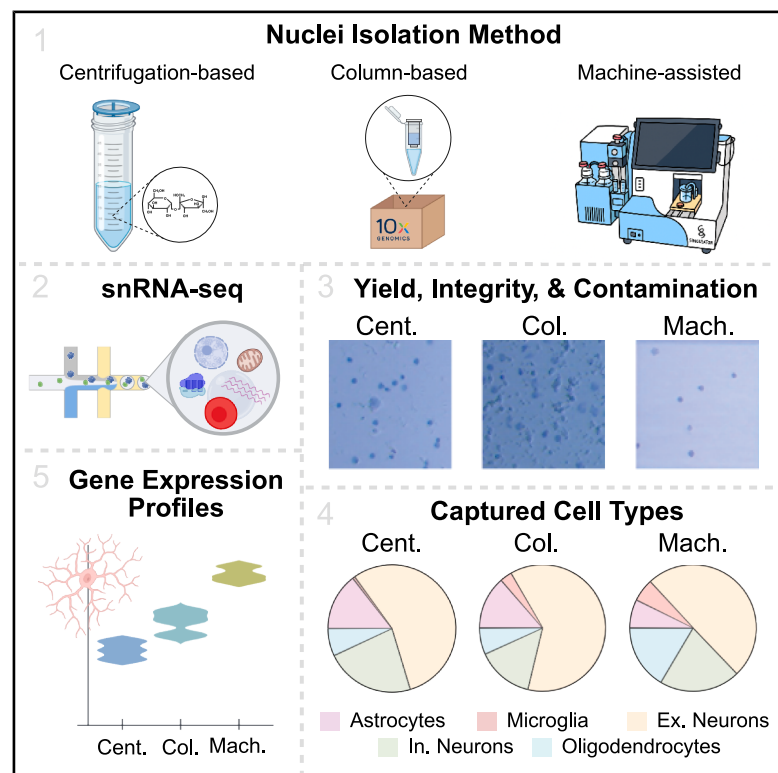


Comparative analysis of nuclei isolation methods for brain single-nucleus RNA sequencing

Graphical abstract



Authors

Holly N. Kersey, Dominic J. Acri,
 Luke C. Dabin, Kelly A. Hartigan,
 Richard Mustaklem, Jung Hyun Park,
 Jungsu Kim

Correspondence

jk123@iu.edu

In brief

Kersey et al. systematically evaluate three nuclei isolation methods for brain snRNA-seq, demonstrating that protocol choice markedly affects data quality metrics, including nuclei yield and ambient RNA contamination levels, as well as cell type proportions. Notably, a machine-assisted approach minimizes technical variability, providing consistent transcriptional signatures across glial and neuronal populations.

Highlights

- Comparison of three nuclei isolation methods for brain single-nucleus RNA sequencing
- Isolation protocol affects brain cell type proportions and transcriptional homogeneity
- Glial and neuronal gene expression patterns differ among isolation protocols
- Machine-assisted isolation yields the most intact nuclei with minimal contamination



Resource

Comparative analysis of nuclei isolation methods for brain single-nucleus RNA sequencing

Holly N. Kersey,^{1,2,5} Dominic J. Acri,^{1,2,5} Luke C. Dabin,^{2,3,5} Kelly A. Hartigan,^{1,2} Richard Mustaklem,^{2,3} Jung Hyun Park,^{2,4} and Jungsu Kim^{2,3,6,*}

¹Medical Neurosciences Graduate Program, Indiana University School of Medicine, Indianapolis, IN, USA

²Stark Neurosciences Research Institute, Indiana University School of Medicine, Indianapolis, IN, USA

³Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA

⁴Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN, USA

⁵These authors contributed equally

⁶Lead contact

*Correspondence: jk123@iu.edu

<https://doi.org/10.1016/j.crmeth.2026.101337>

MOTIVATION Existing nuclei isolation techniques for snRNA-seq vary in their ability to preserve nuclear integrity, minimize ambient RNA contamination, and optimize recovery rates. This poses a challenge for researchers in choosing the most suitable approach for their particular experimental requirements. To address this critical issue, our study directly compared three nuclei isolation methods and evaluated their performance in terms of yield, purity, and downstream sequencing quality. By providing a comprehensive assessment, we aim to guide researchers in selecting the most appropriate isolation protocol for their snRNA-seq experiments, ensuring optimal results and advancing the study of complex brain tissues at the single-nucleus level.

