sun, Y., Jacob, F., Lu, L., Hu, P., et al. (2022). Molecular landscapes of human hippocampal immature neurons across lifespan. Nature *607*, 527–533. https://doi.org/10.1038/s41586-022-04912-w.

23. Wang, W., Wang, M., Yang, M., Zeng, B., Qiu, W., Ma, Q., Jing, X., Zhang, Q., Wang, B., Yin, C., et al. (2022). Transcriptome dynamics of hippocampal neurogenesis in macaques across the lifespan and aged humans. Cell Res. *32*, 729–743. https://doi.org/10.1038/S41422-022-00678-Y.

24. Franjic, D., Skarica, M., Ma, S., Arellano, J.I., Tebbenkamp, A.T.N., Choi, J., Xu, C., Li, Q., Morozov, Y.M., Andrijevic, D., et al. (2022). Transcriptomic taxonomy and neurogenic trajectories of adult human, macaque, and pig hippocampal and entorhinal cells. Neuron *110*. 452–469.e14. https://doi.org/10.1016/J.NEURON.2021.10.036.

25. Mereu, E., Lafzi, A., Moutinho, C., Ziegenhain, C., McCarthy, D.J., Álvarez-Varela, A., Batlle, E., Sagar, G., Grün, D., Lau, J.K., et al. (2020). Benchmarking single-cell RNA-sequencing protocols for cell atlas projects. Nat. Biotechnol. *38*, 747–755. https://doi.org/10.1038/S41587-020-0469-4.

26. Svensson, V., da Veiga Beltrame, E., and Pachter, L. (2020). A curated database reveals trends in single-cell transcriptomics. Database (Oxford) *2020*, baaa073. https://doi.org/10.1093/DATABASE/BAAA073.

27. Kiselev, V.Y., Andrews, T.S., and Hemberg, M. (2019). Challenges in unsupervised clustering of single-cell RNA-seq data. Nat. Rev. Genet. *20*, 273–282. https://doi.org/10.1038/S41576-018-0088-9.

28. Svensson, V., Vento-Tormo, R., and Teichmann, S.A. (2018). Exponential scaling of single-cell RNA-seq in the past decade. Nat. Protoc. *13*, 599–604. https://doi.org/10.1038/NPROT.2017.149.

29. Habib, N., Avraham-Davidi, I., Basu, A., Burks, T., Shekhar, K., Hofree, M., Choudhury, S.R., Aguet, F., Gelfand, E., Ardlie, K., et al. (2017). Massively parallel single-nucleus RNA-seq with DroNc-seq. Nat. Methods *14*, 955–958. https://doi.org/10.1038/nmeth.4407.

30. Ayhan, F., Kulkarni, A., Berto, S., Sivaprakasam, K., Douglas, C., Lega, B.C., and Konopka, G. (2021). Resolving cellular and molecular diversity along the hippocampal anterior-to-posterior axis in humans. Neuron *109*. 2091–2105.e6. https://doi.org/10.1016/J.NEURON.2021.05.003.

31. Abrams, D., Kumar, P., Karuturi, R.K.M., and George, J. (2019). A computational method to aid the design and analysis of single cell RNA-seq experiments for cell type identification. BMC Bioinformatics *20*. https://doi.org/10.1186/S12859-019-2817-2.

32. Davis, A., Gao, R., and Navin, N.E. (2019). SCOPIT: sample size calculations for single-cell sequencing experiments. BMC Bioinformatics *20*, 566. https://doi.org/10.1186/S12859-019-3167-9/TABLES/2.

33. Verwer, R.W.H., Sluiter, A.A., Balesar, R.A., Baayen, J.C., Noske, D.P., Dirven, C.M.F., Wouda, J., van Dam, A.M., Lucassen, P.J., and Swaab, D.F. (2007). Mature astrocytes in the adult human neocortex express the early neuronal marker doublecortin. Brain *130*, 3321–3335. https://doi.org/10.1093/BRAIN/AWM264.

34. Schmid, K.T., Höllbacher, B., Cruceanu, C., Böttcher, A., Lickert, H., Binder, E.B., Theis, F.J., and Heinig, M. (2021). scPower accelerates and optimizes the design of multi-sample single cell transcriptomic studies. Nat. Commun. *12*, 6625. https://doi.org/10.1038/s41467-021-26779-7.

35. Zhang, X., Xu, C., and Yosef, N. (2019). Simulating multiple faceted variability in single cell RNA sequencing. Nat. Commun. *10*, 2611. https://doi.org/10.1038/s41467-019-10500-w.

36. Christian, K.M., Song, H., and Ming, G.L. (2014). Functions and dysfunctions of adult hippocampal neurogenesis. Annu. Rev. Neurosci. *37*, 243–262. https://doi.org/10.1146/annurev-neuro-071013-014134.

37. Maharjan, R., Diaz Bustamante, L.D., Ghattas, K.N., Ilyas, S., Al-Refai, R., and Khan, S. (2020). Role of lifestyle in neuroplasticity and neurogenesis in an aging brain. Cureus *12*, e10639. https://doi.org/10.7759/CUREUS.10639.

38. Sung, P.S., Lin, P.Y., Liu, C.H., Su, H.C., and Tsai, K.J. (2020). Neuroinflammation and neurogenesis in Alzheimer's disease and potential therapeutic approaches. Int. J. Mol. Sci. *21*. https://doi.org/10.3390/IJMS21030701.

39. Ekdahl, C.T., Claasen, J.H., Bonde, S., Kokaia, Z., and Lindvall, O. (2003). Inflammation is detrimental for neurogenesis in adult brain. Proc. Natl. Acad. Sci. USA *100*, 13632–13637. https://doi.org/10.1073/PNAS.2234031100.

