

## Robust transcriptional profiling and identification of differentially expressed genes with low input RNA sequencing of adult hippocampal neural stem and progenitor populations

Jiyeon K. Denninger<sup>1</sup>, Logan A. Walker<sup>2a</sup>, Xi Chen<sup>3</sup>, Altan Turkoglu<sup>2b</sup>, Alex Pan<sup>3</sup>, Zoe Tapp<sup>4</sup>, Sakthi Senthilvelan<sup>1</sup>, Raina Rindani<sup>1</sup>, Olga Kokiko-Cochran<sup>4,5</sup>, Ralf Bundschuh<sup>2,6,7</sup>, Pearly Yan<sup>3,6</sup>, Elizabeth D. Kirby<sup>1,5\*</sup>

<sup>1</sup>The Ohio State University, Department of Psychology, College of Arts and Sciences, Columbus, OH, USA <sup>2</sup>The Ohio State University, Department of Physics, College of Arts and Sciences, Columbus, OH, USA <sup>3</sup>The Ohio State University Comprehensive Cancer Center, College of Medicine, Columbus, OH, USA <sup>4</sup>The Ohio State University, Department of Neuroscience, Institute for Behavioral Medicine Research, Columbus, OH, USA <sup>5</sup>The Ohio State University, Chronic Brain Injury Program, Columbus, OH, USA <sup>6</sup>The Ohio State University, Division of Hematology, Department of Internal Medicine, College of Medicine, Columbus, OH, USA <sup>7</sup>The Ohio State University, Department of Chemistry and Biochemistry, College of Arts and Sciences, Columbus, OH, USA

<sup>a</sup>Current affiliation: University of Michigan, Program in Biophysics, Ann Arbor, MI, USA

<sup>b</sup>Current affiliation: OSU College of Engineering, Department of Electrical and Computer Engineering, Columbus, OH, USA

\* Correspondence: Elizabeth Kirby, Kirby.224@osu.edu

**Keywords:** adult neural stem cell, hippocampus, differential gene expression, single cell RNA sequencing, bulk RNA sequencing

### Abstract

Multipotent neural stem cells (NSCs) are found in several isolated niches of the adult mammalian brain where they have unique potential to assist in tissue repair. Modern transcriptomics offer high-throughput methods for identifying disease or injury associated gene expression signatures in endogenous adult NSCs, but they require adaptation to accommodate the rarity of NSCs. Bulk RNA sequencing (RNAseq) of NSCs requires pooling several mice, which impedes application to labor-intensive injury models. Alternatively, single cell RNAseq can profile hundreds to thousands of cells from a single mouse and is increasingly used to study NSCs. The consequences of the low RNA input from a single NSC on downstream identification of differentially expressed genes (DEGs) remains largely unexplored. Here, to clarify the role that low RNA input plays in NSC DEG identification, we directly compared DEGs in an oxidative stress model of cultured NSCs by bulk and single cell sequencing. While both methods yielded DEGs that were replicable, single cell sequencing DEGs derived from genes with higher relative transcript counts compared to all detected genes and exhibited smaller fold changes than DEGs identified by bulk RNAseq. The loss of high fold-change DEGs in the single cell platform presents an important limitation for identifying disease-relevant genes. To facilitate identification of such genes, we determined an RNA-input threshold that enables transcriptional profiling of NSCs comparable to standard bulk sequencing and used it to establish a workflow for *in vivo* profiling of endogenous NSCs. We then applied this workflow to identify DEGs after lateral fluid percussion injury, a labor-intensive animal model of traumatic brain injury. Our work suggests that single cell RNA sequencing may underestimate the diversity of pathologic DEGs but population level transcriptomic analysis can be adapted to capture more of these DEGs with similar efficacy and diversity as standard bulk sequencing. Together, our data and workflow will be useful for investigators interested in understanding and manipulating adult hippocampal NSC responses to various stimuli.

## Introduction

The subgranular zone (SGZ) of the hippocampal dentate gyrus (DG) is a unique neurogenic niche in the adult mammalian brain (Denoth-Lippuner and Jessberger, 2021; Vicedomini et al., 2020). Neural stem cells (NSCs) in the SGZ give rise to functional new neurons throughout adulthood that contribute to hippocampal memory and affect regulation and could also be a source for endogenous tissue repair after injury or disease (McAvoy and Sahay, 2017; Miller and Sahay, 2019). Over the past decade, transcriptional analysis using high-throughput RNA sequencing (RNAseq) technology has dramatically expanded knowledge of NSC molecular characteristics. For example, studies using prospectively identified stem and progenitor populations have uncovered previously unknown cell lineage relationships (Baser et al., 2019; Berg et al., 2019; Dulken et al., 2017; Llorens-Bobadilla et al., 2015). Other studies using more unbiased approaches have revealed regional and even cell-specific transcriptional differences or transcriptional changes during development, adult neurogenesis, or aging (Artegiani et al., 2017; Dulken et al., 2019; Hochgerner et al., 2018; Mizrak et al., 2019; Shin et al., 2015; Yuzwa et al., 2017; Zywitzka et al., 2018). As studies of the NSC transcriptome expand, researchers are faced with an increasing variety of options for how to accomplish transcriptional profiling of this small, but critical, cell population.

Current major challenges to transcriptional profiling of NSCs include their relative sparsity *in vivo* and their transcriptional similarity to astrocytes. Both of these challenges have made bulk RNAseq of prospectively isolated NSCs a less attractive approach as it requires large cell number input and prospective isolation of the desired population. Instead, single cell RNAseq (scRNAseq) has emerged as the preferred technique to begin overcoming the above barriers. This approach uses the very small amounts of RNA present in single cells to generate thousands of individual cell transcriptomes with massively paralleled sequencing. scRNAseq studies of the adult mouse SVZ and SGZ have identified rare subpopulations of cells, as well as dynamic changes in gene expression at different developmental stages, maturation states, and regional locations (Artegiani et al., 2017; Llorens-Bobadilla et al., 2015; Shin et al., 2015). Several studies have adapted analytical methods from standard bulk sequencing to accommodate the technical challenges presented by using such low input to profile cells from other lineages (Law et al., 2014; Love et al., 2014; Robinson and Oshlack, 2010). In addition, many pioneering studies in other cell types have also developed novel processing and analysis tools specifically designed to facilitate detection of cell heterogeneity or chronological mapping of developmental trajectories using scRNAseq transcriptomes (Finak et al., 2015; Qiu et al., 2017; Shin et al., 2015; Trapnell et al., 2013).

While these recent studies show that scRNAseq is a powerful approach to characterize differences between individual cells, it is not yet clear how effective it is for uncovering population-level changes in gene expression. Identification of differentially expressed genes (DEGs) induced by variables like injury or gene expression manipulation is critical to understanding the mechanisms underlying NSC function in both disease models and in healthy brains. It seems logical that the low input of scRNAseq would affect DEG discovery compared to standard bulk RNAseq, where DEG analysis in transcriptomics was first developed (Arzalluz-Luque et al., 2017; Bhargava et al., 2015). However, beyond a few recent studies directly comparing the efficacy of DEG identification by various statistical methods on single cell datasets, there is little information available directly comparing the quality of DEG analysis in scRNAseq versus bulk RNAseq (Mou et al., 2020; Wang et al., 2019; Ziegenhain et al., 2017). Here, we directly compare scRNAseq

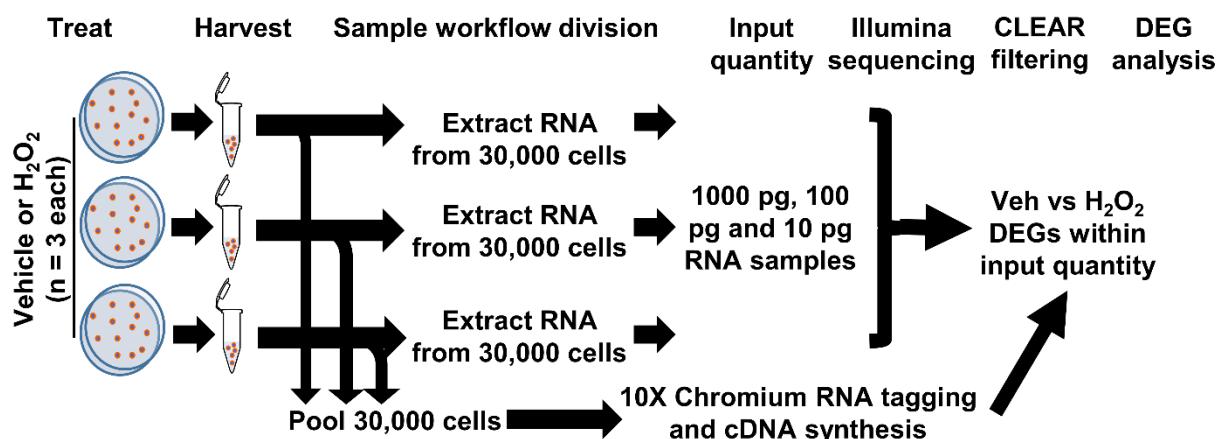
with bulk level RNAseq of cultured NSCs in a model of oxidative stress in vitro to evaluate DEG identification across sequencing approaches. We surprisingly found little overlap in DEGs identified by scRNAseq and bulk RNAseq, despite using the same source samples. While subsequent experiments showed that DEGs from both approaches were replicable, we found that scRNAseq identified DEGs among genes that show a more moderate fold change and high relative transcript count when compared to the bulk RNAseq approach. Because many studies of DEGs would specifically benefit from identification of higher fold change transcripts which are more moderately expressed, we adapted and validated a limiting cell (lc) RNAseq approach for sequencing DG NSCs isolated from individual adult mice with similar reliability as more bulk RNAseq-like approaches. We further demonstrate the utility of this method by applying it to transcriptome profiling of NSCs and their intermediate progenitor cell (IPC) progeny from single adult mouse hippocampi after a lateral fluid percussion injury (LFPI) model of traumatic brain injury (TBI).

## Results

### DEGs identified by scRNAseq versus bulk population level RNAseq are different but accurate.

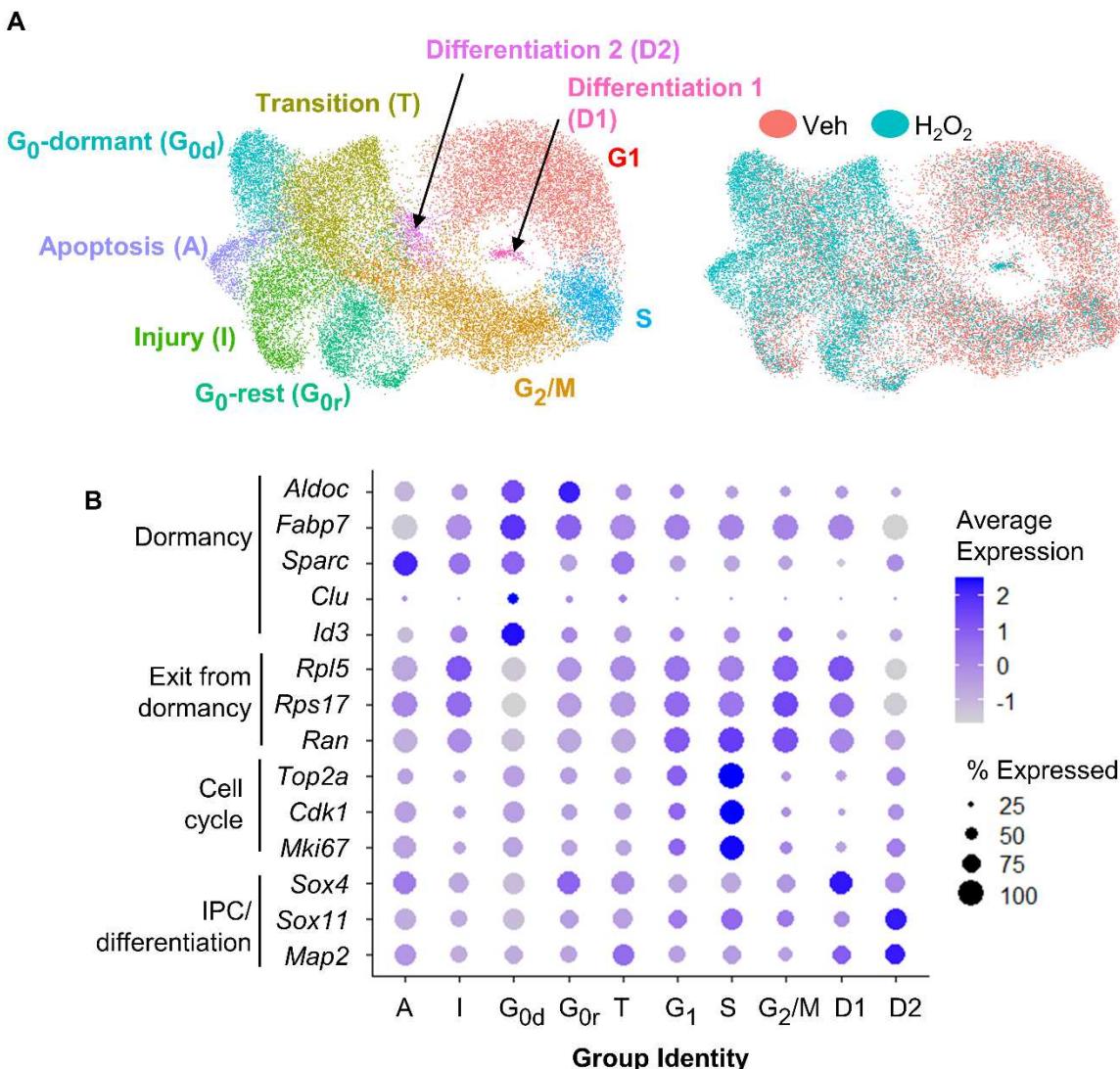
To compare detection of differential gene expression by scRNAseq and bulk RNAseq of cultured cells, we used an in vitro model of oxidative stress with adult mouse DG-derived NSCs. We treated cultured NSCs derived from male and female adult mouse DGs with H<sub>2</sub>O<sub>2</sub> to induce oxidative stress or with vehicle and then harvested cell pellets from 3 biological replicates (1 male, 2 female) per treatment. Harvested cells were then subdivided into two processing streams: RNA extraction for bulk RNAseq of pooled cells or direct RNAseq of individual cells (scRNAseq) using a 10X Chromium platform (**Figure 1**).