SUMMARY

Single-nucleus RNA sequencing (snRNA-seq) enables resolving cellular heterogeneity in complex tissues by using nuclei instead of cells, overcoming limitations of single-cell RNA sequencing and enabling analysis of frozen and hard-to-isolate tissues. Despite advances in isolation techniques, systematic evaluations of their effects on nuclear integrity and subsequent data quality remain lacking, a critical gap with profound implications for rigor and reproducibility. To address this, we compared three mechanistically distinct nuclei isolation strategies with brain tissue: a sucrose gradient centrifugation-based method, a spin column-based method, and a machine-assisted platform. All methods captured diverse cell types but revealed considerable protocol-dependent differences in cell type proportions, transcriptional homogeneity, and the preservation of cell-state-specific markers. Moreover, workflows differentially influenced contamination levels from ambient, mitochondrial, and ribosomal RNAs, with the machine-assisted method exhibiting the highest overall data quality. Our findings establish nuclei isolation methodology as a critical experimental variable shaping snRNA-seq data quality and biological interpretation.

INTRODUCTION

Single-nucleus RNA sequencing (snRNA-seq) is an effective approach for investigating gene expression at the single-cell level, particularly in complex tissues such as the brain.^{1–4} This technique addresses the key limitations of traditional single-cell RNA sequencing (scRNA-seq), which can become problematic when applied to brain tissue.^{5–7} For example, conventional cell dissociation protocols often compromise the integrity of neurons through mechanical shearing or induce the activation of glia.^{6,8–12} These factors contribute to cell type isolation bias, the potential loss of rare or fragile cell pop-

ulations, and *ex vivo* gene expression artifacts.¹³ snRNA-seq overcomes the need for intact cell dissociation, thereby minimizing dissociation-associated challenges and providing a more physiologically accurate representation of the brain's cellular heterogeneity.^{12,14,15} Additionally, the use of nuclei, rather than whole cells, enables the study of samples that were previously difficult to analyze via scRNA-seq, such as postmortem, frozen, or biobank tissue, including those obtained from individuals with neurodegenerative conditions such as Alzheimer's disease (AD).^{12,16} This advancement opens new avenues for investigating cellular and molecular changes in complex disorders, potentially leading to improved



understanding of disease mechanisms and the development of targeted therapies.¹⁷

Given the benefits of snRNA-seq, assessing and improving the quality of preclinical snRNA-seq datasets has become a growing research priority. Recent efforts in single-cell biology have yielded comprehensive atlas-level datasets, such as the Human Cell Atlas, BRAIN Initiative Cell Consensus Network, and Allen Brain Atlas.^{18–21} These and other single-cell initiatives have uncovered potential disease mechanisms and identified new drug targets.^{22–26} Funding agencies, such as the National Institutes of Health (NIH) and Chan Zuckerberg Initiative (CZI), have invested heavily in single-cell research to advance our understanding of human health, disease, and treatment.^{27–31} snRNA-seq can further enhance these efforts, offering unique advantages in terms of tissue compatibility. However, the potential of snRNA-seq in advancing single-cell genomics critically depends on the quality of nuclear preparations. Therefore, there is an urgent need to preserve high-quality nuclei preparations to ensure that valuable samples and substantial sequencing costs are not wasted due to technical artifacts or contamination by ambient RNAs that necessitate resequencing. Our study addresses this critical need by comparing nuclei isolation protocols to identify methods that best maintain nuclear integrity and purity throughout the snRNA-seq workflow.