40. van Wageningen, T.A., Gerrits, E., Palacin I Bonson, S., Huitinga, I., Eggen, B.J.L., and van Dam, A.M. (2021). Exploring reported genes of microglia RNA-sequencing data: uses and considerations. Glia *69*, 2933–2946. https://doi.org/10.1002/GLIA.24078.

41. Leng, K., Li, E., Eser, R., Piergies, A., Sit, R., Tan, M., Neff, N., Li, S.H., Rodriguez, R.D., Suemoto, C.K., et al. (2021). Molecular characterization of selectively vulnerable neurons in Alzheimer's disease. Nat. Neurosci. *24*, 276–287. https://doi.org/10.1038/S41593-020-00764-7.

42. Ding, J., Adiconis, X., Simmons, S.K., Kowalczyk, M.S., Hession, C.C., Marjanovic, N.D., Hughes, T.K., Wadsworth, M.H., Burks, T., Nguyen, L.T., et al. (2020). Systematic comparison of single-cell and single-nucleus RNA-sequencing methods. Nat. Biotechnol. *38*, 737–746. https://doi.org/10.1038/S41587-020-0465-8.

43. Bakken, T.E., Hodge, R.D., Miller, J.A., Yao, Z., Nguyen, T.N., Aevermann, B., Barkan, E., Bertagnolli, D., Casper, T., Dee, N., et al. (2018). Single-nucleus and single-cell transcriptomes compared in matched cortical cell types. PLoS One *13*, e0209648. https://doi.org/10.1371/JOURNAL.PONE.0209648.

44. Wang, X., Allen, M., Li, S., Quicksall, Z.S., Patel, T.A., Carnwath, T.P., Reddy, J.S., Carrasquillo, M.M., Lincoln, S.J., Nguyen, T.T., et al. (2020). Deciphering cellular transcriptional alterations in Alzheimer's disease brains. Mol. Neurodegener. *15*, 38. https://doi.org/10.1186/S13024-020-00392-6.

45. Lake, B.B., Codeluppi, S., Yung, Y.C., Gao, D., Chun, J., Kharchenko, P.V., Linnarsson, S., and Zhang, K. (2017). A comparative strategy for single-nucleus and single-cell transcriptomes confirms accuracy in predicted cell-type expression from nuclear RNA. Sci. Rep. *7*, 6031. https://doi.org/10.1038/S41598-017-04426-W.

46. Olah, M., Menon, V., Habib, N., Taga, M.F., Ma, Y., Yung, C.J., Cimpean, M., Khairallah, A., Coronas-Samano, G., Sankowski, R., et al. (2020). Single cell RNA sequencing of human microglia uncovers a subset associated with Alzheimer's disease. Nat. Commun. *11*, 6129. https://doi.org/10.1038/s41467-020-19737-2.

47. Smajić, S., Prada-Medina, C.A., Landoulsi, Z., Ghelfi, J., Delcambre, S., Dietrich, C., Jarazo, J., Henck, J., Balachandran, S., Pachchek, S., et al. (2022). Single-cell sequencing of human midbrain reveals glial activation and a Parkinson-specific neuronal state. Brain *145*, 964–978. https://doi.org/10.1093/BRAIN/AWAB446.

48. Gupta, A., Shamsi, F., Altemose, N., Dorlhiac, G.F., Cypess, A.M., White, A.P., Yosef, N., Patti, M.E., Tseng, Y.H., and Streets, A. (2022). Characterization of transcript enrichment and detection bias in single-nucleus RNA-seq for mapping of distinct human adipocyte lineages. Genome Res. *32*, 242–257. https://doi.org/10.1101/GR.275509.121.

49. Zaghlool, A., Niazi, A., Björklund, Å.K., Westholm, J.O., Ameur, A., and Feuk, L. (2021). Characterization of the nuclear and cytosolic transcriptomes in human brain tissue reveals new insights into the subcellular distribution of RNA transcripts. Sci. Rep. *11*, 4076. https://doi.org/10.1038/S41598-021-83541-1.

50. Thrupp, N., Sala Frigerio, C., Wolfs, L., Skene, N.G., Fattorelli, N., Poovathingal, S., Fourne, Y., Matthews, P.M., Theys, T., Mancuso, R., et al. (2020). Single-nucleus RNA-seq is not suitable for detection of microglial activation genes in humans. Cell Rep. *32*, 108189. https://doi.org/10.1016/J.CELREP.2020.108189.

51. Boekhoorn, K., Joels, M., and Lucassen, P.J. (2006). Increased proliferation reflects glial and vascular-associated changes, but not neurogenesis in the presenile Alzheimer hippocampus. Neurobiol. Dis. *24*, 1–14. https://doi.org/10.1016/J.NBD.2006.04.017.

52. Li, J.Z., Vawter, M.P., Walsh, D.M., Tomita, H., Evans, S.J., Choudary, P.V.v., Lopez, J.F., Avelar, A., Shokoohi, V., Chung, T., et al. (2004). Systematic changes in gene expression in postmortem human brains associated with tissue pH and terminal medical conditions. Hum. Mol. Genet. *13*, 609–616. https://doi.org/10.1093/HMG/DDH065.

53. Dachet, F., Brown, J.B., Valyi-Nagy, T., Narayan, K.D., Serafini, A., Boley, N., Gingeras, T.R., Celniker, S.E., Mohapatra, G., and Loeb, J.A. (2021). Selective time-dependent changes in activity and cell-specific gene expression in human postmortem brain. Sci. Rep. *11*, 6078. https://doi.org/10.1038/s41598-021-85801-6.