In scRNAseq, a total of 26,916 cells were captured and sequenced. UMAP analysis of both H<sub>2</sub>O<sub>2</sub>-treated and control samples revealed 10 different subpopulations characterized by gene expression profiles linked to GO terms consistent with specific stages of the cell cycle, phases of quiescence, differentiation, or response to injury (**Figure 2A-B** and **Supplementary**



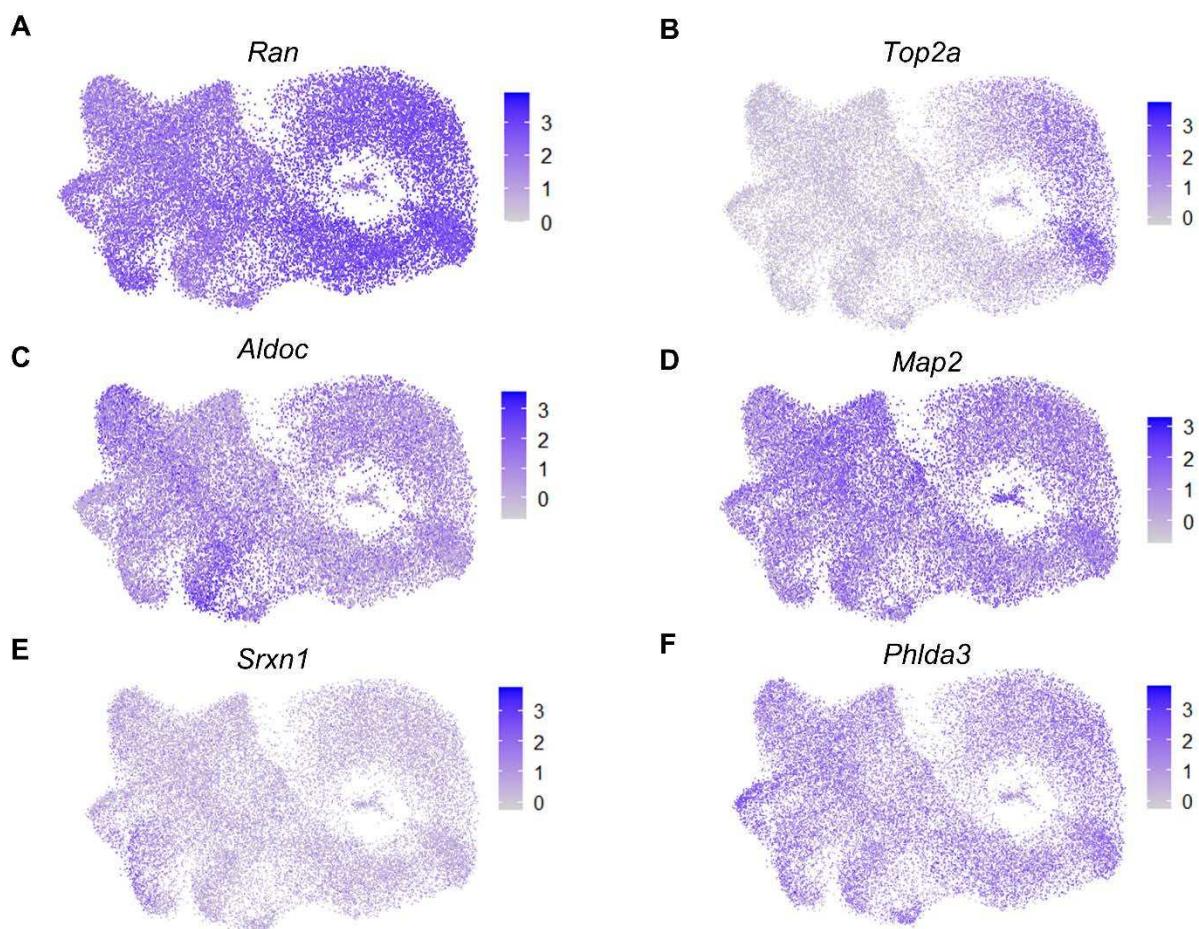
**FIGURE 1.** Workflow for cultured NSCs in scRNAseq and bulk RNAseq. NSCs were treated with H<sub>2</sub>O<sub>2</sub> or vehicle in triplicate. 10,000 cells from each biologic replicate were pooled for a total of 30,000 cells applied to the 10X Chromium scRNAseq platform and subsequent DEG analysis. RNA was extracted from 30,000 cells of each biologic replicate and 10 pg, 100 pg, and 1000 pg (1 ng) were used in RNAseq with subsequent CLEAR filtering and DEG analysis.

**Table 1**). The majority of cells were in G<sub>1</sub> (22%), G<sub>2</sub>/M (21%), or S (6.5%) phase of the cell cycle. 17.5% of cells were in one of two detected quiescent phases (G<sub>0d</sub> and G<sub>0r</sub>). 16% of the cells appeared to be in an intermediate state that we characterize as transitioning to/from the cell cycle (T) and the remaining 17% of cells were differentiating (3%, D1 and D2) or responding to injury (14%, I and A). Cells from both treatment groups were present in all clusters (**Figure 2A**). However, vehicle treated cells were mostly concentrated in the cycling clusters while H<sub>2</sub>O<sub>2</sub> treatment resulted in a notable shift away from cycling clusters to quiescent and apoptotic/injured clusters.



**FIGURE 2.** scRNAseq of cultured NSCs in an in vitro model of oxidative stress. **(A)** UMAP of vehicle and H<sub>2</sub>O<sub>2</sub>-treated cultured NSCs yielded 10 subpopulations defined by gene expression profiles consistent with GO terms associated with various stages of the cell cycle, levels of quiescence or response to injury (left). UMAP comparison of H<sub>2</sub>O<sub>2</sub>-treated versus vehicle-treated NSCs indicated a shift towards quiescence, apoptosis, senescence, and injury response following oxidative stress (right). **(B)** Dot plot visualization of average expression of and percent of cells expressing select genes known to be associated with quiescence versus activation further confirmed subpopulation identities.

To confirm the cluster identities of the dormant and cycling cells, we looked at expression of known markers of dormancy and progression through the cell cycle. The largest subpopulation, which we designated as G<sub>1</sub> based on GO analysis, was characterized by moderate to high expression of genes involved in exit from dormancy, such as *Rpl15*, *Ran* and *Rps17* (Harris et al., 2021), as well as moderate expression of cell cycle genes such as *Top2a*, *Cdk1* and *Mki67* (**Figure 2A-B** and **Supplementary Figure 1A-B**) (Diril et al., 2012; Nielsen et al., 2020; Sun and Kaufman, 2018). The S phase population was characterized by a distinct upregulation of cell cycle gene expression (**Figure 2B** and **Supplementary Figure 1B**). The G<sub>2</sub>/M cluster showed continued high expression of exit from dormancy genes coupled with sharp downregulation of cell cycle genes relative to G<sub>1</sub> and S (**Figure 2A-B** and **Supplementary Figure 1A-B**). Two clusters of G<sub>0</sub>-like cells were found, both of which showed upregulation of genes linked with GO terms such as ion homeostasis and metabolic processes. The two clusters differed most notably in expression of genes associated with transition to/from deeper quiescence. Specifically, G<sub>0</sub>-dormant was characterized by high expression of quiescence-associated genes *Fabp7*, *Aldoc*, *Sparc*, *Clu* and *Id3* (Artegiani et al., 2017; Borrett et al., 2020; Dulken et al., 2017; Urbán et al.,

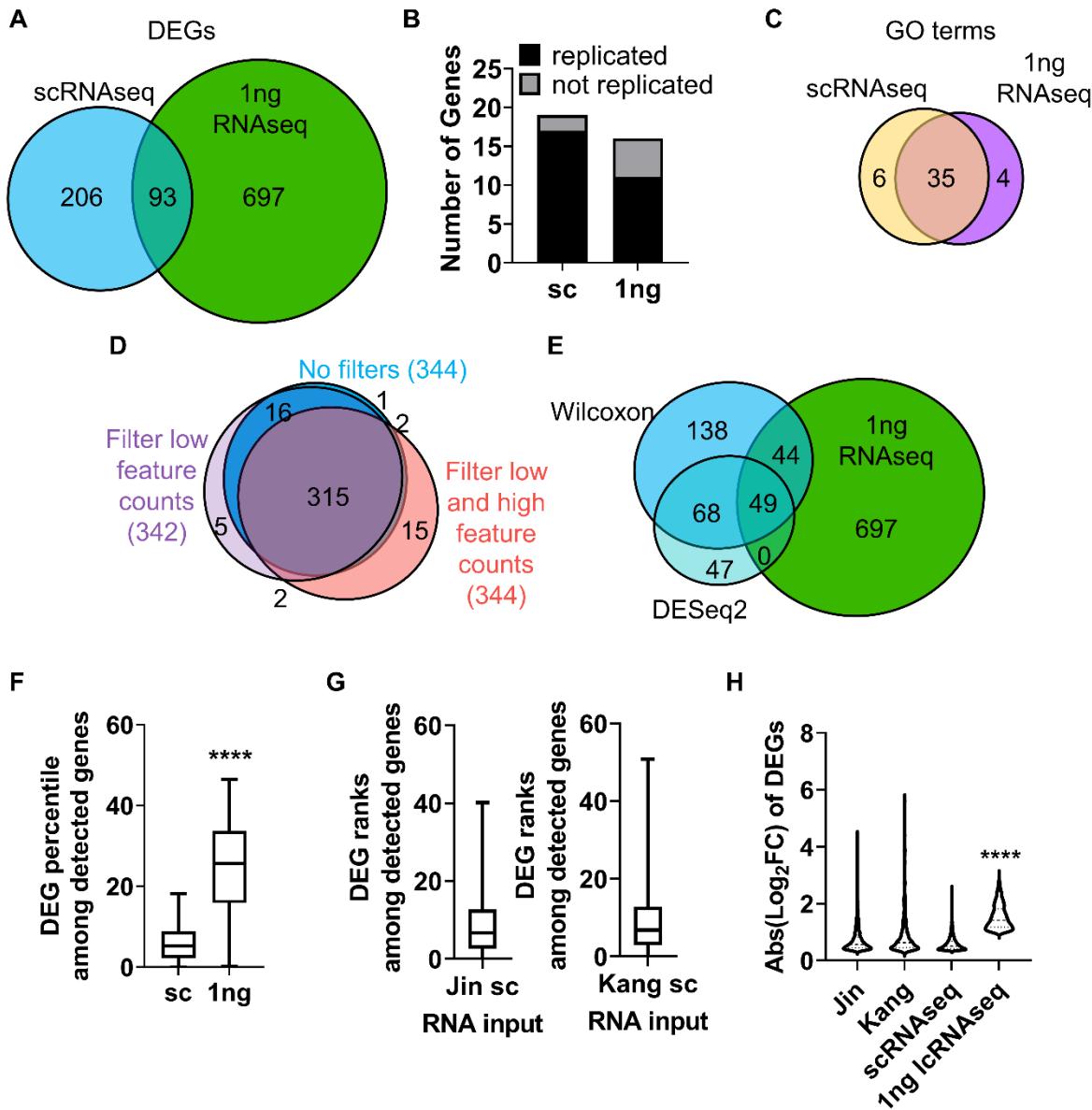


**Supplementary Figure 1.** Feature plots of characteristic genes for subpopulations identified by scRNAseq of NSCs with and without H<sub>2</sub>O<sub>2</sub>-induced oxidative stress in vitro. **(A)** *Ran* as a representative gene for exit from dormancy **(B)** *Top2a* as a representative gene for cycling cells. **(C)** *Aldoc* as a representative gene for dormancy/quiescence. **(D)** *Map2* as a representative gene for differentiating cells. **(E)** *Srxn1* as a representative gene for injured cells. **(F)** *Phlda3* as a representative gene for apoptosis.

2019), coupled with low expression of exit from dormancy-associated genes *Rpl5*, *Rps17* and *Ran* (**Figure 2A-B** and **Supplementary Figure 1A and C**) (Harris et al., 2021). G<sub>0</sub>-rest, in contrast, showed an upregulation of exit from dormancy genes and slight suppression of quiescence genes (**Figure 2A-B**). The transition group (T) was intermediate in expression of markers of dormancy and cell cycle, supporting its assignment as a transitional state between quiescent G<sub>0</sub> states and the cell cycle (**Figure 2A-B**). The two clusters of differentiating cells (D1 and D2) showed upregulation of genes associated with progenitor cell differentiation (*Sox4*, *Sox11* and *Map2*) (**Figure 2A-B** and **Supplementary Figure 1D**) (Artegiani et al., 2017; Hochgerner et al., 2018; Shin et al., 2015). Last, both injured and apoptotic clusters showed expression of genes associated with cell injury and death processes such as *Srxn1* and *Phlda3* respectively (**Figure 2A-B** and **Supplementary Figure 1E and F**) (Bell et al., 2015; Kawase et al., 2009).

We next identified DEGs between H<sub>2</sub>O<sub>2</sub> and vehicle treated cells using both the scRNAseq data and RNAseq from pooled cells using 1 ng total input RNA, an amount which yields sequencing data in the bulk RNA sequencing range (Walker et al., 2020). For scRNAseq, all clusters were combined within treatment for DEG analysis using the standard Wilcoxon test. 299 DEGs were identified between H<sub>2</sub>O<sub>2</sub> and vehicle treated NSCs using scRNAseq (**Supplementary Table 2**). 1ng RNAseq data was pre-processed using the coverage-based limiting-cell experiment analysis (CLEAR) pipeline to eliminated unreliable, lowly expressed transcripts (Walker et al., 2020) then DEGs were identified using DESeq2. The 1ng RNAseq comparison yielded 790 DEGs between H<sub>2</sub>O<sub>2</sub> and vehicle treated NSCs (**Figure 3A** and **Supplementary Table 2**). Comparison with the DEGs identified between sequencing methods revealed only 93 genes that were common to both platforms. Not surprisingly, most of the non-overlapping genes (697) were unique to the higher RNA input platform of 1 ng RNA, implying an expected greater sensitivity for DEG detection with greater RNA input. More unexpectedly, though, scRNAseq identified 206 unique DEGs compared to the 1ng input RNAseq. This low overlap between sequencing approaches from the same source NSCs implied either that one method was calling numerous false DEGs or that the two platforms had strongly different biases in what DEGs they can detect.