Low-quality nuclei isolation protocols can introduce artifacts that compromise the proper data interpretation and experimental reliability.^{32,33} Contaminants, such as ambient RNA, ribosomal RNA, and mitochondrial RNA, can cause technical artifacts that mask biological effects. For example, ambient RNA from lysed cells, a ubiquitous issue of scRNA- and snRNA-seq datasets, can lead to false-positives in gene expression profiles, obscuring true cell type-specific markers.^{34,35} Excessive levels of contaminating mitochondrial and ribosomal RNA can overwhelm sequencing data, thereby reducing the depth of informative transcripts and hindering the detection of low-abundance genes.³⁶ Furthermore, such artifacts can significantly compromise various aspects of downstream analyses, including accurate cell type identification, differential expression analysis, and integration with other omics information. While computational tools can mitigate some variability in sample preparation, obtaining high-quality nuclei isolations before library preparation is essential for minimizing technical noise and ensuring reliable, reproducible data.^{37–40}

Isolating nuclei from complex tissues such as the brain poses challenges in preserving cellular diversity, minimizing RNA degradation, and reducing extranuclear contamination. To address these challenges, we compared three mechanistically distinct nuclei isolation protocols: (1) manual homogenization followed by sucrose gradient centrifugation, (2) a spin column-based method, and (3) a machine-assisted platform.^{41–43} Each nuclei isolation method for brain tissue processing has distinct advantages and limitations. The sucrose gradient centrifugation method is a well-established and cost-effective technique. However, this method can suffer from person-to-person variability in hand grinding and gradient preparation and may require ultracentrifugation. The column-based method provides comparable scalability without the need for specialized machinery, providing faster processing times than sucrose gradient preparation. How-

ever, this protocol still faces potential variability in tissue grinding and requires specific consumable columns. The machine-assisted platform provides an automated approach that reduces processing time and minimizes sample-to-sample and person-to-person variability. However, it requires the purchase of specialized equipment and specific consumable cartridges. The choice of method ultimately depends on factors such as available equipment, sample size, desired throughput, and tolerance for variability, requiring researchers to carefully consider these aspects when selecting a nuclei isolation approach for their specific experimental needs.

Using the mouse brain cortex, we assessed the efficacy of these methods by preparing nuclei suspensions, capturing RNA using 10x Genomics' Chromium, and sequencing snRNA-seq libraries. Given the heterogeneity of brain tissue, we sought to preserve various neuronal and glial populations while minimizing bias and artifacts. Each isolation method was evaluated for nuclei yield, contamination, and marker gene expression profiles. Our comparative analysis provides valuable insights into nuclei isolation strategies for brain tissues, enhancing our ability to study complex neurological processes and disorders at the single-nucleus level.

RESULTS

Nuclei yield and viability differ across isolation protocols

We evaluated three distinct nuclei isolation protocols optimized for snRNA-seq: manual homogenization followed by sucrose gradient centrifugation, a commercially available spin column-based method, and a machine-assisted platform (Figure 1A). All isolations were performed using cortical tissue from 6-month-old C57BL/6J mice ($N = 2$ biological replicates/method). Cortical tissue was dissected from the anterior portion of one hemisphere, and approximately 30 mg was weighed to normalize the input for each method. Representative bright-field microscopy images revealed clear differences in preparation quality among the methods. The centrifugation-based method produced defined individual nuclei with minimal background debris (Figure 1B). In contrast, the column-based method resulted in densely packed nuclei with notable aggregation and substantial debris contamination even after multiple optimization attempts (Figure 1C). The machine-assisted method yielded well-separated, intact nuclei with negligible debris (Figure 1D). Among the six samples selected for sequencing, the centrifugation-based and machine-assisted methods provided similar yields of 2 million nuclei or approximately 60,000 nuclei/mg of input (Figure 1E). In contrast, the column-based approach yielded 25% fewer nuclei with similar input material. Of the total number of nuclei recovered, the machine-assisted method maintained the structural integrity of almost 100% of the nuclei. This result indicates that the extracted nuclei remained intact upon recovery (Figure 1F). The centrifugation-based method preserved 85% of nuclei in their intact form, whereas the column-based method yielded only 35% structurally intact nuclei. These findings demonstrate that, compared with the column-based approach, the centrifugation-based and machine-assisted methods are more effective in terms of both nuclear yield and purity.