54. Madissoon, E., Wilbrey-Clark, A., Miragaia, R.J., Saeb-Parsy, K., Mahbubani, K.T., Georgakopoulos, N., Harding, P., Polanski, K., Huang, N., Nowicki-Osuch, K., et al. (2019). scRNA-seq assessment of the human lung, spleen, and esophagus tissue stability after cold preservation. Genome Biol. *21*, 1. https://doi.org/10.1186/s13059-019-1906-x.

55. Ferreira, P.G., Muñoz-Aguirre, M., Reverter, F., Sá Godinho, C.P., Sousa, A., Amadoz, A., Sodaei, R., Hidalgo, M.R., Pervouchine, D., Carbonell-Caballero, J., et al. (2018). The effects of death and post-mortem cold ischemia on human tissue transcriptomes. Nat. Commun. *9*, 490. https://doi.org/10.1038/s41467-017-02772-x.

56. Pozhitkov, A.E., Neme, R., Domazet-Lošo, T., Leroux, B.G., Soni, S., Tautz, D., and Noble, P.A. (2017). Tracing the dynamics of gene transcripts after organismal death. Open Biol. *7*, 160267. https://doi.org/10.1098/RSOB.160267.

57. van den Brink, S.C., Sage, F., Vértesy, Á., Spanjaard, B., Peterson-Maduro, J., Baron, C.S., Robin, C., and van Oudenaarden, A. (2017). Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. Nat. Methods *14*, 935–936. https://doi.org/10.1038/nmeth.4437.

58. Marsh, S.E., Walker, A.J., Kamath, T., Dissing-Olesen, L., Hammond, T.R., de Soysa, T.Y., Young, A.M.H., Murphy, S., Abdulraouf, A., Nadaf, N., et al. (2022). Dissection of artifactual and confounding glial signatures by single-cell sequencing of mouse and human brain. Nat. Neurosci. *25*, 306–316. https://doi.org/10.1038/s41593-022-01022-8.

59. Ayhan, F., Douglas, C., Lega, B.C., and Konopka, G. (2021). Nuclei isolation from surgically resected human hippocampus. Star Protoc. *2*, 100844. https://doi.org/10.1016/J.XPRO.2021.100844.

60. Hochgerner, H., Zeisel, A., Lönnerberg, P., and Linnarsson, S. (2018). Conserved properties of dentate gyrus neurogenesis across postnatal development revealed by single-cell RNA sequencing. Nat. Neurosci. *21*, 290–299. https://doi.org/10.1038/s41593-017-0056-2.

61. Pancheva, A., Wheadon, H., Rogers, S., and Otto, T.D. (2022). Using topic modeling to detect cellular crosstalk in scRNA-seq. PLoS Comput. Biol. *18*, e1009975. https://doi.org/10.1371/JOURNAL.PCBI.1009975.

62. Hao, Z.Z., Wei, J.R., Xiao, D., Liu, R., Xu, N., Tang, L., Huang, M., Shen, Y., Xing, C., Huang, W., et al. (2022). Single-cell transcriptomics of adult macaque hippocampus reveals neural precursor cell populations. Nat. Neurosci. *25*, 805–817. https://doi.org/10.1038/s41593-022-01073-x.

63. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*. 3573–3587.e29. https://doi.org/10.1016/J.CELL.2021.04.048.

64. Luecken, M.D., and Theis, F.J. (2019). Current best practices in single-cell RNA-seq analysis: a tutorial. Mol. Syst. Biol. *15*, e8746. https://doi.org/10.15252/MSB.20188746.

65. Qiu, P. (2020). Embracing the dropouts in single-cell RNA-seq analysis. Nat. Commun. *11*, 1169. https://doi.org/10.1038/s41467-020-14976-9.

66. DeMeo, B., and Berger, B. (2022). Recovering single-cell heterogeneity through information-based dimensionality reduction. https://doi.org/10.1101/2021.01.19.427303.

67. Duò, A., Robinson, M.D., and Soneson, C. (2018). A systematic performance evaluation of clustering methods for single-cell RNA-seq data. F1000Res *7*, 1141. https://doi.org/10.12688/F1000RESEARCH.15666.3.

68. Fang, J., Chan, C., Owzar, K., Wang, L., Qin, D., Li, Q.-J., and Xie, J. (2022). Clustering deviation index (CDI): a robust and accurate unsupervised measure for evaluating scRNA-seq data clustering. https://doi.org/10.1101/2022.01.03.474840.

69. Kiselev, V.Y., Kirschner, K., Schaub, M.T., Andrews, T., Yiu, A., Chandra, T., Natarajan, K.N., Reik, W., Barahona, M., Green, A.R., and Hemberg, M. (2017). SC3: consensus clustering of single-cell RNA-seq data. Nat. Methods *14*, 483–486. https://doi.org/10.1038/NMETH.4236.

70. Žurauskiene, J., and Yau, C. (2016). pcaReduce: hierarchical clustering of single cell transcriptional profiles. BMC Bioinformatics *17*, 140. https://doi.org/10.1186/S12859-016-0984-Y.

71. Freytag, S., Tian, L., Lönnstedt, I., Ng, M., and Bahlo, M. (2018). Comparison of clustering tools in R for medium-sized 10x Genomics single-cell RNA-sequencing data. F1000Res *7*, 1297. https://doi.org/10.12688/F1000RESEARCH.15809.2.

72. Jindal, A., Gupta, P., Jayadeva, and SenGupta, D. (2018). Discovery of rare cells from voluminous single cell expression data. Nat. Commun. *9*, 4719. https://doi.org/10.1038/S41467-018-07234-6.