To determine how accurate and replicable the DEGs from scRNAseq and 1ng RNAseq were, we repeated H<sub>2</sub>O<sub>2</sub>/vehicle treatment of cultured NSCs in two more independent replications and analyzed gene expression of the top upregulated DEGs identified by scRNAseq and 1 ng RNAseq using quantitative real time PCR (qRT-PCR). Of the top 20 genes, we achieved effective primers for 19 of the top 20 scRNAseq DEGs and 16 of the 1 ng RNAseq DEGs. The majority of genes identified by both scRNAseq and 1 ng RNAseq were confirmed to be upregulated with qRT-PCR and there was no significant difference between methods in the number of genes replicated (**Figure 3B** and **Supplementary Figure 2A**). To further understand what distinguished DEGs called by scRNAseq versus 1ng RNAseq, we compared overlap of GO terms for upregulated genes in each platform. GO terms showed a high degree of overlap (35 out of 45 distinct categories) in the biological processes represented by the genes identified with both scRNAseq and 1 ng RNAseq (**Figure 3C** and **Supplementary Table 2**). This overlap in biological processes being triggered by oxidative stress, coupled with the high replicability of individual DEGs, implies that both platforms likely generated accurate DEGs that reflect true changes in cell activity. However, the low degree of overlap in the individual DEGs identified by the 2 platforms indicates that some form of bias affected the type of genes identified by each platform.



**FIGURE 3.** RNA input leads to bias in DEG discovery. **(A)** Venn diagram of DEGs identified by scRNAseq and 1 ng RNAseq with CLEAR filtering. **(B)** qRT-PCR analysis corroborated the majority of DEGs identified by scRNAseq and 1 ng RNAseq in cultured NSCs following H<sub>2</sub>O<sub>2</sub>-induced oxidative stress.  $\chi^2$  contingency test (df = 1) = 2.331, p = 0.127. **(C)** Venn diagram of GO terms associated with DEGs identified by scRNAseq and 1 ng RNAseq. **(D)** Venn diagram of DEGs identified by scRNAseq with no filtering, filtering for low feature counts, and filtering for both low and high feature counts (i.e. the default for Seurat analysis). **(E)** Venn diagram of DEGs identified by scRNAseq using the Wilcoxon test or DESeq2 compared with DEGs identified by 1ng RNAseq. **(F)** DEGs were ranked by percentile of average transcript count level relative to all detected gene counts. The highest count genes are the lowest number percentile (i.e. highest count relative to all detected genes would be in the 1<sup>st</sup> percentile). Box plot shows median and quartiles of percentiles of DEGs. DEGs identified by scRNAseq averaged the 6<sup>th</sup> percentile while DEGs identified by 1ng RNAseq averaged the 25<sup>th</sup> percentile. \*\*\*\*p<0.0001 unpaired t-test. **(G)** DEGs identified in scRNAseq data from Jin et al. (left) and Kang et al. (right) were ranked by transcript level percentile relative to all detected genes in those datasets. Similar to our scRNAseq data, the majority of DEGs were limited to the top 10% of expressed genes. **(H)** Comparison of fold changes in average transcript count between treatment groups for DEGs from all datasets indicated that DEGs identified by 1 ng RNAseq showed significantly larger fold changes in gene expression than DEGs identified by any of the scRNAseq datasets. \*\*\*\*p<0.0001 Dunnett's multiple comparison test.

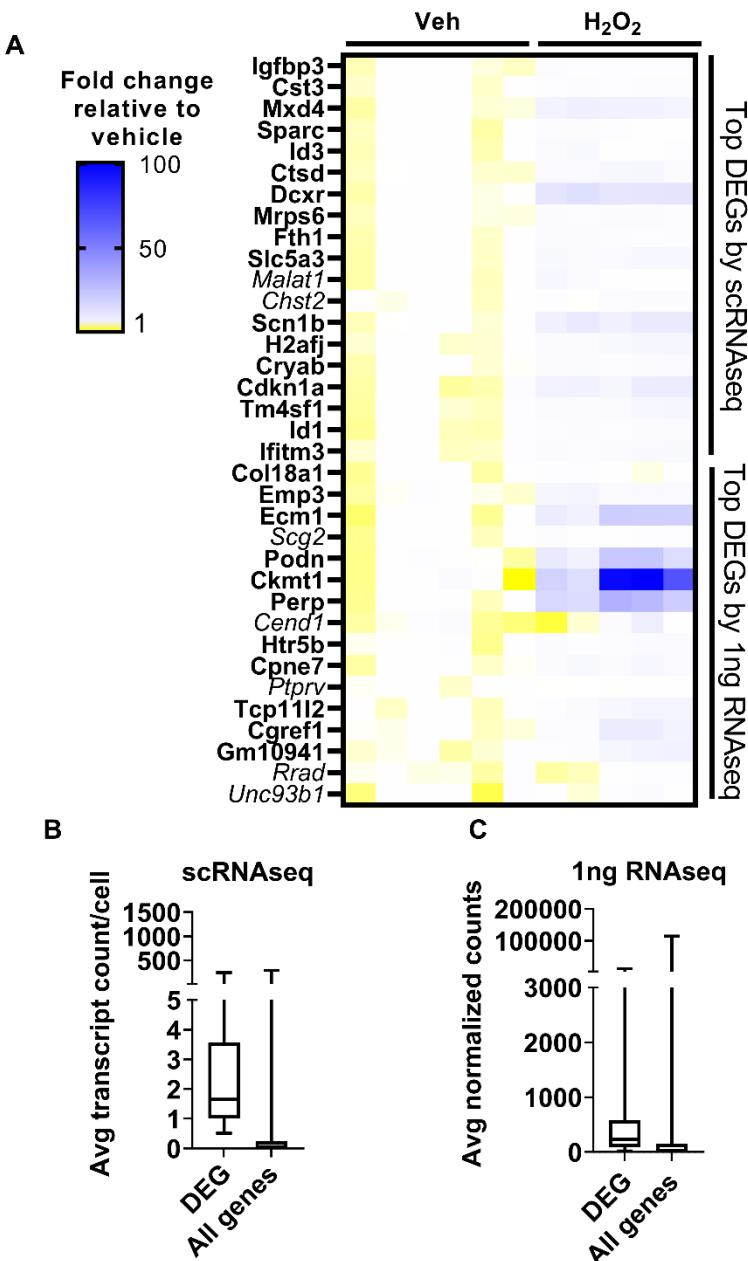
### RNA input amount drives bias in DEG discovery.

The first possibility we considered for the difference in DEG identification between sequencing platforms was that it was an artefact of the different data analysis and statistical techniques used to identify DEGs in scRNASeq versus 1ng RNASeq data. scRNASeq analysis involves several layers of filtering to compensate for technical limitations of low capture efficiency, high dropouts, and doublets. Two commonly used quality controls (QCs) set minimum and maximum thresholds for Unique Molecular Identifiers (UMIs) and feature counts to filter out damaged cells or doublets (Stuart et al., 2019). However, such filtering relies on predetermined values and could overzealously filter out valid cells on the ends of the spectrum of true cellular differences in RNA content and complexity. To determine if QC limits altered DEG discovery, DEG analysis was performed on scRNASeq data with and without filtering for low (1000 counts/cell) and high (2500 counts/cell) feature counts (**Figure 3D**). Removing both high and low feature counts resulted in a total of 344 DEGs with 92% overlap compared to fully filtered data (**Figure 3D**). Filtering out cells with only low feature counts yielded 342 DEGs with almost 93% of those genes also identified when all QCs are employed (**Figure 3D**). This comparison revealed high overlap in the DEGs identified with and without the various filters, indicating that QC filtering of scRNASeq data was not introducing the divergence in DEG discovery between sequencing approaches.

We next considered whether the different statistical tests used in scRNASeq and 1ng RNASeq might have contributed to differential DEG identification. For scRNASeq data, we used the default in Seurat pipeline implementation of the Wilcoxon test. Bulk RNASeq, in contrast, is typically (and was in our case) analyzed using DESeq2. To determine if the choice of statistical test affected the DEGs identified, we compared DEGs called by DESeq2 in both scRNASeq and 1ng RNASeq. DEG discovery using DESeq2 of scRNASeq data resulted in far fewer genes (a little over half the number identified using the Wilcoxon test), but 71% of those genes were common to both statistical methods (**Figure 3E**). DESeq2 has been noted to be more restrictive (Mou et al., 2020) so the fewer DEGs from that analysis is not surprising. While the two different methods of statistical analysis of scRNASeq data yielded highly overlapping sets of genes, they both only marginally overlapped (about 30%) with DEGs identified with DESeq2 in the 1 ng RNASeq dataset (**Figure 3E**). This comparison suggests that use of Wilcoxon versus DESeq tests is not likely the source of DEG discordance between scRNASeq and 1ng RNASeq approaches. Rather, these data suggest that there is something inherently different in the data generated with scRNASeq versus 1 ng RNASeq that is leading to different biases in DEG discovery with each approach.

To further explore the differences in DEGs identified in scRNASeq versus 1ng RNASeq, we next looked at the relative transcript count level of DEGs. Gene expression level within a cell impacts the likelihood of transcript capture and is known to significantly influence DEG analysis in single cell studies (Mou et al., 2020). Indeed, the low RNA input of scRNASeq is expected to result in high drop out and a consequent overall reduction of detected genes and therefore DEGs. The identification of DEGs in scRNASeq that were not identified via 1ng RNASeq in our data, however, is more unexpected. To determine if gene expression level led to the discordance in DEG discovery here, the averaged transcript counts per cell of DEGs identified in each sequencing approach were compared to those for all detected genes within that same approach (**Supplementary Figure 2B and C**). Genes that were not expressed in any cells were excluded to limit bias from zero inflation, especially in scRNASeq. To enable direct comparison between these two datasets, we then converted the transcript counts of DEGs within each sequencing

platform to percentiles relative to the average transcript counts of all detected genes. In this comparison, genes with the highest average transcript count per sample comprise the top percentiles (e.g. 1st %ile) of counts (**Figure 3F**). DEGs identified by 1 ng RNAseq had a higher median transcript count percentile value (~25%ile) and spanned a wider percentile range of



**Supplementary Figure 2.** Gene expression levels impact DEG identification. **(A)** qRT-PCR of the top DEGs identified by scRNAseq or 1 ng RNAseq of cultured NSCs treated with  $\text{H}_2\text{O}_2$  or vehicle show significant upregulation for most genes. Genes not significantly different between groups by Mann-Whitney test are italicized. N = 5-6 well replicates from 2 independent experiments. **(B)** Box plots show the median and quartiles of transcript count per cell for DEGs and all genes in scRNAseq data. **(C)** Box plots show the median and quartiles of transcript count per cell for DEGs and all genes in 1ng RNAseq data. Genes with zero expression in all cells were eliminated for (B) and (C).

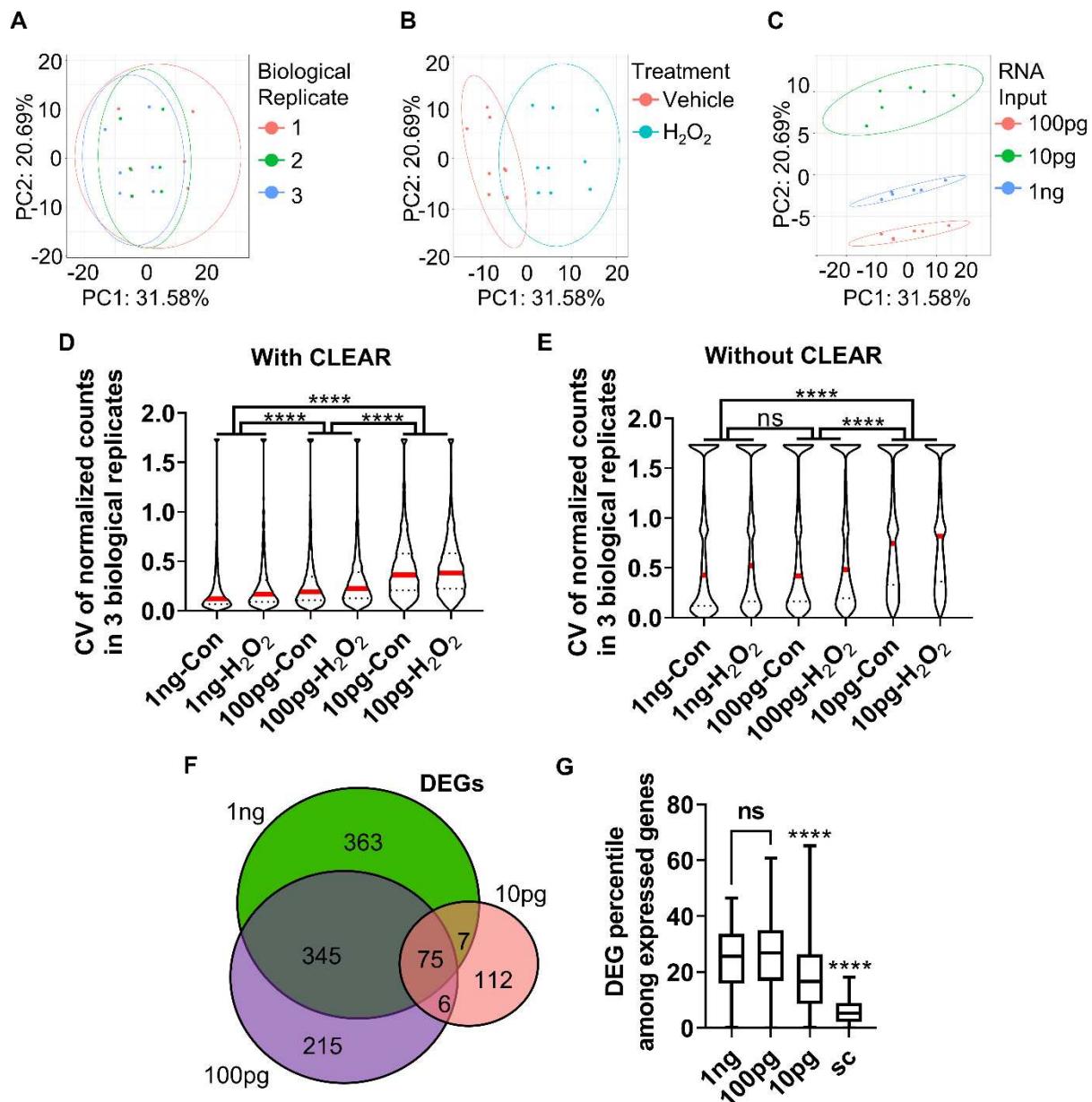
transcript count levels than those identified by scRNAseq, which were limited to the top 5-10% of detected transcripts (**Figure 3F**). It is important to note that these comparisons were made after eliminating genes with no counts, showing that a strong bias for high count genes persists in scRNAseq data even when undetected genes are excluded. To determine how widespread this bias for high count genes might be, we analyzed two other, published scRNAseq datasets from different model systems. Kang et al. performed droplet-based scRNAseq of human peripheral blood mononuclear cells treated with vehicle or interferon beta ex vivo (Kang et al., 2018). Jin et al. performed scRNAseq of all CD45 highly positive cells isolated from young and aged mouse dentate gyrus (Jin et al., 2021). Despite using different species, source tissue and manipulations, we found that scRNAseq DEG discovery in these two studies was similarly limited to the top 5-10% of high count transcripts (**Figure 3G**).