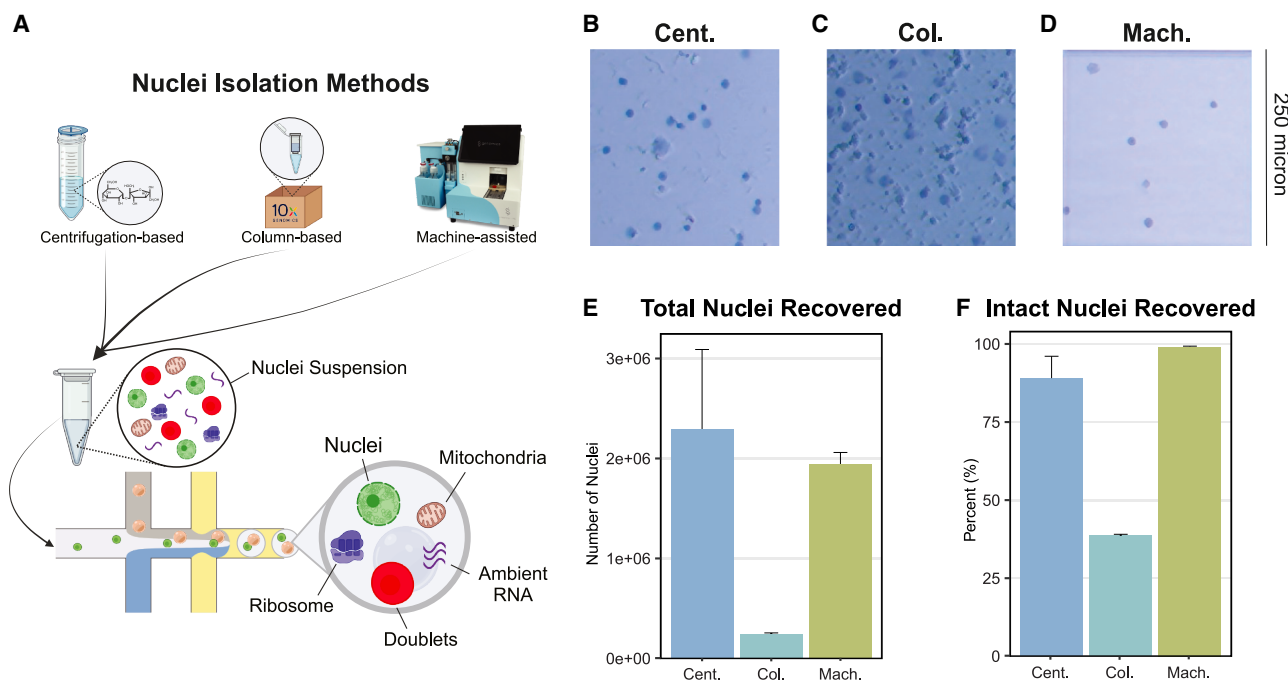


Figure 1. Nuclei yield and viability differ across isolation protocols

(A–D) (A) Graphical abstract depicting the workflow of each method and highlighting potential contaminants, including mitochondria, ribosomes, and ambient RNA, which can be captured in droplet-based sequencing depending on the isolation approach. Representative images of trypan blue-stained nuclei demonstrating the morphology and quality of nuclei obtained from the (B) centrifugation-based (Cent.), (C) column-based (Col.), and (D) machine-assisted (Mach.) isolation methods. All images were taken with the same microscope settings. Scale bar represents 250 μ m.

(E) The total number of nuclei recovered from each method, as quantified using the Countess Automated Cell Counter, reveals differences in yield among the protocols. Error bars represent standard deviation.

(F) From the total nuclei population isolation, the proportion of structurally intact nuclei recovered from each method. Error bars represent standard deviation.

Isolation technique influences the cell types captured in snRNA-seq

Next, we performed snRNA-seq experiments to evaluate the impact of the different isolation methods on the transcriptional profiles using 10x Genomics' Chromium single-cell 3' gene expression assay. In total, 98,452 nuclei were captured from the six samples, forming 36 distinct clusters (Figure 2A). The defined clusters included astrocytes (A), microglia (M), excitatory neurons (eN), inhibitory neurons (iN), oligodendrocytes (O), and other cell types (Ot). Clusters were annotated using the single-cell Mouse Cell Atlas (Table S1) and further refined with PanglaoDB.^{44,45} Key marker genes exhibited distinct expression patterns across cell types (Figure 2B). Notably, *Gja1* and *Slc1a3* were strongly expressed in astrocyte clusters, whereas *Sv2b* and *Slc17a7* were robustly expressed across excitatory neuron populations. *Gad1* and *Gad2* expressions were restricted to inhibitory neuron clusters. Across all methods, most of the nuclei captured were excitatory neurons (53.9%), followed by inhibitory neurons (17.2%) (Figure 2C). Notably, the column-based method captured the highest mean percentage of excitatory neurons (58.4%) but the lowest percentage of inhibitory neurons (13.6%) (Figure 2C). Glial cells and other cell types comprised smaller proportions of the total population, with mean percentages varying according to the isolation method. Interestingly, the centrifugation-based method captured the largest propor-