73. Bej, S., Galow, A.M., David, R., Wolfien, M., and Wolkenhauer, O. (2021). Automated annotation of rare-cell types from single-cell RNA-sequencing data through synthetic oversampling. BMC Bioinformatics *22*, 557. https://doi.org/10.1186/S12859-021-04469-X.

74. Wegmann, R., Neri, M., Schuierer, S., Bilican, B., Hartkopf, H., Nigsch, F., Mapa, F., Waldt, A., Cuttat, R., Salick, M.R., et al. (2019). CellSIUS provides sensitive and specific detection of rare cell populations from complex single-cell RNA-seq data. Genome Biol. *20*, 142. https://doi.org/10.1186/S13059-019-1739-7.

75. Grün, D., Muraro, M.J., Boisset, J.C., Wiebrands, K., Lyubimova, A., Dharmadhikari, G., van den Born, M., van Es, J., Jansen, E., Clevers, H., et al. (2016). De novo prediction of stem cell identity using single-cell transcriptome data. Cell Stem Cell *19*, 266–277. https://doi.org/10.1016/J.STEM.2016.05.010.

76. Dong, R., and Yuan, G.C. (2020). GiniClust3: a fast and memory-efficient tool for rare cell type identification. BMC Bioinformatics *21*, 158. https://doi.org/10.1186/S12859-020-3482-1/FIGURES/1.

77. Imoto, Y., Segi-Nishida, E., Suzuki, H., and Kobayashi, K. (2017). Rapid and stable changes in maturation-related phenotypes of the adult hippocampal neurons by electroconvulsive treatment. Mol. Brain *10*, 8. https://doi.org/10.1186/S13041-017-0288-9.

78. Zocher, S., Overall, R.W., Berdugo-Vega, G., Rund, N., Karasinsky, A., Adusumilli, V.S., Steinhauer, C., Scheibenstock, S., Händler, K., Schultze, J.L., et al. (2021). De novo DNA methylation controls neuronal maturation during adult hippocampal neurogenesis. EMBO J. *40*, e107100. https://doi.org/10.15252/EMBJ.2020107100.

79. Shin, J., Berg, D.A., Zhu, Y., Shin, J.Y., Song, J., Bonaguidi, M.A., Enikolopov, G., Nauen, D.W., Christian, K.M., Ming, G.L., and Song, H. (2015). Single-cell RNA-seq with waterfall reveals molecular cascades

underlying adult neurogenesis. Cell Stem Cell *17*, 360–372. https://doi.org/10.1016/J.STEM.2015.07.013.