We next analyzed the relative fold change in counts of DEGs identified by scRNAseq and 1ng RNAseq. We found that scRNAseq in our experiment and in the two independent, unrelated comparison datasets yielded DEGs with lower fold changes than 1 ng RNAseq did. scRNAseq dataset DEGs showed an average fold change mostly between 1.5 and 2.5 (**Figure 3H**). Our 1 ng RNAseq data, in contrast, had a 3.0 fold change on average. Combined with the above findings on relative count level of DEGs, these findings suggest that the divergence in DEG detection between our scRNAseq and 1ng RNAseq data is driven by bias in scRNAseq data for detection of DEGs that derive from high count genes that show a more moderate fold change and in 1ng RNAseq data for DEGs from more moderate count genes that show a higher fold change between groups.

#### RNAseq determination of optimal RNA input for DEG discovery.

The comparison of scRNAseq and 1ng RNAseq shows that while both methods accurately identify DEGs, RNA input biased the types of DEGs that were detected. In studies seeking to identify specific causative genes implicated in a manipulation (e.g. disease model, drug treatment, altered gene expression), discovering genes that have high fold change and low expression level is particularly advantageous. Our data suggest that scRNAseq may be distinctly ill-suited for this purpose and a bulk RNAseq approach would be better. However, in the case of in vivo NSCs, the number of cells isolated from a single mouse is lower (100s-1000s) than that which is typically necessary for bulk RNAseq (100,000s). We used our in vitro oxidative stress model to determine the consequences of reduced RNA input, like that which may be encountered in a limiting cell (1c) RNAseq analysis of in vivo NSCs, for DEG detection. From the same samples, we compared transcriptional profiles derived from 1ng, 100pg and 10pg of RNA input. Similar to our analysis of the 1ng samples above, 100 and 10 pg RNAseq data was pre-processed using the CLEAR pipeline to eliminated unreliable, lowly expressed transcripts (Walker et al., 2020). Principal component analysis (PCA) showed that biological replicates completely overlapped with each other, indicating that the cultured NSCs were fairly uniform populations of cells, regardless of passage number or sex of the cells (**Figure 4A**). There was a clear separation between H<sub>2</sub>O<sub>2</sub>- and vehicle-treated NSCs on principal component (PC) 1 (**Figure 4B**). The clearest separation, however, was by RNA input level, along PC2 (**Figure 4C**).

Comparison of the median coefficient of variation (CV) between biological replicates showed that CV increased significantly with decreasing RNA input (**Figure 4D**). However, the difference in median CV was much more moderate between 1 ng and 100 pg (~46% increase from 1ng to 100pg) versus 1 ng and 10 pg (~77% increase). These data indicate lower precision and reproducibility as RNA input is reduced, particularly when it drops below 100 pg (**Figure 4D**).



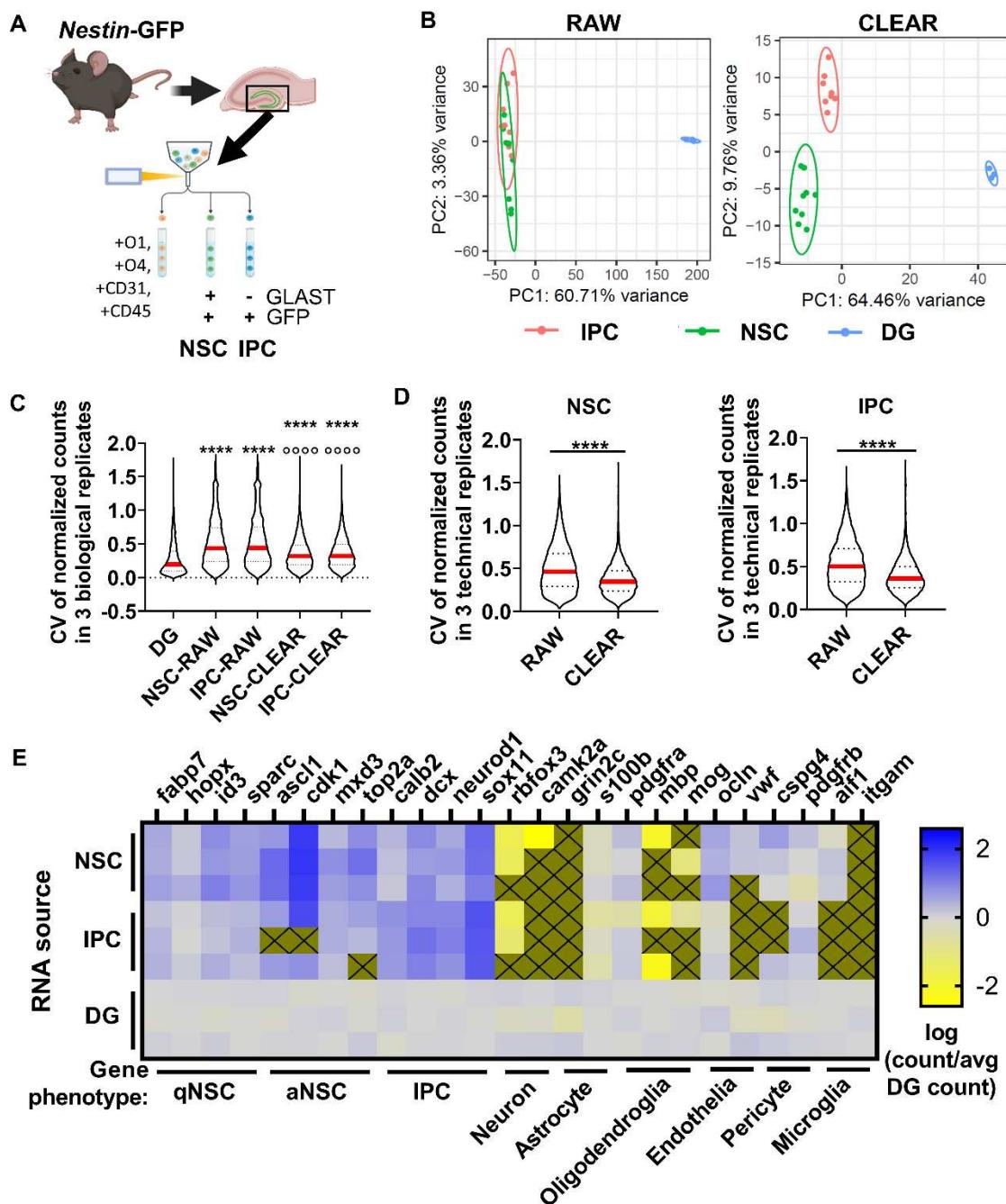
**FIGURE 4.** Determination of optimal input RNA amount with RNAseq and CLEAR that preserves unbiased DEG identification. **(A)** Biological replicates of cultured NSCs overlap in PCA. **(B)** PCA of cultured NSCs shows that cells undergoing oxidative stress diverge from control cells along principal component (PC) 1. **(C)** PCA of samples based on RNA input amount leads to significant separation between samples along PC2. **(D)** CVs were inversely related to RNA input amount. The difference in median CV between all RNA input levels were statistically significant. Median (red line) and quartiles (dotted line) are shown within violin plots of CV for all genes detected in all samples. **(E)** Analysis of the CV with vehicle and  $\text{H}_2\text{O}_2$ -treated samples of all three RNA inputs before CLEAR filtering shows substantially greater CVs at all RNA input levels, confirming the utility of CLEAR preprocessing. Median (red line) and quartiles (dotted line) are shown within violin plots of CV for all genes detected in all samples. **(F)** Venn diagram of DEGs identified in all three RNA input levels show majority overlap between 1ng and 100pg but low overlap with 10pg. **(G)** Comparison of transcript count percentiles for DEGs identified at the three different RNA input levels show no significant difference in DEG percentile between 1ng and 100pg RNA inputs. \*\*\*p<0.0001 One-Way ANOVA with Kruskal-Wallis test and Dunn's multiple comparisons in (D), (E), and (G).

Notably, CV was substantially larger in all conditions when data were analyzed without CLEAR pre-processing, emphasizing the utility of this step for extraction of reliable data in RNAseq analysis (**Figure 4E**). Comparison of DEGs identified in the CLEAR processed 10 pg, 100 pg, and 1 ng RNAseq datasets revealed 41.5% of DEGs overlapped between the 100 pg and 1 ng datasets, while only 9.0% of the 10 pg and 1 ng DEGs overlapped. 100 pg and 10 pg shared only 10.6% of DEGs (**Figure 4F**). Analysis of DEG gene counts as percentiles relative to average transcript counts for all genes further revealed that the 100pg RNA input DEGs had a median transcript count level comparable to that of the 1ng input while the 10pg RNA input dataset exhibited a significantly higher percentile of gene expression level (**Figure 4G**). Cumulatively, these data indicate that using 100pg of input RNA preserves data quality and many DEG characteristics of sequencing at the 1ng+ level without requiring its substantially higher number of cells. 10 pg of RNA, in contrast, shows greater variability across biological replicates and has less breadth in the count level of detected DEGs.

#### RNAseq enables DEG discovery from FACS isolated NSCs and IPCs from individual mouse hippocampi.

Standard bulk sequencing of adult hippocampal NSCs, a particularly sparse *in vivo* population, requires pooling of several mice to generate sufficient quantities of input RNA. Our findings suggest that when using CLEAR pre-processing, RNA input can be decreased substantially in an lcrNAseq approach and still yield reliable DEGs from a wide range of gene count levels. To test whether an RNAseq approach would be useful for transcriptional sequencing of *in vivo* NSCs, we used FACS to isolate NSCs and IPCs from 3 individual NestinGFP transgenic mice (Mignone et al. 2004) (**Figure 5A**). Using immunofluorescence of fixed tissue sections of adult NestinGFP mice, we confirmed that GFAP+SOX2+ radial glia like (RGL) NSCs and SOX2+ IPCs, but not DCX+ neuroblasts/immature neurons or NEUN+ mature neurons, expressed NestinGFP (**Supplementary Figure 3A-D**). We also confirmed NestinGFP expression in CD31+ endothelia and OLIG2+ oligodendroglial cells, as expected based on previous work (**Supplementary Figure 3E-F**) (Artegiani et al., 2017). To exclude the endothelia and oligodendroglial cells, we selected for cells immunonegative for CD31, O1 and O4. To specifically capture NSCs and IPCs separately, we used EAAT1 immunolabeling, with NestinGFP+EAAT1+ cells representing NSCs and NestinGFP+EAAT1- cells representing IPCs. To maximize RNA integrity, cells were sorted directly into lysis buffer and converted to cDNA libraries without an intervening RNA isolation step. Direct cDNA synthesis prevented measurement of RNA yield to compare with our *in vitro* studies, but we used the equivalent of a 60 cell RNA input amount (estimated to approximate 100-200pg input) to generate 3 technical replicates for each biological (mouse) replicate with a 300 cell complexity level for RNAseq and CLEAR filtering (**Figure 5A**). For comparison, we also sequenced in parallel 1 ng of RNA isolated from whole DG of 3 separate mice.