tion of astrocytes (13.9%), whereas the machine-assisted method attained the largest proportions of microglia (5.6%) and oligodendrocytes (15.9%). The uniformity of identified cell populations is a crucial quality control matrix for accurate cell type identification. To evaluate this, population homogeneity was assessed via the ratio of global unshifted entropy (ROGUE), a metric that quantifies the transcriptional consistency within cell populations (Figure 2D).⁴⁶ Most cell types exhibited moderate homogeneity, with ROGUE values between 0.5 and 0.75. In particular, astrocytes displayed elevated ROGUE values (>0.72) across all three methodologies, indicating highly uniform transcriptional profiles within this cell population. In all other cell types, both the machine-assisted and column-based methods produced comparable scores, between 0.65 and 0.8, whereas the centrifugation-based method yielded lower ROGUE values, <0.68, implying less uniformity within these populations. These data demonstrate that although all three nuclear isolation methods successfully capture diverse cell types from mouse cortical tissue, they exhibit differences in cell type proportions and transcriptional homogeneity.

Differences in quality-control metrics among isolation protocols

We evaluated the quality of nuclear preparations using standard metrics commonly employed in snRNA- and scRNA-seq

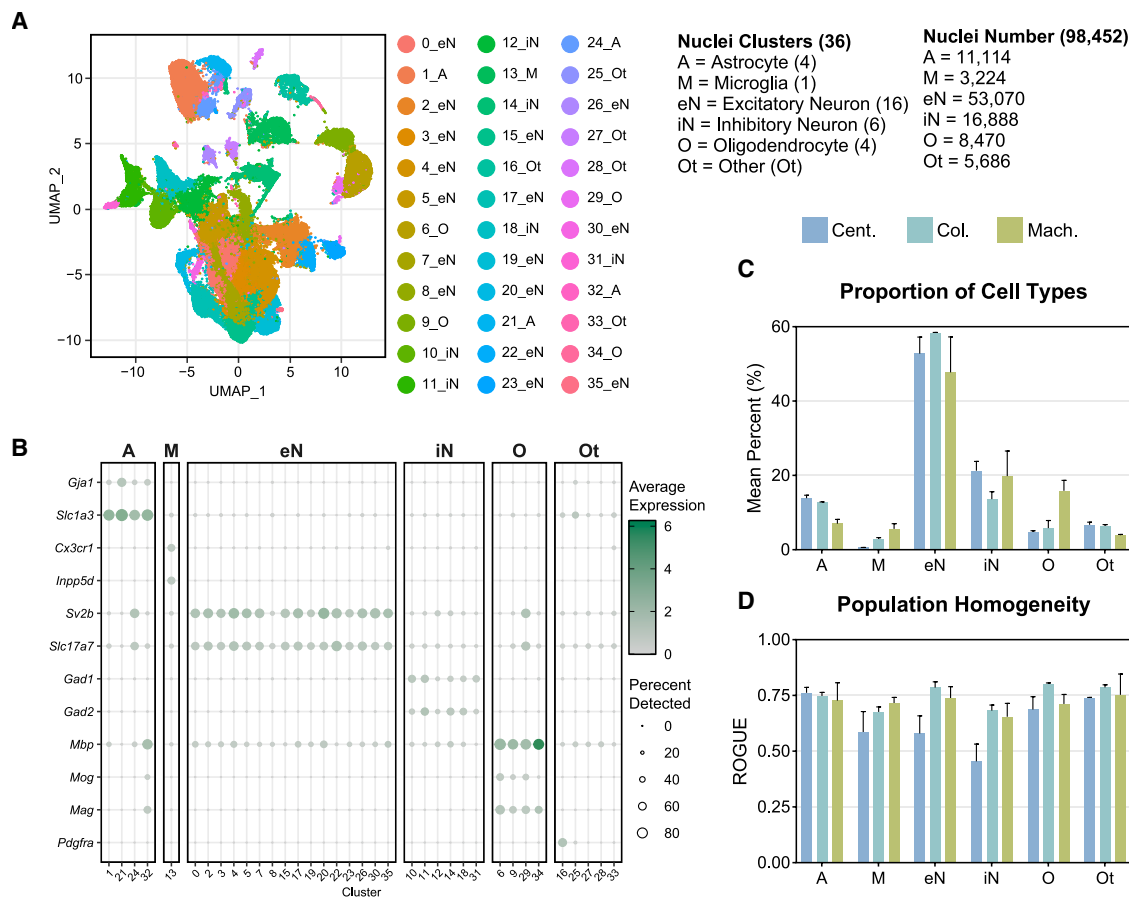


Figure 2. Isolation technique influences the cell types captured in snRNA-seq

(A) Uniform manifold approximation and projection (UMAP) plot of 98,452 nuclei including astrocyte (A), microglia (M), excitatory neurons (eN), inhibitory neurons (iN), oligodendrocytes (O), and other (Ot) clusters.

(B) Dot plot of key marker genes for each cell type. Average scaled expression level scales with white (low) to green (high) shading. The size of dots represents the percentage of nuclei in each cell type expressing the gene.

(C) The mean percent of each type of nuclei captured relative to the total population for each nuclei isolation method.