80. Dulken, B.W., and Brunet, A. (2018). Same path, different beginnings. Nat. Neurosci. *21*, 159–160. https://doi.org/10.1038/s41593-017-0063-3.
81. Artegiani, B., Lyubimova, A., Muraro, M., van Es, J.H., van Oudenaarden, A., and Clevers, H. (2017). A single-cell RNA sequencing study reveals cellular and molecular dynamics of the hippocampal neurogenic niche. Cell Rep. *21*, 3271–3284. https://doi.org/10.1016/J.CELREP.2017.11.050.
82. Penning, A., Tosoni, G., Abiega, O., Bielefeld, P., Gasperini, C., de Pietri Tonelli, D., Fitzsimons, C.P., and Salta, E. (2021). Adult neural stem cell regulation by small non-coding RNAs: physiological significance and pathological implications. Front. Cell. Neurosci. *15*, 781434. https://doi.org/10.3389/FNCEL.2021.781434.
83. Bonaguidi, M.A., Wheeler, M.A., Shapiro, J.S., Stadel, R.P., Sun, G.J., Ming, G.L., and Song, H. (2011). In vivo clonal analysis reveals self-renewing and multipotent adult neural stem cell characteristics. Cell *145*, 1142–1155. https://doi.org/10.1016/J.CELL.2011.05.024.
84. Dulken, B.W., Buckley, M.T., Navarro Negredo, P., Saligrama, N., Cayrol, R., Leeman, D.S., George, B.M., Boutet, S.C., Hebestreit, K., Pluvinage, J.V., et al. (2019). Single-cell analysis reveals T cell infiltration in old neurogenic niches. Nature *571*, 205–210. https://doi.org/10.1038/S41586-019-1362-5.
85. Arellano, J.I., Morozov, Y.M., Micali, N., and Rakic, P. (2021). Radial glial cells: new views on old questions. Neurochem. Res. *46*, 2512–2524. https://doi.org/10.1007/S11064-021-03296-Z.
86. Cebrian-Silla, A., Nascimento, M.A., Redmond, S.A., Mansky, B., Wu, D., Obernier, K., Romero Rodriguez, R.R., Gonzalez-Granero, S., García-Verdugo, J.M., Lim, D.A., et al. (2021). Single-cell analysis of the ventricular-subventricular zone reveals signatures of dorsal and ventral adult neurogenesis. eLife *10*, e67436. https://doi.org/10.7554/eLife.67436.
87. Zywitza, V., Misios, A., Bunatyan, L., Willnow, T.E., and Rajewsky, N. (2018). Single-cell transcriptomics characterizes cell types in the subventricular zone and uncovers molecular defects impairing adult neurogenesis. Cell Rep. *25*. 2457–2469.e8. https://doi.org/10.1016/J.CELREP.2018.11.003.
88. Zhang, H., Li, J., Ren, J., Sun, S., Ma, S., Zhang, W., Yu, Y., Cai, Y., Yan, K., Li, W., et al. (2021). Single-nucleus transcriptomic landscape of primate hippocampal aging. Protein Cell *12*, 695–716. https://doi.org/10.1007/S13238-021-00852-9.
89. Kremer, L.P., Cerrizuela, S., Eid, M., Shukairi, A., Ellinger, T., Straub, J., Dehler, S., Korkmaz, A., Weichenhan, D., Plass, C., et al. (2022). Single-cell triple-omics uncovers DNA methylation as key feature of stemness in the healthy and ischemic adult brain. https://doi.org/10.1101/2022.07.13.499860.
90. Kohler, S.J., Williams, N.I., Stanton, G.B., Cameron, J.L., and Greenough, W.T. (2011). Maturation time of new granule cells in the dentate gyrus of adult macaque monkeys exceeds six months. Proc. Natl. Acad. Sci. USA *108*, 10326–10331. https://doi.org/10.1073/PNAS.1017099108.
91. Ngwenya, L.B., Heyworth, N.C., Shwe, Y., Moore, T.L., and Rosene, D.L. (2015). Age-related changes in dentate gyrus cell numbers, neurogenesis, and associations with cognitive impairments in the rhesus monkey. Front. Syst. Neurosci. *9*, 102. https://doi.org/10.3389/FNSYS.2015.00102.
92. Merz, K., and Lie, D.C. (2013). Evidence that doublecortin is dispensable for the development of adult born neurons in mice. PLoS One *8*, e62693. https://doi.org/10.1371/JOURNAL.PONE.0062693.
93. Ming, G.L. li, and Song, H. (2011). Adult neurogenesis in the mammalian brain: significant answers and significant questions. Neuron *70*, 687–702. https://doi.org/10.1016/J.NEURON.2011.05.001.
94. Knoth, R., Singec, I., Ditter, M., Pantazis, G., Capetian, P., Meyer, R.P., Horvat, V., Volk, B., and Kempermann, G. (2010). Murine features of neurogenesis in the human hippocampus across the lifespan from 0 to 100 years. PLoS One *5*, e8809. https://doi.org/10.1371/JOURNAL.PONE.0008809.
95. Mayer, C., Hafemeister, C., Bandler, R.C., Machold, R., Batista Brito, R., Jaglin, X., Allaway, K., Butler, A., Fishell, G., and Satija, R. (2018). Developmental diversification of cortical inhibitory interneurons. Nature *555*, 457–462. https://doi.org/10.1038/NATURE25999.
96. Nowakowski, T.J., Bhaduri, A., Pollen, A.A., Alvarado, B., Mostajo-Radji, M.A., di Lullo, E., Haeussler, M., Sandoval-Espinosa, C., Liu, S.J., Velmeshev, D., et al. (2017). Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. Science *358*, 1318–1323. https://doi.org/10.1126/SCIENCE.AAP8809.
97. Telley, L., and Jabaudon, D. (2018). A mixed model of neuronal diversity. Nature *555*, 452–454. https://doi.org/10.1038/D41586-018-02539-4.
98. Zhong, S., Ding, W., Sun, L., Lu, Y., Dong, H., Fan, X., Liu, Z., Chen, R., Zhang, S., Ma, Q., et al. (2020). Decoding the development of the human hippocampus. Nature *577*, 531–536. https://doi.org/10.1038/s41586-019-1917-5.
99. Huang, W., Bhaduri, A., Velmeshev, D., Wang, S., Wang, L., Rottkamp, C.A., Alvarez-Buylla, A., Rowitch, D.H., and Kriegstein, A.R. (2020). Origins and proliferative states of human oligodendrocyte precursor cells. Cell *182*. 594–608.e11. https://doi.org/10.1016/J.CELL.2020.06.027.
100. Neftel, C., Laffy, J., Filbin, M.G., Hara, T., Shore, M.E., Rahme, G.J., Richman, A.R., Silverbush, D., Shaw, M.L., Hebert, C.M., et al. (2019). An integrative model of cellular states, plasticity, and genetics for glioblastoma. Cell *178*. 835–849.e21. https://doi.org/10.1016/J.CELL.2019.06.024.
101. Squair, J., Courtine, G., and Skinnider, M. (2020). Enabling reproducible re-analysis of single-cell data. Genome Biol. *22*, 215. https://doi.org/10.5281/ZENODO.4772064.
102. Füllgrabe, A., George, N., Green, M., Nejad, P., Aronow, B., Fexova, S.K., Fischer, C., Freeberg, M.A., Huerta, L., Morrison, N., et al. (2020). Guidelines for reporting single-cell RNA-seq experiments. Nat. Biotechnol. *38*, 1384–1386. https://doi.org/10.1038/s41587-020-00744-z.
103. Vinsland, E., and Linnarsson, S. (2022). Single-cell RNA-sequencing of mammalian brain development: insights and future directions. Development *149*, dev200180. https://doi.org/10.1242/dev.200180.
104. Krienen, F.M., Goldman, M., Zhang, Q., del Rosario, C.H.R., Florio, M., Machold, R., Saunders, A., Levandowski, K., Zaniewski, H., et al. (2020). Innovations present in the primate interneuron repertoire. Nature *586*, 262–269. https://doi.org/10.1038/s41586-020-2781-z.
105. Bakken, T.E., Jorstad, N.L., Hu, Q., Lake, B.B., Tian, W., Kalmbach, B.E., Crow, M., Hodge, R.D., Krienen, F.M., Sorensen, S.A., et al. (2021). Comparative cellular analysis of motor cortex in human, marmoset and mouse. Nature *598*, 111–119. https://doi.org/10.1038/s41586-021-03465-8.
106. Lanjakornsiripan, D., Pior, B.J., Kawaguchi, D., Furutachi, S., Tahara, T., Katsuyama, Y., Suzuki, Y., Fukazawa, Y., and Gotoh, Y. (2018). Layer-specific morphological and molecular differences in neocortical astrocytes and their dependence on neuronal layers. Nat. Commun. *9*, 1623. https://doi.org/10.1038/S41467-018-03940-3.
107. Batiuk, M.Y., Martirosyan, A., Wahis, J., de Vin, F., Marneffe, C., Kusserow, C., Koeppen, J., Viana, J.F., Oliveira, J.F., Voet, T., et al. (2020). Identification of region-specific astrocyte subtypes at single cell resolution. Nat. Commun. *11*, 1220. https://doi.org/10.1038/S41467-019-14198-8.
108. Squair, J.W., Gautier, M., Kathe, C., Anderson, M.A., James, N.D., Hutson, T.H., Hudelle, R., Qaiser, T., Matson, K.J.E., Barraud, Q., et al. (2021). Confronting false discoveries in single-cell differential expression. Nat. Commun. *12*, 5692. https://doi.org/10.1038/s41467-021-25960-2.
109. Wang, T., Li, B., Nelson, C.E., and Nabavi, S. (2019). Comparative analysis of differential gene expression analysis tools for single-cell RNA