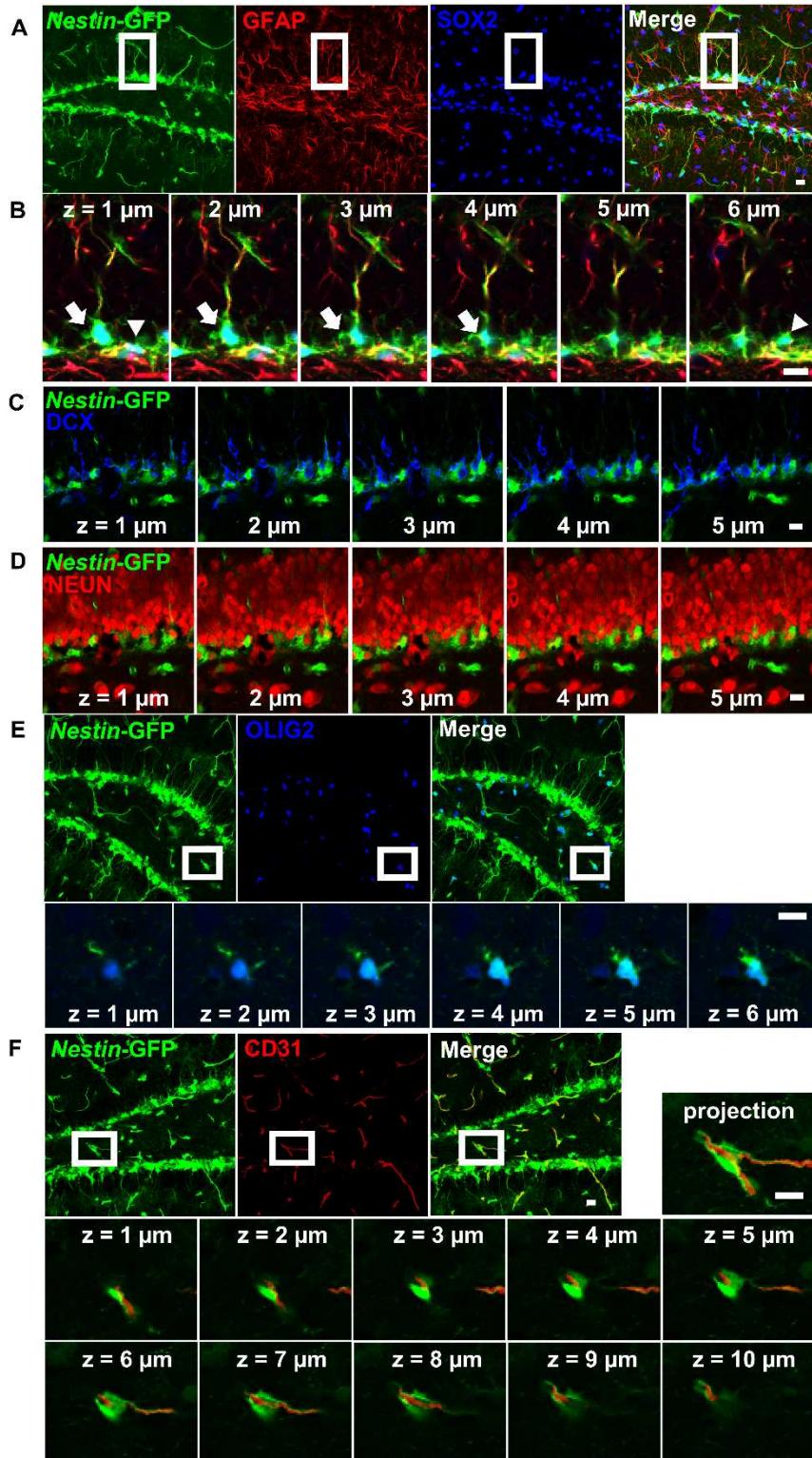
Principal component analysis of NSCs, IPCs, and whole DG revealed that CLEAR pre-processing both decreased the percent variance within NSCs and IPCs compared to PCA performed on raw data, and also separated NSCs from IPCs into non-overlapping populations (**Figure 5B**). Unfiltered (RAW) CVs for NSCs and IPCs across biological replicates were also more than double the whole DG CV (**Figure 5C**). CLEAR filtering reduced CVs by over 25% for NSCs and IPCs. CLEAR filtering also reduced the CV between technical replicates within the NSCs (by 25%) and IPCs (by 28%) (**Figure 5D**). Because tissue processing and handling can introduce variability when assessing freshly isolated cells, it was not surprising that the technical



**FIGURE 5.** DEG identification in vivo. **(A)** Workflow for DEG identification from hippocampal NSCs and IPCs of individual adult mice. DGs from Nestin-GFP+ mice were isolated and homogenized into single cell suspensions enabling FACS isolation of NSCs and IPCs for RNAseq. **(B)** PCA of transcriptome using unfiltered RAW and CLEAR filtered data. CLEAR filtering improves separation and variance between NSC and IPC samples. **(C)** CV for NSC and IPC biological replicates before and after CLEAR filtering. Dunn's multiple comparisons \*\*\*\*p<0.0001 vs DG °°°p<0.0001 vs raw RGL and raw IPC. **(D)** CV for NSC and IPC technical replicates (3/mouse) before and after CLEAR filtering. Unpaired t-test \*\*\*\*p<0.0001. **(E)** Heatmap shows expression of phenotypic genes in each of 3 biological replicates for NSCs, IPCs and whole DG, which was sequenced from 3 separate Wt mice in parallel. High expression of RGL phenotypic markers in NSCs and high expression of IPC phenotypic markers in IPCs confirmed cell type identities. Low expression of neuronal, astrocytic, oligodendroglial, endothelial, pericyte and microglial phenotypic genes confirmed exclusion of other major DG cell types. Gold boxes with an X were not detected.

RNA input in vitro. However, the significant improvement in CVs after CLEAR application confirm its utility in improving transcriptional data from limited starting material.

FACS-isolated NSC and IPC population identities were confirmed with expression of characteristic cell type markers for NSCs and IPCs (Figure 5E). Higher expression of quiescent



**Supplementary Figure 3.**  
Immunostaining of adult Nestin-GFP+ mouse DG demonstrating in vivo DG cell types. (A) RGL NSCs are Nestin-GFP+, GFAP+, SOX2+ while IPCs are Nestin-GFP+ and SOX2+ but GFAP-. (B) Serial z-stack images show the radial glial like process extending from an NSC (white arrow) in the subgranular layer through the granule cell layer of the DG. IPCs (arrowhead) are do not have a radial glial like process. SOX2+GFAP+ stellate morphology astrocytes were not observed to co-label with Nestin-GFP. (C) DCX+ immature neurons are not Nestin-GFP+. (D) NEUN+ mature neurons are similarly Nestin-GFP-. (E) Olig2+ oligodendroglia were co-labeled for Nestin-GFP. (F) CD31+ endothelial cells also co-labeled for Nestin-GFP.

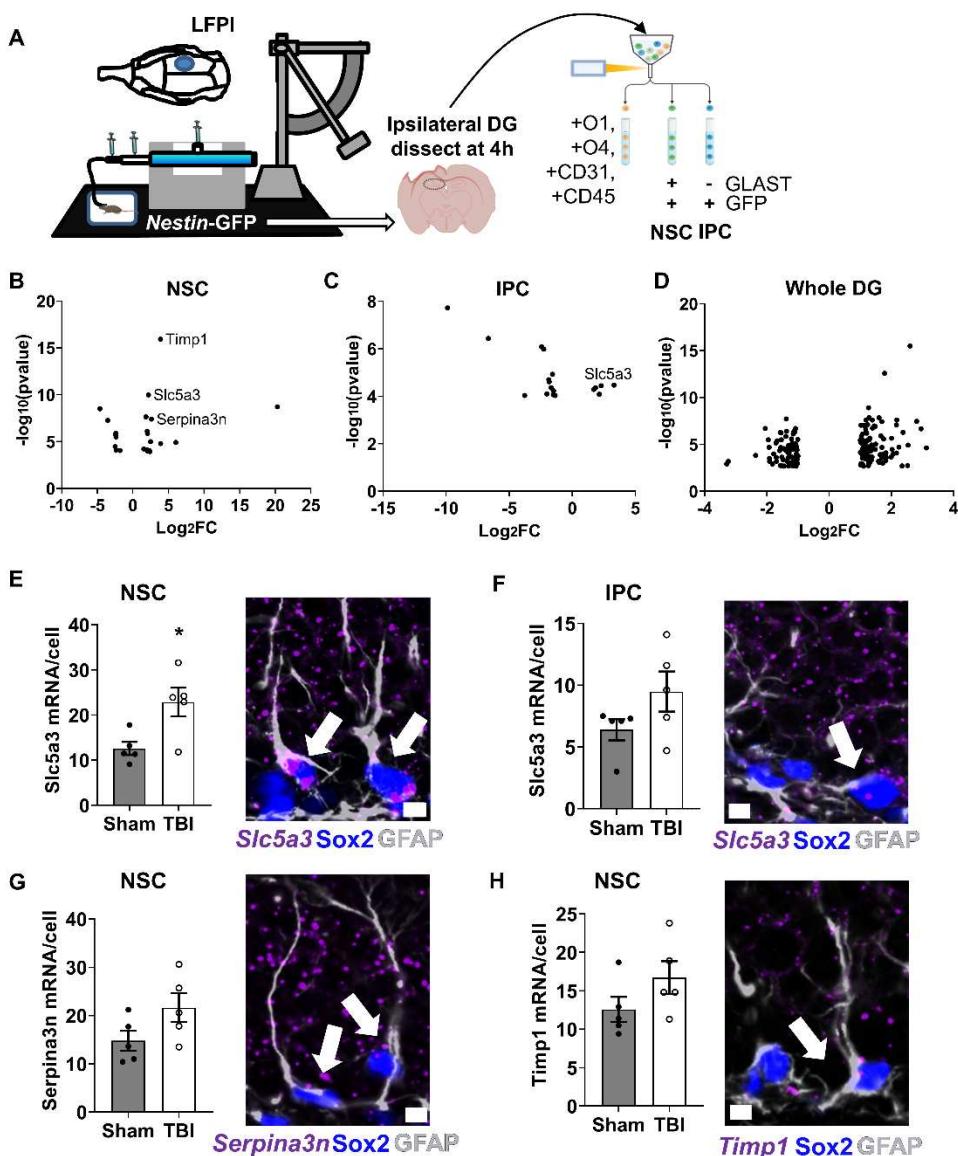
radial glial-like cell (qRGL) markers *Vim* and *Id3* as well as activated radial glial-like cell (aRGL) markers *Neurog2* and *Ascl1* were observed in the NSC samples compared to the IPC and whole DG samples, confirming accurate FACS isolation of NSCs (**Figure 5E**). Likewise, higher expression of intermediate progenitor cell (IPC) markers such as *Calb2* and *Sox11* in the IPC samples compared to the NSC and whole DG samples confirm accurate FACS isolation of IPCs (**Figure 5E**). Exclusion of other cell populations was confirmed via expression of phenotypic genes for neurons, astrocytes, oligodendroglia, endothelia, pericytes and microglia (**Figure 5E**). 177 DEGs were identified between NSCs and IPCs acutely isolated from adult mouse DGs (**Supplementary Table 3**). All together, these findings indicate that RNAseq of adult DG NSCs and IPCs can be achieved from a single mouse per sample replicate with data quality similar to that derived from more bulk-like RNA sequencing.

#### RNAseq identifies DEGs in NSCs and IPCs induced by lateral fluid percussive brain injury.

As proof of principle, we applied our RNAseq workflow to a LFPI model of TBI and identified DEGs in adult mouse NSCs and IPCs *in vivo* (**Figure 6A**). NSCs and IPCs from the DGs ipsilateral to injury of 3 individual mice were FACS isolated 4 hours after LFPI or sham treatment. Whole DGs from separate, but similarly treated, mice were also processed for comparison by 1 ng RNA input sequencing. A total of 6319 genes were identified in NSCs, IPCs, and whole DGs (**Supplementary Table 4**). 23 DEGs were identified in NSCs after LFPI with 15 significantly upregulated and 8 significantly downregulated (**Figure 6B**). In the IPC population, 5 genes were significantly upregulated while 13 genes were downregulated in LFPI mice compared to sham mice (**Figure 6C**). In whole DG, 188 DEGs were identified with 106 significantly upregulated and 82 significantly downregulated in LFPI mice compared to sham mice (**Figure 6D**). Comparison of the DEGs identified in NSCs and IPCs following LFPI did not overlap with DEGs identified on a whole DG level and NSC and IPC DEGs only shared one gene in common (*Slc5a3*), emphasizing the importance of examining individual cell types, even for cells as closely related as NSCs and their IPC progeny. RNAscope fluorescent *in situ* hybridization (FISH) combined with immunohistochemical staining verified transcriptional upregulation of *Slc5a3* in NSCs and showed a similar trend in upregulation in IPCs following LFPI (**Figure 6E and F**). We also verified two other upregulated DEGs using RNAscope FISH: *Serpina3n* (**Figure 6G**) and *Timp1* both trended higher in expression in NSCs of LFPI mice than sham 4h after injury. (**Figure 6H**). These findings indicate that RNAseq of acutely isolated, *in vivo* NSCs and IPCs can reliably identify DEGs in an injury model.

#### **Discussion**

Endogenous adult hippocampal NSCs provide a source of both cellular and biochemical support for tissue homeostasis. Characterizing these cells at baseline, as well as after injury, may lead to therapeutically relevant strategies for promoting optimal brain function. However, studying adult DG NSCs is challenging due to their relatively low cell number, residence within a complex niche, and their inherent heterogeneity. To overcome this challenge, researchers have increasingly turned to scRNAseq approaches for transcriptional profiling (Hochgerner et al., 2018; Knight and Serrano, 2018; Kulkarni et al., 2019; Shin et al., 2015; Zeng et al., 2018). However, we show here that, despite the many strengths of this approach, scRNAseq may not be the ideal method to answer certain types of research questions. Specifically, we show that scRNAseq analysis of differential gene expression may miss genes that are more moderately expressed and show large fold changes in expression. We show that an lRNAseq approach can be used to help circumvent this problem and discover DEGs from a broader count range and with greater fold



**FIGURE 6.** RNAseq and CLEAR filtering enable accurate DEG identification in NSCs and IPCs in mice with lateral fluid percussive brain injury (LFPI). **(A)** Mice (n=3 per group) were given lateral fluid percussive injuries or sham treatment. Four hours after injury, DGs from the ipsilateral side of the brain were isolated and processed for RNAseq as described in Figure 5A. **(B)** Volcano plot of DEGs identified in NSCs after LFPI. **(C)** Volcano plot of DEGs identified in IPCs after LFPI. **(D)** Volcano plot of DEGs identified in whole DG after LFPI. **(E)** Quantification of *Slc5a3* mRNA via RNAscope in NSCs (identified via GFAP+SOX2+ immunolabeling) confirmed upregulation after LFPI (left).  $p=0.0185$ . Mean  $\pm$  SEM of mean RNA puncta per NSC area is shown with 5 individual mice replicates. Representative image with arrows pointing to NSCs with *Slc5a3* expression (right). **(F)** Quantification of *Slc5a3* mRNA in IPCs similarly showed a trend in increased expression after LFPI (left).  $p=0.1324$ . Representative image with arrow pointing to IPC with *Slc5a3* expression (right). **(G)** Quantification of *Serpina3n* mRNA in NSCs show a trend for upregulation after LFPI (left).  $p=0.0954$ . Representative image with arrow pointing to NSC with *Serpina3n* expression (right). **(H)** Quantification of *Timp1* mRNA in NSCs showed increased expression that did not reach significance ( $p=0.1627$ ) after LFPI (left). Representative image with arrow pointing to NSC with *Timp1* expression (right). (E-H) n=5 per group, unpaired t-test, \* $p<0.05$ . Scale bars = 5  $\mu$ m.

changes, and that such an approach can be adapted for transcriptional profiling and DEG identification in NSCs acutely isolated from adult mouse DG.