(D) The homogeneity of nuclei populations as reported by the ROGUE value, an entropy-based statistic that measures the uniformity of given populations, using ROGUE. Color represents the isolation method (centrifugation based, Cent.; column based, Col.; machine assisted, Mach.). Error bars represent standard deviation.

experiments. We predicted the presence of doublets, which are technical artifacts where mRNAs from two different nuclei are tagged with the same barcode and cannot be demultiplexed, using DoubletFinder. DoubletFinder uses a subset of a scRNA-seq dataset to generate artificial doublets, combines them with the remaining data, and employs principal-component analysis to measure the proportion of artificial nearest neighbors for each cell, thresholding these values to produce final doublet predictions.³⁸ The column-based method yielded inconsistent doublet rates among samples, with predicted percentages ranging from a low of 3.2% to a high of 5.7% (Figure 3A). In the context of snRNA-seq analysis, a consistent doublet rate is generally preferable as it indicates more reliable and reproducible sample processing. The remaining samples produced by the centrifugation-based and machine-assisted methods had comparable doublet rates, ranging between 4.5% and 5.5%. To quantify ambient RNA contamination, we utilized SoupX,

which calculates the contamination fraction parameter “rho.” This parameter estimates the proportion of unique molecular identifiers attributable to ambient RNA by analyzing empty droplets and determining cell-specific contamination levels.³⁷ This approach identifies and removes nuclei-free RNAs, which can confound the biological interpretation of single-nuclei transcriptomic data. Column-based samples presented significantly elevated rho values (>0.25), nearly 2-fold higher than those of both centrifugation-based and machine-assisted methods, indicating substantial ambient RNA contamination (Figure 3B). The machine-assisted method again displayed consistency, with rho values of approximately 0.14 for both samples. The samples produced via the centrifugation-based method presented some degree of intersample variability, with rho values between 0.07 and 0.17. Recent studies have suggested that *Malat1* (metastasis-associated lung adenocarcinoma transcript 1), a long noncoding nuclear RNA, can act as a quality metric