sequencing data. BMC Bioinformatics *20*, 40. https://doi.org/10.1186/S12859-019-2599-6/TABLES/7.

110. Bond, A.M., Ming, G.L., and Song, H. (2022). What is the relationship between hippocampal neurogenesis across different stages of the lifespan? Front. Neurosci. *16*, 891713. https://doi.org/10.3389/FNINS.2022.891713/XML/NLM.

111. Snyder, J.S. (2019). Recalibrating the relevance of adult neurogenesis. Trends Neurosci. *42*, 164–178. https://doi.org/10.1016/J.TINS.2018.12.001.

112. Pollen, A.A., Bhaduri, A., Andrews, M.G., Nowakowski, T.J., Meyerson, O.S., Mostajo-Radji, M.A., di Lullo, E., Alvarado, B., Bedolli, M., Dougherty, M.L., et al. (2019). Establishing cerebral organoids as models of human-specific brain evolution. Cell *176*. 743–756.e17. https://doi.org/10.1016/J.CELL.2019.01.017.

113. Zhu, Y., Sousa, A.M.M., Gao, T., Skarica, M., Li, M., Santpere, G., Esteller-Cucala, P., Juan, D., Ferrández-Peral, L., Gulden, F.O., et al. (2018). Spatiotemporal transcriptomic divergence across human and macaque brain development. Science *362*, eaat8077. https://doi.org/10.1126/science.aat8077.

114. Kong, W., Fu, Y.C., Holloway, E.M., Garipler, G., Yang, X., Mazzoni, E.O., and Morris, S.A. (2022). Capybara: A computational tool to measure cell identity and fate transitions. Cell Stem Cell *29*, 635–649.e11. https://doi.org/10.1016/J.STEM.2022.03.001.

115. Xie, B., Jiang, Q., Mora, A., and Li, X. (2021). Automatic cell type identification methods for single-cell RNA sequencing. Comput. Struct. Biotechnol. J. *19*, 5874–5887. https://doi.org/10.1016/J.CSBJ.2021.10.027.

116. Allen, D.E., Donohue, K.C., Cadwell, C.R., Shin, D., Keefe, M.G., Sohal, V.S., and Nowakowski, T.J. (2022). Fate mapping of neural stem cell niches reveals distinct origins of human cortical astrocytes. Science *376*, 1441–1446. https://doi.org/10.1126/SCIENCE.ABM5224.

117. Wagner, D.E., and Klein, A.M. (2020). Lineage tracing meets single-cell omics: opportunities and challenges. Nat. Rev. Genet. *21*, 410–427. https://doi.org/10.1038/s41576-020-0223-2.

118. Luecken, M.D., Büttner, M., Chaichoompu, K., Danese, A., Interlandi, M., Mueller, M.F., Strobl, D.C., Zappia, L., Dugas, M., Colomé-Tatché, M., and Theis, F.J. (2021). Benchmarking atlas-level data integration in single-cell genomics. Nat. Methods *19*, 41–50. https://doi.org/10.1038/s41592-021-01336-8.

119. Guo, T., Chen, Y., Shi, M., Li, X., and Zhang, M.Q. (2022). Integration of single cell data by disentangled representation learning. Nucleic Acids Res. *50*, e8. https://doi.org/10.1093/NAR/GKAB978.

120. Li, C., Virgilio, M.C., Collins, K.L., and Welch, J.D. (2022). Multi-omic single-cell velocity models epigenome-transcriptome interactions and improves cell fate prediction. Nat. Biotechnol. *41*, 387–398. https://doi.org/10.1038/S41587-022-01476-Y.

121. la Rosa, C., Ghibaudi, M., and Bonfanti, L. (2019). Newly generated and non-newly generated "immature" neurons in the mammalian brain: a possible reservoir of young cells to prevent brain aging and disease? J. Clin. Med. *8*, 685. https://doi.org/10.3390/JCM8050685.

122. Becht, E., McInnes, L., Healy, J., Dutertre, C.A., Kwok, I.W.H., Ng, L.G., Ginhoux, F., and Newell, E.W. (2018). Dimensionality reduction for visualizing single-cell data using UMAP. Nat. Biotechnol. *37*, 38–47. https://doi.org/10.1038/NBT.4314.