Using a model of oxidative stress in cultured NSCs, we identified DEGs from the same source samples using both an scRNAseq and 1ng RNAseq approach. By using cells from the

same biological replicates for both sequencing platforms, we avoided potential variation in gene expression induced by a difference in tissue processing. Yet, we found little overlap in the identification of DEGs by these two methods. Both methods yielded accurate DEGs that could be confirmed with qRT-PCR. Furthermore, GO analysis of the DEGs identified using the two different approaches implicated mostly the same biological pathways being triggered by injury. However, the individual genes identified by each method were mostly unique within each platform. This lack of overlap was maintained when scRNAseq data were analysed without filtering for high or low feature counts and when the same statistical method was used for DEG discovery in both datasets, suggesting it was not an artefact of data processing/analysis.

When we probed the difference in the kinds of DEGs identified by scRNAseq versus 1 ng RNAseq, we found that scRNAseq identified high count genes with lower fold changes while 1 ng RNAseq identified genes with a wider range of count levels, including many genes with moderate counts and higher fold changes. Two independent scRNAseq datasets from different tissues and experimental designs showed a notably similar DEG profile in terms of relative count level and fold change as in our data. This similarity across such different experiments suggests that the restriction of DEGs to high count genes with lower fold change may be an inherent limitation of scRNAseq via 10x Chromium approach. Importantly, this bias in count level and fold change emerged in comparison to transcripts detected above 0 in each respective dataset. scRNAseq, of course, yields fewer detected genes than a bulk approach but this difference in DEG profile emerged even among genes that would be recorded in the dataset as detected and therefore having been evaluated for potential to be a DEG.

There are several likely contributors to the selectivity of scRNAseq data for generating DEGs from high count transcripts. The first and most obvious potential contributor is the high zero-count rate in scRNAseq data, due to the very low RNA input from a single cell. Transcripts with counts near threshold of detection will be characterized by many 0-count genes and therefore have relatively high variability from cell to cell. This high variability may make such genes less likely to be detected as significantly different between two conditions. Discovery of genes with higher counts, and therefore lower zero-influenced variability, would be favored. By being higher count, fold change is similarly likely to be more constrained by a ceiling effect, yielding lower fold change DEGs. Analytical processing compensations for heteroskedasticity may also play a role. Heteroskedasticity is the phenomenon whereby genes with relatively low counts exhibit higher fold changes. Thus, general DEG analysis methods, such as DESeq2, and scRNAseq analysis pipelines, such as Seurat, correct for this problem by applying a variance-stabilizing preprocessing step that transforms the data and minimizes the effect of count-based technical noise on ratio-based outputs such as log fold change in gene expression (Ahlmann-Eltze and Huber, 2021; Love et al., 2014). Although heteroskedasticity is recognized for any count-based data, standard bulk level RNAseq is inherently more robust and delivers generalized expression data for hundreds of thousands of cells. This facilitates accurate correction for heteroskedasticity, while in scRNAseq, low biological replicates and very low input RNA amounts may lead to over-correction of true variation in gene expression by variance-stabilizing preprocessing (Love et al., 2014; Mou et al., 2020).

While our findings suggest that scRNAseq and low-input population level RNAseq are both similarly accurate, as measured by replicability, they suggest different utility of each approach depending on experimental goal. If characterization of cellular heterogeneity is the goal, scRNAseq is obviously superior. However, discovery of genes that are moderate to lowly

expressed in cells yet exhibit higher fold changes upon stimulation is also a common goal, particularly when searching for candidate molecular mediators or druggable targets in models of injury and disease. Relying on scRNAseq alone for DEG discovery could therefore limit the scope of understanding of disease or injury-associated transcriptional signatures. The difference in DEG characteristics between scRNAseq and 1ng RNAseq shown here emphasizes the importance of identifying the ultimate goals and readouts of an experiment to choose the method that best addresses the needs of the study.

Although several studies have provided invaluable insight into hippocampal NSC biology using scRNAseq (Artegiani et al., 2017; Hochgerner et al., 2018; Shin et al., 2015), standard bulk RNAseq to study endogenous adult hippocampal NSCs has been challenging, as evidenced by the dearth of such studies. Though NSCs have been successfully profiled when combined with their IPC progeny via bulk RNAseq approach (Adusumilli et al., 2021), NSCs alone are sufficiently rare that their “bulk” level transcriptional profiling is especially difficult. Here, we optimized a protocol that enables population level transcriptional analysis of this rare cell type from individual mice. First, we used cultured adult hippocampal NSCs to determine 100 pg as a lower RNA input amount that enables profiling of DEGs that are comparable to those obtained with standard bulk sequencing. Using this threshold of 100 pg RNA input as a guide for *in vivo* analysis, we isolated adult DG NSCs and IPCs from individual Nestin-GFP reporter mice to profile the transcriptomes of each population at a 300 cell complexity level.

To ensure accurate identification of DEGs from this low input level, we applied CLEAR filtering which was previously shown to minimize technical noise due to limited RNA input (Walker et al., 2020). In brief, CLEAR preprocessing removes transcripts from analysis that are detected below a threshold. That threshold is determined by observing where on an mRNA transcript sequence RNAseq read fragments map. When reads show preferential mapping to only 3' and 5' ends of their mRNA transcripts, it indicates strong RNA degradation, a feature which predominates as transcript count drops. CLEAR filtering removes transcripts with counts below the threshold where most transcripts start to show this pattern of mapping more exclusively towards the 3' and 5' ends of their mRNA sequences. CLEAR filtering of our data improved the coefficient of variation between biological replicates for normalized counts at all RNA input amounts *in vitro* and *in vivo*, and was essential for effective transcriptional separation of *in vivo*-isolated NSCs and IPCs via PCA. We showed that this workflow, with CLEAR filtering, accurately profiled the transcriptomes of NSCs and IPCs. This ability to capture population level complexity with a low amount of RNA input is valuable when studying rare cell populations in complex experimental models of disease or injury where maximizing biological replicates is critical but limited by time and labor costs.

To demonstrate the utility of our approach for transcriptional profiling of *in vivo* NSCs, we applied it in a proof-of-principle experiment identifying TBI-associated DEGs from *in vivo* NSCs and IPCs. Using our workflow, we obtained cell type specific DEGs four hours after the LFPI model of TBI. In response to LFPI, NSCs and IPCs showed mostly unique DEGs, with only one exception: both upregulated *Slc5a3*, a sodium-coupled inositol transporter protein that maintains osmotic pressure in cells and regulates intracellular myo-inositol levels (Andronic et al., 2015). We confirmed transcriptional upregulation of *Slc5a3*, as well as two other DEGs, *Serpina3n* and *Timp1*, with RNAscope *in situ* hybridization paired with immunofluorescent staining to identify NSCs and IPCs. These findings suggest that our population-level approach to RNAseq of isolated

NSCs and IPCs can effectively identify replicable changes in gene transcription after an injury stimulus.

There are several limitations to this study. First, we only compared bulk RNAseq to scRNAseq using a 10x Chromium platform. It is therefore possible that other approaches to scRNAseq would yield different results than what we found. However, many of the limitations we noted seemed inherent to working with very low RNA input levels and were not exclusive to our experimental platform (in vitro NSCs) or any one approach to data analysis. Second, the specific cell/RNA input levels that we identified as yielding more bulk-like range in DEG count level and fold change may not apply outside of our selected cell population (adult NSCs). RNA content, cellular heterogeneity and capture efficacy of different cell types and tissue sources will likely influence the appropriate cell input needed in other models. Third, transcriptomics does not equate with proteomics. Particularly in NSCs, there appears to be substantial translational priming, in which mRNA is produced but not translated (Denninger et al., 2020; Kjell et al., 2020). Thus, both transcriptional and proteomic analyses are needed to accurately characterize NSCs in health and disease.

In conclusion, we present a comparison of two different approaches to transcriptional profiling, scRNAseq and population level RNAseq, of adult hippocampal NSCs, a rare cell type that is difficult to study *in vivo*. We found that each had their strengths, as well as weaknesses, which should be balanced with the needs of each specific study. We have shown here that our method for *in vivo* transcriptional profiling in a bulk-like IcRNAseq approach can provide valuable information about rare cell populations that are traditionally difficult to study *in vivo*. Thus, we present our workflow as an addition to the transcriptional toolbox for studying limited *in vivo* cell types moving forward.

## Materials and Methods

### Animals

NestinGFP mice (Mignone et al., 2004) (Jackson Labs, #033927) and Wt C57Bl/6J mice (Jackson Labs, #000664) were housed in a 12 hour light-dark cycle with food and water *ad libitum*. To isolate brains for fluorescence *in situ* hybridization, adult mice (6-9 wk old) were anesthetized with an intraperitoneal injection of ketamine (87.5 mg/kg) and xylazine (12.5 mg/kg). Mice were then transcardially perfused with ice cold PBS followed by cold 4% PFA. For immunofluorescence and whole DG RNA isolation, mice were only perfused with ice cold PBS. For IcRNAseq experiments, mice were only perfused with ice cold HBSS without calcium or magnesium. This study was approved by the Institutional Animal Care and Use Committee (IACUC) at the Ohio State University in accordance with institutional and national guidelines.

### DG isolation and FACS for RNAseq

To isolate DGs for subsequent fluorescence activated cell sorting (FACS), adult mice (6-9 wk old) were anesthetized and perfused with HBSS as described above. Following perfusion, brains were removed and placed in cold HBSS on ice. To expose the hippocampus, brains were bisected along the midsagittal line and the cerebellum and diencephalic structures were removed. Under a dissection microscope (Zeiss), the DG was excised using a beveled syringe needle and placed in ice cold HBSS without calcium or magnesium. DGs were then mechanically dissociated with sterile scalpel blades before enzymatic dissociation with a pre-warmed papain (Roche 10108014001)/dispase (Stem Cell Technologies 07913)/DNase (Stem Cell Technologies

NC9007308) (PDD) cocktail at 37°C for 20 min. Afterwards, the tissue was again mechanically disrupted by trituration for 1 min. Dissociated cells were collected by centrifugation at 500g for 5 min before resuspending in HBSS without calcium/magnesium. Cells were then filtered through a 35µm nylon filter before staining with fluorescent antibodies (**Supplementary Table 5**) on ice for 30 minutes. During the last 10 minutes of staining, Hoechst dye was added for live/dead discrimination. All cells were washed twice following staining and immediately sorted as NSC or IPC populations based on fluorescent markers with the FACSaria III (BD Biosciences). CD31-, CD45-, O1-, and O4 negative live cells were designated as NSCs if double positive for GLAST and GFP or IPCs if GFP positive and GLAST negative. Three technical replicates of 300 cells each were sorted from each individual mouse into 1.5mL microcentrifuge tubes containing cell lysis buffer buffer from the Clontech SMART-Seq HT (Takara) kit for direct cDNA synthesis and RNAseq library generation.

### Cell culture

NSCs were isolated from adult DGs of C57Bl6/J mice as described in (Babu et al. 2011). Two separate lines, one from 4 pooled C57Bl6/J male mice and one from 4 pooled C57Bl6/J female mice, were used in experiments between passage 5 and 15. NSCs were cultured on poly-D-lysine (Sigma) and laminin (Invitrogen) coated plates in Neurobasal A media (Invitrogen) with 1x B27 supplement without vitamin A (Gibco), 1x glutamax (Invitrogen) and 20 ng/ml each of EGF and FGF2 (Peprotech). There were no inherent differences in morphology or proliferation between NSC cultures and both lines differentiated into neurons and glia upon culture in differentiation conditions, as we previously showed (Denninger et al., 2020). All cultures were verified to be mycoplasma-free. For oxidative stress experiments, NSCs in triplicate were treated at 70% confluence with 500 µM H<sub>2</sub>O<sub>2</sub> (Sigma) or equal volume of vehicle (PBS) for 24 hours. All cells were harvested with brief accutase treatment and one wash with HBSS for RNA isolation or scRNAseq.

### RNAseq of cultured cells and whole DG

RNA from 30,000 cultured adult NSCs or whole DGs were isolated with the Clontech Nucleospin RNA XS Plus isolation kit (Takara 740990.10) per manufacturer protocol. RNA quality (RNA Integrity Number or RIN) and quantity was assessed using Agilent BioAnalyzer RNA 6000 Pico Kit and the Invitrogen Qubit RNA HS Assay kit (Invitrogen), respectively. All cultured samples used a RIN value of 10 while whole DG samples had RIN values over 8. RNA from cultured NSCs was serially diluted to 10-, 100-, and 1000-pg for RNA input quantity studies. Whole DG libraries for bulk RNAseq were generated with the NEBNext Ultra II Directional RNA Library prep kit (New England Biolabs). The Clontech SMART-Seq HT (Takara) kit was used for global preamplification of cultured and 20% (~60 cell input) of the 300 FACS-isolated NSC cell lysate for low input RNAseq. Purified library products were then used in HiSeq 4000 paired-end sequencing (Illumina) to a depth of 15–20 million 2 × 150 bp clusters. FASTQ files generated for each library were trimmed using AdapterRemoval v2.2.0 (Schubert et al., 2016), ensuring that all sequencing adapters were removed and that the average quality score for each read was above Q20 (representing 1 in 100 Illumina base error rate). Reads which were aligned by HISAT2 v2.0.6 (Kim et al., 2015) against rRNA, mtDNA, or PhiX bacteriophage (Illumina spike-in control) sequences, retrieved from NCBI RefSeq (O’Leary et al., 2016), were removed from each FASTQ file, as these do not represent gene expression signal. All remaining reads were aligned against the reference mouse genome GRCm38p4 with HISAT2. The resulting BAM alignment files were sorted and indexed before further analysis.

Alignments were quantified using the featureCounts utility from the Subread package v1.5.1 (Liao et al., 2014, 2013) in unstranded mode using GENCODE (Harrow et al., 2012, 2006) mouse gene reference version M14 in GTF format. Custom Python scripts were used to produce a formatted gene expression counts table from the raw output of featureCounts. RNAseq Quality metrics were derived using a modification of the QuaCRS quality control workflow (Kroll et al., 2014) which includes running RNA-SeQC v1.1.8.1 (DeLuca et al., 2012), FASTQC v0.11.5, and RSeQC v2.6.2 (Wang et al., 2012). Finally, coverage maps of each BAM file were derived using the Bedtools ‘genomcov’ utility v2.27.0 (Quinlan and Hall, 2010).