123. Marsh, S., Salmon, M., and Hoffman, P. (2023). samuel-marsh/scCustomize: version 1.1.1. https://doi.org/10.5281/ZENODO.7534950.

124. Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., and Vilo, J. (2019). g:profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Res. *47*, W191–W198. https://doi.org/10.1093/NAR/GKZ369.

125. Chatzi, C., Zhang, Y., Shen, R., Westbrook, G.L., and Goodman, R.H. (2016). Transcriptional profiling of newly generated dentate granule cells using TU tagging reveals pattern shifts in gene expression during circuit integration. eNeuro *3*, 589–596. https://doi.org/10.1523/ENEURO.0024-16.2016.

126. Berg, D.A., Su, Y., Jimenez-Cyrus, D., Patel, A., Huang, N., Morizet, D., Lee, S., Shah, R., Ringeling, F.R., Jain, R., et al. (2019). A common embryonic origin of stem cells drives developmental and adult neurogenesis. Cell *177*. 654–668.e15. https://doi.org/10.1016/J.CELL.2019.02.010.

127. Jonas, P., and Lisman, J. (2014). Structure, function, and plasticity of hippocampal dentate gyrus microcircuits. Front. Neural Circuits *8*, 107. https://doi.org/10.3389/fncir.2014.00107.

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

129. Marsh, S., Salmon, M., and Hoffman, P. (2023). samuel-marsh/scCustomize: version 1.1.1. https://doi.org/10.5281/ZENODO.7534950.

Please cite this article in press as: Tosoni et al., Mapping human adult hippocampal neurogenesis with single-cell transcriptomics: Reconciling controversy or fueling the debate?, Neuron (2023), https://doi.org/10.1016/j.neuron.2023.03.010

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
scRNA-seq data of mouse dentate gyrus	Hochgerner et al.[60]	GEO: GSE104323
snRNA-seq data of adult macaque dentate gyrus	Franjic et al.[24]	GEO: GSE186538
scRNA-seq data of adult macaque dentate gyrus	Hao et al.[62]	ArrayExpress: E-MTAB-10225
snRNA-seq data of adult human dentate gyrus	Franjic et al.[24]	GEO: GSE186538
snRNA-seq data of adult human hippocampus	Wang et al.[23]	GEO: GSE163737
Software and algorithms		
R version 4.1.3 (2022-03-10)	R project	https://www.r-project.org/
Seurat (4.1.1)	Hao et al.[63]	https://satijalab.org/seurat/index.html
UMAP	Becht et al.[122]	https://github.com/lmcinnes/umap
scCustomize	Marsh[123]	https://samuel-marsh.github.io/scCustomize/index.html
g:Profiler version e106_eg53_p16_65fcd97	Raudvere et al.[124]	https://biit.cs.ut.ee/gprofiler/orth
Howmanycells	Satija lab	https://satijalab.org/howmanycells/
Cell Ranger (6.0.2)	10X Genomics	https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/6.0/what-is-cell-ranger

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Evgenia Salta (e.salta@nin.knaw.nl).

Materials availability

This study did not generate new unique reagents.

Data and code availability

This paper analyzes existing, publicly available data. Accession numbers for the datasets are listed in the key resources table. This paper does not report original code. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

METHOD DETAILS

Meta-analysis

A systematic literature search was performed to identify relevant single-cell transcriptomic studies related to developmental neurogenesis, adult neurogenesis, hippocampus, and neural stem cells. Eight independent searches were conducted in PubMed, including only studies published in the past ten years, ranging from 2011 to 2021, which led to a total of 1734 hits (Figure 4A).

After duplicate article removal, an initial filtering step was performed. For inclusion, the studies had to fulfill the following criteria: mouse or human research articles, DG-related studies in mice, brain- or neurodevelopment-related studies in humans, presence of gene expression data. Subsequently, exclusion criteria were applied to assess the eligibility of the articles, which were filtered based on data availability. Thus, papers not including publicly available datasets, datasets consisting of solely raw data (e.g., GEO available datasets), or differentially expression data comparing cell types across different experimental conditions, were excluded. In total, 33 datasets were finally selected for data extraction.

The extracted gene sets were organized into species-specific and cell type-specific groups, and subsequently underwent an initial processing. For both species, gene expression data corresponding to the following cell types were collected: astrocytes, mature granule cells (GCs), immature neurons (ImN), neuroblasts (NBs), neural progenitor cells (NPCs), and radial glia-like cells (RGLs).

Subsequently, data selection was carried out depending on the number of genes provided per cell type. More specifically, only the top 1000 expressed genes, according to fold change ranking, were selected from gene lists with more than 1000 genes, whereas the entire gene list was selected from gene sets providing less than 1000 genes. Gene lists with less than 100 genes were excluded. For mouse genes: dentate gyrus-specific gene sets were available and used for astrocytes, ImN, NBs, NPCs and RGLs. For GCs, non-dentate gyrus-only datasets were also used. In addition, the GC dataset from Chatzi et al.[125] and the NPC dataset from Berg et al.[126] were excluded due to marker-based selection of cellular subtypes. For human genes: hippocampus-specific datasets were available and used for astrocytes and GCs. For immature GCs, NBs, NPCs and RGLs, both hippocampal and cortical gene sets were used. Marker-based, cellular subtype-specific datasets were also excluded. See Table S3 for final datasets included and calculated numbers and percentages of overlapping genes.