RNAseq coverage maps were processed with the CLEAR v1.0 (Walker et al., 2020) workflow to determine which genes were reliably quantified. In brief, for each transcript in the UCSC GRCm38 release (Kuhn et al., 2013), a parameter ( $\mu_i$ ) is calculated, which represents the positional mean of the reads covering the transcript, normalized to the length of each sequence. Sequential bins of 250  $\mu_i$  values each, ordered by descending expression, are fit to a sum of two beta distributions (Gupta and Nadarajah, 2004) for the determination of two free parameters, which are thresholded to determine genes which “pass” CLEAR. Unless otherwise noted, these passing genes are used for downstream analysis. DEGs were derived from filtered RNAseq expression counts tables following the DESeq2 v1.20.0 protocol (Love et al., 2014) implemented in R v3.5.0. DEGs were then processed with the pcaExplorer v2.6.0 (Marini and Binder, 2019) visualization package to produce principal component analysis (PCA) projections using the default settings on r-log transformed counts.

### scRNAseq

30,000 NSCs were pooled from triplicate  $\text{H}_2\text{O}_2$ - or vehicle-treated cultures and loaded onto the 10X Genomics single cell sequencing platform using the standard kit. The 3' RNA-seq library was sequenced using paired-end 150bp approach on an Illumina HiSeq 4000 sequencer. Data from vehicle-treated cells was previously published in (Denninger et al., 2020) and similar analysis was performed here, but on the combined data, including both vehicle-treated and  $\text{H}_2\text{O}_2$ -treated cells. CellRanger v3.0.2 (Zheng et al., 2017) was used to demultiplex, align, and deduplicate sequencing reads in BCL files. Single-cell data in feature-barcode matrices were then processed using Seurat v3.0.1's default pipeline (Butler et al., 2018; Stuart et al., 2019) to identify unsupervised cell clusters and generate a uniform manifold approximation and projection (UMAP) plot. Seurat's feature expression analysis was used to visualize genes markers known to be expressed in cell populations displayed in the UMAP plot.

### qRT-PCR

Cultured NSCs were lysed in culture plates and RNA was isolated with the BioRad Aurum<sup>TM</sup> Total RNA Mini Kit according to the manufacturer protocol. Isolated RNA was quantified and assessed for quality using the BioTek Epoch Microplate Spectrophotometer. cDNA was synthesized with the BioRad iScript<sup>TM</sup> cDNA Synthesis Kit in the ThermoFisher Applied Biosystems 2720 Thermal Cycler according to manufacturer protocol. qRT-PCR was performed in the BioRad CFX96 Touch Real-Time PCR Detection System with BioRad SsoAdvanced Universal SYBR Green Supermix and primers listed in Supplementary Table 5.  $\Delta\Delta\text{Ct}$  values generated with normalization to housekeeping gene Rpl7 then converted to fold change (relative to vehicle).

### Lateral fluid percussion injury

All surgical procedures were performed as previously described (Tapp et al., 2020). Briefly, 6–9-week-old mice were anesthetized with 4% isoflurane gas in an induction chamber for 4 min. Mice were then positioned in a stereotaxic frame before making a sagittal incision to expose the cranium. Midway between bregma and lambda on the right parietal bone, a 3.0-mm craniectomy was trephined, leaving the intact dura mater exposed. A modified portion of a Luer-Loc syringe (3.0-mm inside diameter) was secured over the craniectomy site with cyanoacrylate adhesive. Mice were placed in their home cages on a heating pad to recover. Once mice resumed normal activity, mice were returned to the vivarium for 24 hours. The next day, mice were anesthetized with 4% isoflurane in an induction chamber for 4 min. Using the modified Luer-Loc syringe, mice were connected to the fluid percussion injury device (Custom Design & Fabrication). For mice designated to the TBI group, a prepositioned pendulum was released onto the end of the LFPI device to deliver a fluid pulse onto the exposed dura mater, inducing a moderate LFPI. Sham mice were attached to the LFPI device but did not receive a fluid pulse. The modified syringe and adhesive were removed following LFPI or sham treatment. The incision was stapled closed. All animals were placed on a heating pad and injury severity was assessed with the self-righting reflex test. After the subjects demonstrated the righting reflex, they were returned to their home cages on a heating pad. 4 hours after injury, mice were perfused for either histology, RNA isolation, or FACS.

### Histology (IF and RNAscope)

For identification of in vivo cell types, brains from NestinGFP mice were harvested and fixed overnight at 4°C in 4% PFA before overnight equilibration in 30% sucrose. Serial 40µm sections were rinsed in PBS before blocking with 1% normal donkey serum (Jackson) and 0.3% Triton X-100 in PBS for 30 minutes at room temperature. Sections were incubated with primary antibodies (**Supplementary Table 5**) overnight at 4°C in blocking solution on an orbital shaker. The next day, sections were washed with PBS and incubated with fluorescently conjugated secondary antibodies diluted in blocking solution (**Supplemental Table 5**) for 2h at room temperature and counterstained with Hoechst 33342 (1:2000) for nuclear visualization. Sections were then washed and mounted onto slides before coverslipping with ProLong Gold anti-fade solution (Molecular Probes). Slides were imaged in 1-µm z-stacks on an LSM700 confocal microscope (Zeiss) with a 40× oil objective.

For RNAscope, four hours after LFPI or sham injury, brains from NestinGFP mice were harvested and fixed overnight at 4°C in 4% PFA before serial overnight equilibration in 10%, 20%, and 30% sucrose. Fixed equilibrated tissue was snap frozen in OCT in a dry ice/100% ethanol bath and stored at -70°C. 12 µm cryosections, 1 section per slide, were prepared with a cryostat. Slides were stored at -70°C with desiccant until staining. RNA in situ hybridization was performed with RNAscope Multiplex Fluorescent v2 Assay (Advanced Cell Diagnostics) according to manufacturer recommendations for using fixed frozen tissue samples with the following modifications to enable concurrent immunohistochemical staining. The pretreatment steps were replaced with a 15 min modified citrate buffer (Dako) antigen retrieval step in a steamer at 95°C. To enable subsequent immunohistochemical staining, the protease III step was excluded. Probes for mouse Slc5a3 (ACD custom design NPR-0006102), mouse Serpina3n (ACD 430191), and mouse Timp1 (ACD 316841) RNA were hybridized to tissue prior to immunohistochemical staining for GFAP and SOX2 protein. Immunostaining for GFAP and SOX2 was conducted as described above with the following exceptions. Blocking was performed with 10% normal donkey

serum in TBS-1% BSA. Antibody incubations were performed in TBS-1% BSA. All washes were performed with TBST. DAPI provided by the RNAscope Multiplex Fluorescent kit was used for nuclear counterstaining. Images were acquired with a Zeiss Axio Observer Z1 microscope with Apotome for optical sectioning using a 20x air objective. Full z-stacks were acquired for analysis. NSCs were identified based on SOX2 positivity and GFAP+ apical processes extending from the nucleus in 1  $\mu$ m z-stack images from  $n = 4$  mice. IPCs were similarly identified based on SOX2 positivity but without GFAP+ apical processes. mRNA puncta for all 3 genes were counted manually throughout the depth of cell nuclei and length of processes.

### **Statistics**

Statistics were performed as described in each figure legend.  $\chi^2$  contingency test was used to compare reproducibility of DEGs identified in cultured NSCs after oxidative stress by scRNAseq and 1 ng RNAseq. For the qRT-PCR confirmation of top DEGs by platform, a Mann-Whitney test was performed to identify individual genes with significant upregulation. To compare percentiles of gene expression by DEGs identified via scRNAseq versus 1 ng RNAseq, all genes with no expression in any cell were first eliminated. Averaged transcript counts for DEGs were compared to averaged transcript counts of all detected genes within each sequencing platform. Unpaired t-test was used to compare scRNAseq and 1ng RNAseq of our cultured NSCs. Dunnett's multiple comparison test was used to compare our data with data published in [Kang et al. 2018] and [Jin et al. 2021]. To compare coefficient of variation between samples with or without CLEAR filtering, a one-way ANOVA with the Kruskal-Wallis test followed by Dunn's multiple comparisons or unpaired t-tests were applied. NSC and IPC mean expression per cell of Slc5a3, Serpina3n, and Timp1 were compared using unpaired t-tests. All analyses were performed using Prism (v9.0; GraphPad Software) and  $p < 0.05$  was considered significant.

### **Conflict of Interest**

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

### **Author Contributions**

JKD: Execution of research and writing. JKD and EDK: Primary contributors to research design, analysis, and writing. EDK: Funding acquisition and supervision. LAW, XC, AT, AP, RB, PY: Contributed to design, execution, analysis, and writing of methods for RNAseq studies. ZT and OKC: Performed LFPI studies. SS and RR: Provided general support in execution of research.

### **Funding**

Funding provided by R00NS089938 to EDK from the NIH/NINDS, R50 CA211524 to PY from NIH/NCI, and P30CA016058 to the OSU Comprehensive Cancer Center Shared Genomics Resource from NIH/NCI.

### **Acknowledgments**

We thank all of the members of the Kirby lab for excellent discussions during the execution of this project as well as planning and writing of this work. We also thank the OSU Chronic Brain Injury Initiative (CBI), the OSUCCC Genomics Shared Resource, the Kokiko-Cochran lab, and the OSU animal care staff for their support.

## References

Adusumilli, V.S., Walker, T.L., Overall, R.W., Klatt, G.M., Zeidan, S.A., Zocher, S., Kirova, D.G., Ntitsias, K., Fischer, T.J., Sykes, A.M., Reinhardt, S., Dahl, A., Mansfeld, J., Rünker, A.E., Kempermann, G., 2021. ROS Dynamics Delineate Functional States of Hippocampal Neural Stem Cells and Link to Their Activity-Dependent Exit from Quiescence. *Cell Stem Cell* 28, 300-314.e6. <https://doi.org/10.1016/j.stem.2020.10.019>

Ahlmann-Eltze, C., Huber, W., 2021. Transformation and Preprocessing of Single-Cell RNA-Seq Data (preprint). *Bioinformatics*. <https://doi.org/10.1101/2021.06.24.449781>

Andronic, J., Shirakashi, R., Pickel, S.U., Westerling, K.M., Klein, T., Holm, T., Sauer, M., Sukhorukov, V.L., 2015. Hypotonic Activation of the Myo-Inositol Transporter SLC5A3 in HEK293 Cells Probed by Cell Volumetry, Confocal and Super-Resolution Microscopy. *PLOS ONE* 10, e0119990. <https://doi.org/10.1371/journal.pone.0119990>

Artegiani, B., Lyubimova, A., Muraro, M., van Es, J.H., van Oudenaarden, A., 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>

Arzalluz-Luque, Á., Devailly, G., Mantsoki, A., Joshi, A., 2017. Delineating biological and technical variance in single cell expression data. *Int. J. Biochem. Cell Biol.* 90, 161–166. <https://doi.org/10.1016/j.biocel.2017.07.006>

Baser, A., Skabkin, M., Kleber, S., Dang, Y., Gülcüler Balta, G.S., Kalamakis, G., Göpferich, M., Ibañez, D.C., Schefzik, R., Lopez, A.S., Bobadilla, E.L., Schultz, C., Fischer, B., Martin-Villalba, A., 2019. Onset of differentiation is post-transcriptionally controlled in adult neural stem cells. *Nature* 566, 100–104. <https://doi.org/10.1038/s41586-019-0888-x>

Bell, K.F.S., Al-Mubarak, B., Martel, M.-A., McKay, S., Wheelan, N., Hasel, P., Márkus, N.M., Baxter, P., Deighton, R.F., Serio, A., Bilican, B., Chowdhry, S., Meakin, P.J., Ashford, M.L.J., Wyllie, D.J.A., Scannevin, R.H., Chandran, S., Hayes, J.D., Hardingham, G.E., 2015. Neuronal development is promoted by weakened intrinsic antioxidant defences due to epigenetic repression of Nrf2. *Nat. Commun.* 6, 7066. <https://doi.org/10.1038/ncomms8066>

Berg, D.A., Su, Y., Jimenez-Cyrus, D., Patel, A., Huang, N., Morizet, D., Lee, S., Shah, R., Ringeling, F.R., Jain, R., Epstein, J.A., Wu, Q.-F., Canzar, S., Ming, G.-L., Song, H., Bond, A.M., 2019. A Common Embryonic Origin of Stem Cells Drives Developmental and Adult Neurogenesis. *Cell* 177, 654-668.e15. <https://doi.org/10.1016/j.cell.2019.02.010>

Bhargava, V., Head, S.R., Ordoukhalian, P., Mercola, M., Subramaniam, S., 2015. Technical Variations in Low-Input RNA-seq Methodologies. *Sci. Rep.* 4, 3678. <https://doi.org/10.1038/srep03678>

Borrett, M.J., Innes, B.T., Jeong, D., Tahmasian, N., Storer, M.A., Bader, G.D., Kaplan, D.R., Miller, F.D., 2020. Single-Cell Profiling Shows Murine Forebrain Neural Stem Cells Reacquire a Developmental State when Activated for Adult Neurogenesis. *Cell Rep.* 32, 108022. <https://doi.org/10.1016/j.celrep.2020.108022>

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

DeLuca, D.S., Levin, J.Z., Sivachenko, A., Fennell, T., Nazaire, M.-D., Williams, C., Reich, M., Winckler, W., Getz, G., 2012. RNA-SeQC: RNA-seq metrics for quality control and process optimization. *Bioinformatics* 28, 1530–1532. <https://doi.org/10.1093/bioinformatics/bts196>

Denninger, J.K., Chen, X., Turkoglu, A.M., Sarchet, P., Volk, A.R., Rieskamp, J.D., Yan, P., Kirby, E.D., 2020. Defining the adult hippocampal neural stem cell secretome: In vivo versus in vitro transcriptomic differences and their correlation to secreted protein levels. *Brain Res.* 1735, 146717. <https://doi.org/10.1016/j.brainres.2020.146717>

Denoth-Lippuner, A., Jessberger, S., 2021. Formation and integration of new neurons in the adult hippocampus. *Nat. Rev. Neurosci.* 22, 223–236. <https://doi.org/10.1038/s41583-021-00433-z>

Diril, M.K., Ratnacaram, C.K., Padmakumar, V.C., Du, T., Wasser, M., Coppola, V., Tessarollo, L., Kaldis, P., 2012. Cyclin-dependent kinase 1 (Cdk1) is essential for cell division and suppression of DNA re-replication but not for liver regeneration. *Proc. Natl. Acad. Sci.* 109, 3826–3831. <https://doi.org/10.1073/pnas.1115201109>

Dulken, B.W., Buckley, M.T., Navarro Negredo, P., Saligram, N., Cayrol, R., Leeman, D.S., George, B.M., Boutet, S.C., Hebestreit, K., Pluvnig, J.V., Wyss-Coray, T., Weissman, I.L., Vogel, H., Davis, M.M., Brunet, A., 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>

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

Finak, G., McDavid, A., Yajima, M., Deng, J., Gersuk, V., Shalek, A.K., Slichter, C.K., Miller, H.W., McElrath, M.J., Prlic, M., Linsley, P.S., Gottardo, R., 2015. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol.* 16, 278. <https://doi.org/10.1186/s13059-015-0844-5>

Gupta, A.K., Nadarajah, S., 2004. *Handbook of beta distribution and its applications*, 1st ed. CRC Press, Boca Raton.