Power analysis for snRNA-seq datasets

The numbers of DCX^+ - cells were obtained from each of the indicated datasets according to information provided by the authors in the respective manuscripts (Figure 1C). The total number of granule cells (GCs) was estimated at 15 million.[127] The proportion of GCs to the total number of dentate gyrus cells was estimated at 23%.[22,24] Proportions of DCX^+ - cells to the total number of GCs or of dentate gyrus cells were calculated using the aforementioned estimates. Based on the calculated proportions, the expected number of DCX^+ - cells per 30,000 GCs was calculated for each dataset, assuming linearity (i.e., expected number when employing immunolabeling in tissue sections). The value of 30,000 GCs was selected as the number of GCs sequenced in the most powered published snRNA-seq studies to date.[22,24] Probability estimation for snRNA-seq was performed using How Many Cells | Satija Lab online software (https://satijalab.org/howmanycells/) calculating how many cells need to be sampled to detect at least n cells of each type. For a given cell type, the probability of seeing at least n cells in a sample of size k follows the cumulative distribution function of a negative binomial NBcdf (k; n, p), with p being the relative abundance. As n, we used the calculated expected number of DCX^+ - cells per 30,000 GCs (in immunolabeling settings) for each dataset. Rare cell clusters were defined as those consisting of 300 cells or less. The number of rare cell types (m) was inferred from Franjic et al. and estimated to be 20. As minimum fraction of rarest cell type in each dataset (p) we employed the proportion of DCX^+ - cells per total dentate gyrus cells.

Re-analysis of previously published datasets

Data extraction and preprocessing

To impute the sequencing depth of the Wang et al. study,[23] Fastq files were retrieved from NCBI GEO (see Table S1 for accession number). Reads were aligned to the human reference genome GRCh38 with the Cell Ranger software v6.0.2 using *–include-introns* option, and subsequently a cell-by-gene count matrix was generated.

Data extraction and processing

Preprocessed UMI count matrices from published datasets (summarized in Table S1) were retrieved from the respective repositories (key resources table) and processed independently, using the same criteria used in the original publications, into the R package Seurat (4.1.1). Briefly, preprocessed UMI count matrices were normalized using NormalizeData function in the R package Seurat with the scaling factor equal to 10,000.[128] To facilitate the visualization of all cells with similar transcriptomic profile in the two-dimensional space the top 2,000 highly variable genes were obtained by the Seurat function FindVariableFeatures via the default variance stabilizing process. Dimension reduction analysis in the space of those variable features was performed with the functions RunPCA and RunUMAP.[122] Clustering analysis was performed using the FindNeighbors and FindClusters functions. Cell annotation information was retrieved from the available metadata of the respective datasets[24,60,62] or from marker genes provided by the authors[23] and used for all visualizations in the two-dimensional space and downstream analyses.

Median number of genes

Median number of genes (nFeature_RNA) per cell type was calculated using Median_Stats function[129] and visualized in the two-dimensional space and as violin plots.

Marker identification

When not explicitly provided in the original study, markers for the different cell types of interest were identified using FindMarkers function of Seurat with an adjusted p-value threshold of 0.05. The top 4 and top 5 expressed markers were selected according to the $\log_2$ fold change value. The full lists of markers used per cell type in Figures 3, 5, 6, S3, and S4 can be found in Table S1, where yellow-highlighted genes represent the top 4 and top 5 markers employed in Figures 5I and S3.

Assessing cell type specificity of marker genes

The cell type-specific expression of the identified or provided markers was assessed using two different approaches. To assess the expression specificity of the top 4 or 5 markers, cells co-expressing at least 1 UMI of the top markers were subset using the subset() function of Seurat (results reported in Figure 5I and visualized in the two-dimensional space (Figure S3)).

To assess the cell type enrichment of the comprehensive lists of markers in different datasets, the AddModuleScore() function of Seurat was employed, and the results were visualized in the two-dimensional space (Figures 3, 5, 6, and S2). The performance/cell type specificity of the comprehensive lists of markers in different datasets was assessed by calculating precision, recall/sensitivity and F_1-score on the imputed module scores, as described below.

Precision, Sensitivity and F_1-score calculation

To assess the ability of a given gene set to correctly identify (only) the cell type of interest, precision, recall/sensitivity and F_1-scores were calculated according to the following formulas:

$$Precision(PPV) = \frac{TP}{TP + FP}$$

$$Sensitivity(TPR) = \frac{TP}{P} = \frac{TP}{TP + FN}$$

$$F_1 = 2 \times \frac{PPV \times TPR}{PPV + TPR}$$

TP = True Positives
FP = False Positives
FN = False Negatives
P = Condition Positive
PPV = Positive Predicted Value
TPR = True Positive Rate

As threshold/cutoff for binary identification of cells as 'positive' or 'negative', the mean of the module score in the cluster of interest was used. True positives (TP) and false positives (FP) correspond to the number of cells reporting a module score higher than the cutoff within and outside the cluster of interest, respectively. False negatives (FN) and true negatives (TN) correspond to the number of cells with a module score lower than the cutoff within and outside the cluster of interest, respectively.

Orthologue gene conversion

To compare genes obtained from different species (mouse, macaque, human), genes were converted into the respective orthologues. Orthologue gene conversion was performed using g:Profiler[124] g:Orth, g:Profiler version *e106_eg53_p16_65fcd97*, database updated on 18/05/2022).

Inflammation score

The inflammation score of astrocytes was calculated by applying the AddModuleScore() function of Seurat using the gene set of 'positive regulation of inflammatory response' (GO:0050729). The score plotted in Figure S1 represents the average inflammation score of each astrocytic cluster ± SEM.

Statistical analysis

All statistical analyses and resulting p-values are indicated in the text or figure legends and were performed with the R language for statistical computing (v4.1.3; https://www.r-project.org/) or using GraphPad Prism software.