Harris, L., Rigo, P., Stiehl, T., Gaber, Z.B., Austin, S.H.L., Masdeu, M. del M., Edwards, A., Urbán, N., Marciniak-Czochra, A., Guillemot, F., 2021. Coordinated changes in cellular behavior ensure the lifelong maintenance of the hippocampal stem cell population. *Cell Stem Cell* 28, 863–876.e6. <https://doi.org/10.1016/j.stem.2021.01.003>

Harrow, J., Denoeud, F., Frankish, A., Reymond, A., Chen, C.-K., Chrast, J., Lagarde, J., Gilbert, J.G., Storey, R., Swarbreck, D., Rossier, C., Ucla, C., Hubbard, T., Antonarakis, S.E., Guigo, R., 2006. GENCODE: producing a reference annotation for ENCODE. *Genome Biol.* 9.

Harrow, J., Frankish, A., Gonzalez, J.M., Tapanari, E., Diekhans, M., Kokocinski, F., Aken, B.L., Barrell, D., Zadissa, A., Searle, S., Barnes, I., Bignell, A., Boychenko, V., Hunt, T., Kay, M., Mukherjee, G., Rajan, J., Despacio-Reyes, G., Saunders, G., Steward, C., Harte, R., Lin, M., Howald, C., Tanzer, A., Derrien, T., Chrast, J., Walters, N., Balasubramanian, S., Pei, B., Tress, M., Rodriguez, J.M., Ezkurdia, I., van Baren, J., Brent, M., Haussler, D., Kellis, M., Valencia, A.,

Reymond, A., Gerstein, M., Guigo, R., Hubbard, T.J., 2012. GENCODE: The reference human genome annotation for The ENCODE Project. *Genome Res.* 22, 1760–1774. <https://doi.org/10.1101/gr.135350.111>

Hochgerner, H., Zeisel, A., Lönnberg, P., 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>

Jin, W.-N., Shi, K., He, W., Sun, J.-H., Van Kaer, L., Shi, F.-D., Liu, Q., 2021. Neuroblast senescence in the aged brain augments natural killer cell cytotoxicity leading to impaired neurogenesis and cognition. *Nat. Neurosci.* 24, 61–73. <https://doi.org/10.1038/s41593-020-00745-w>

Kang, H.M., Subramaniam, M., Targ, S., Nguyen, M., Maliskova, L., McCarthy, E., Wan, E., Wong, S., Byrnes, L., Lanata, C.M., Gate, R.E., Mostafavi, S., Marson, A., Zaitlen, N., Criswell, L.A., Ye, C.J., 2018. Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. *Nat. Biotechnol.* 36, 89–94. <https://doi.org/10.1038/nbt.4042>

Kawase, T., Ohki, R., Shibata, T., Tsutsumi, S., Kamimura, N., Inazawa, J., Ohta, T., Ichikawa, H., Aburatani, H., Tashiro, F., Taya, Y., 2009. PH Domain-Only Protein PHLDA3 Is a p53-Regulated Repressor of Akt. *Cell* 136, 535–550. <https://doi.org/10.1016/j.cell.2008.12.002>

Kim, D., Langmead, B., Salzberg, S.L., 2015. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. <https://doi.org/10.1038/nmeth.3317>

Kjell, J., Fischer-Sternjak, J., Thompson, A.J., Friess, C., Sticco, M.J., Salinas, F., Cox, J., Martinelli, D.C., Ninkovic, J., Franze, K., Schiller, H.B., Götz, M., 2020. Defining the Adult Neural Stem Cell Niche Proteome Identifies Key Regulators of Adult Neurogenesis. *Cell Stem Cell* 26, 277–293.e8. <https://doi.org/10.1016/j.stem.2020.01.002>

Knight, V.B., Serrano, E.E., 2018. Expression analysis of RNA sequencing data from human neural and glial cell lines depends on technical replication and normalization methods. *BMC Bioinformatics* 19, 412. <https://doi.org/10.1186/s12859-018-2382-0>

Kroll, K.W., Mokaram, N.E., Pelletier, A.R., Frankhouser, D.E., Westphal, M.S., Stump, P.A., Stump, C.L., Bundschuh, R., Blachly, J.S., Yan, P., 2014. Quality Control for RNA-Seq (QuaCRS): An Integrated Quality Control Pipeline. *Cancer Inform.* 13s3, CIN.S14022. <https://doi.org/10.4137/CIN.S14022>

Kuhn, R.M., Haussler, D., Kent, W.J., 2013. The UCSC genome browser and associated tools. *Brief. Bioinform.* 14, 144–161. <https://doi.org/10.1093/bib/bbs038>

Kulkarni, A., Anderson, A.G., Merullo, D.P., Konopka, G., 2019. Beyond bulk: a review of single cell transcriptomics methodologies and applications. *Curr. Opin. Biotechnol.* 58, 129–136. <https://doi.org/10.1016/j.copbio.2019.03.001>

Law, C.W., Chen, Y., Shi, W., Smyth, G.K., 2014. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* 15, R29. <https://doi.org/10.1186/gb-2014-15-2-r29>

Liao, Y., Smyth, G.K., Shi, W., 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930. <https://doi.org/10.1093/bioinformatics/btt656>

Liao, Y., Smyth, G.K., Shi, W., 2013. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res.* 41, e108–e108. <https://doi.org/10.1093/nar/gkt214>

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

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. <https://doi.org/10.1186/s13059-014-0550-8>

Marini, F., Binder, H., 2019. pcaExplorer: an R/Bioconductor package for interacting with RNA-seq principal components. *BMC Bioinformatics* 20, 331. <https://doi.org/10.1186/s12859-019-2879-1>

McAvoy, K.M., Sahay, A., 2017. Targeting Adult Neurogenesis to Optimize Hippocampal Circuits in Aging. *Neurotherapeutics* 14, 630–645. <https://doi.org/10.1007/s13311-017-0539-6>

Mignone, J.L., Kukekov, V., Chiang, A.-S., Steindler, D., Enikolopov, G., 2004. Neural stem and progenitor cells in nestin-GFP transgenic mice. *J. Comp. Neurol.* 469, 311–324. <https://doi.org/10.1002/cne.10964>

Miller, S.M., Sahay, A., 2019. Functions of adult-born neurons in hippocampal memory interference and indexing. *Nat. Neurosci.* 22, 1565–1575. <https://doi.org/10.1038/s41593-019-0484-2>

Mizrak, D., Levitin, H.M., Delgado, A.C., Crotet, V., Yuan, J., Chaker, Z., Silva-Vargas, V., Sims, P.A., Doetsch, F., 2019. Single-Cell Analysis of Regional Differences in Adult V-SVZ Neural Stem Cell Lineages. *Cell Rep.* 26, 394-406.e5. <https://doi.org/10.1016/j.celrep.2018.12.044>

Mou, T., Deng, W., Gu, F., Pawitan, Y., Vu, T.N., 2020. Reproducibility of Methods to Detect Differentially Expressed Genes from Single-Cell RNA Sequencing. *Front. Genet.* 10, 1331. <https://doi.org/10.3389/fgene.2019.01331>

Nielsen, C.F., Zhang, T., Barisic, M., Kalitsis, P., Hudson, D.F., 2020. Topoisomerase II $\alpha$  is essential for maintenance of mitotic chromosome structure. *Proc. Natl. Acad. Sci.* 117, 12131–12142. <https://doi.org/10.1073/pnas.2001760117>

O'Leary, N.A., Wright, M.W., Brister, J.R., Ciufo, S., Haddad, D., McVeigh, R., Rajput, B., Robbertse, B., Smith-White, B., Ako-Adjei, D., Astashyn, A., Badretdin, A., Bao, Y., Blinkova, O., Brover, V., Chetverin, V., Choi, J., Cox, E., Ermolaeva, O., Farrell, C.M., Goldfarb, T., Gupta, T., Haft, D., Hatcher, E., Hlavina, W., Joardar, V.S., Kodali, V.K., Li, W., Maglott, D., Masterson, P., McGarvey, K.M., Murphy, M.R., O'Neill, K., Pujar, S., Rangwala, S.H., Rausch, D., Riddick, L.D., Schoch, C., Shkeda, A., Storz, S.S., Sun, H., Thibaud-Nissen, F., Tolstoy, I., Tully, R.E., Vatsan, A.R., Wallin, C., Webb, D., Wu, W., Landrum, M.J., Kimchi, A., Tatusova, T., DiCuccio, M., Kitts, P., Murphy, T.D., Pruitt, K.D., 2016. Reference sequence (RefSeq) database at NCBI: current

status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* 44, D733–D745. <https://doi.org/10.1093/nar/gkv1189>

Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H.A., Trapnell, C., 2017. Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* 14, 979–982. <https://doi.org/10.1038/nmeth.4402>

Quinlan, A.R., Hall, I.M., 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. <https://doi.org/10.1093/bioinformatics/btq033>

Robinson, M.D., Oshlack, A., 2010. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 11, R25. <https://doi.org/10.1186/gb-2010-11-3-r25>

Schubert, M., Lindgreen, S., Orlando, L., 2016. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res. Notes* 9, 88. <https://doi.org/10.1186/s13104-016-1900-2>

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., 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>

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., Satija, R., 2019. Comprehensive Integration of Single-Cell Data. *Cell* 177, 1888–1902.e21. <https://doi.org/10.1016/j.cell.2019.05.031>

Sun, X., Kaufman, P.D., 2018. Ki-67: more than a proliferation marker. *Chromosoma* 127, 175–186. <https://doi.org/10.1007/s00412-018-0659-8>

Tapp, Z.M., Kumar, J.E., Witcher, K.G., Atluri, R.R., Velasquez, J.A., O'Neil, S.M., Dziabis, J.E., Bray, C.E., Sheridan, J.F., Godbout, J.P., Kokiko-Cochran, O.N., 2020. Sleep Disruption Exacerbates and Prolongs the Inflammatory Response to Traumatic Brain Injury. *J. Neurotrauma* 37, 1829–1843. <https://doi.org/10.1089/neu.2020.7010>

Trapnell, C., Hendrickson, D.G., Sauvageau, M., Goff, L., Rinn, J.L., Pachter, L., 2013. Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat. Biotechnol.* 31, 46–53. <https://doi.org/10.1038/nbt.2450>

Urbán, N., Blomfield, I.M., Guillemot, F., 2019. Quiescence of Adult Mammalian Neural Stem Cells: A Highly Regulated Rest. *Neuron* 104, 834–848. <https://doi.org/10.1016/j.neuron.2019.09.026>

Vicidomini, C., Guo, N., Sahay, A., 2020. Communication, Cross Talk, and Signal Integration in the Adult Hippocampal Neurogenic Niche. *Neuron* 105, 220–235. <https://doi.org/10.1016/j.neuron.2019.11.029>

Walker, L.A., Sovic, M.G., Chiang, C.-L., Hu, E., Denninger, J.K., Chen, X., Kirby, E.D., Byrd, J.C., Muthusamy, N., Bundschuh, R., Yan, P., 2020. CLEAR: coverage-based limiting-cell experiment analysis for RNA-seq. *J. Transl. Med.* 18, 63. <https://doi.org/10.1186/s12967-020-02247-6>

Wang, L., Wang, S., Li, W., 2012. RSeQC: quality control of RNA-seq experiments. *Bioinformatics* 28, 2184–2185. <https://doi.org/10.1093/bioinformatics/bts356>

Wang, T., Li, B., Nelson, C.E., 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>

Yuzwa, S.A., Borrett, M.J., Innes, B.T., Voronova, A., Ketela, T., Kaplan, D.R., Bader, G.D., Miller, F.D., 2017. Developmental Emergence of Adult Neural Stem Cells as Revealed by Single-Cell Transcriptional Profiling. *Cell Rep.* 21, 3970–3986. <https://doi.org/10.1016/j.celrep.2017.12.017>

Zeng, Z., Miao, N., Sun, T., 2018. Revealing cellular and molecular complexity of the central nervous system using single cell sequencing. *Stem Cell Res. Ther.* 9, 234. <https://doi.org/10.1186/s13287-018-0985-z>

Zheng, G.X.Y., Terry, J.M., Belgrader, P., Ryvkin, P., Bent, Z.W., Wilson, R., Ziraldo, S.B., Wheeler, T.D., McDermott, G.P., Zhu, J., Gregory, M.T., Shuga, J., Montesclaros, L., Underwood, J.G., Masquelier, D.A., Nishimura, S.Y., Schnall-Levin, M., Wyatt, P.W., Hindson, C.M., Bharadwaj, R., Wong, A., Ness, K.D., Beppu, L.W., Deeg, H.J., McFarland, C., Loeb, K.R., Valente, W.J., Ericson, N.G., Stevens, E.A., Radich, J.P., Mikkelsen, T.S., Hindson, B.J., Bielas, J.H., 2017. Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* 8, 14049. <https://doi.org/10.1038/ncomms14049>

Ziegenhain, C., Vieth, B., Parekh, S., Reinius, B., Guillaumet-Adkins, A., Smets, M., Leonhardt, H., Heyn, H., Hellmann, I., Enard, W., 2017. Comparative Analysis of Single-Cell RNA Sequencing Methods. *Mol. Cell* 65, 631-643.e4. <https://doi.org/10.1016/j.molcel.2017.01.023>

Zywitsa, V., Misios, A., Bunatyan, L., Willnow, T.E., 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>

## Data Availability Statement

RAW RNA-seq sequencing FASTQ files and processed counts tables will be deposited into the NCBI Gene Expression Omnibus (GEO) upon publication.