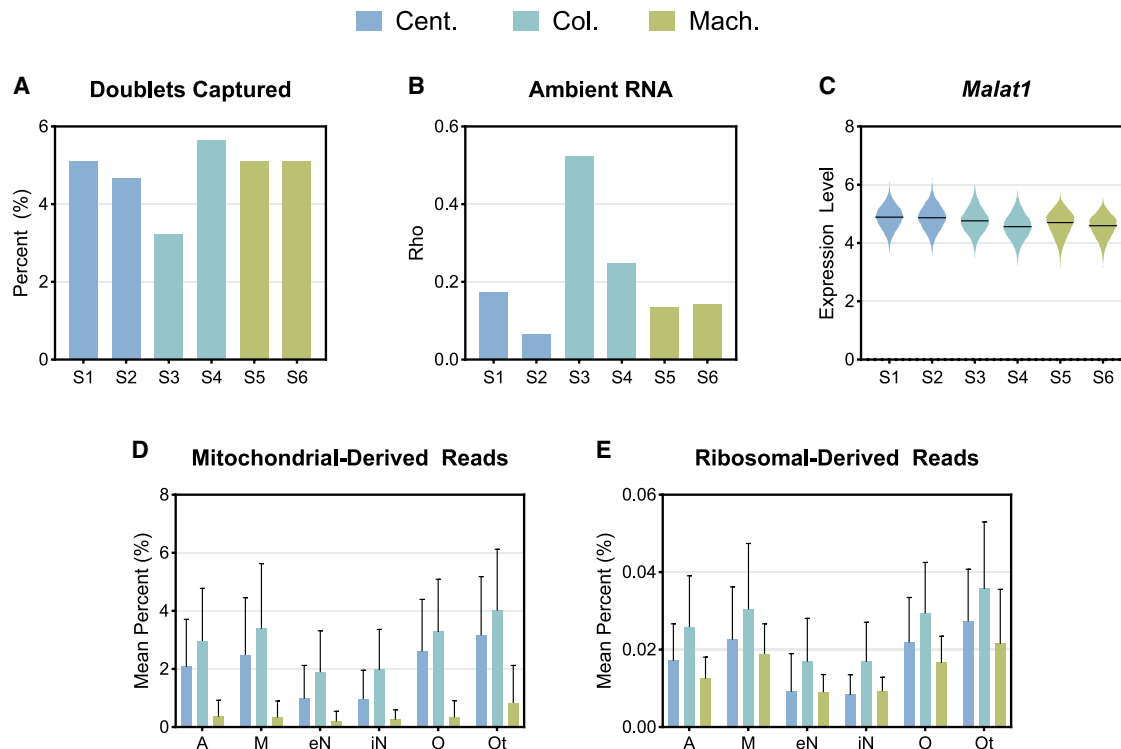


Figure 3. Differences in quality control metrics among isolation protocols

(A) The percentage of nuclei identified as doublets for each sample using DoubletFinder, representing the proportion of doublets relative to the total number of captured nuclei.

(B) The contamination fraction as reported by Rho value, an indicator of ambient RNA contamination, for each sample using SoupX.

(C) Violin plot of *Malat1*, a marker of nuclear integrity, expression for each sample. The line represents the median expression. The mean percentage of (D) mitochondrial-derived reads and (E) ribosomal-derived reads in each cell type per isolation method. Error bars represent standard deviation. Color represents the isolation method (centrifugation based, Cent.; column based, Col.; machine assisted, Mach.).

for snRNA-seq, with low levels identifying nuclei of poor quality.^{47,48} There was no discernable difference between *Malat1* expression, with all samples having median expression levels, the log-transformed RNA counts, of approximately 5 (Figure 3C).

High-quality nuclei preparations should contain minimal to no cytoplasmic content. In snRNA-seq, the presence of non-nuclear material can introduce bias in gene expression profiles and compromise the accuracy of downstream analyses. Therefore, we analyzed mitochondrial and ribosomal transcript levels as indicators of cytoplasmic contamination. Importantly, the machine-assisted method demonstrated low levels of average mitochondrial-derived reads, <0.5% of total reads, across all cell types (Figure 3D). The centrifugation- and column-based methods resulted in elevated average levels of mitochondrial-derived reads, ranging from 1% to 3% and 2% to 4%, respectively. Although ribosomal-derived reads remained minimal across all preparation methods, the machine-assisted and column-based protocols displayed the lowest contamination, with values <0.03% (Figure 3E). The column-based method demonstrated the highest levels of contamination, with ribosomal-derived reads exhibiting an average contamination level of 0.036%. The results indicate that the machine-assisted method yielded nuclei preparations with high consistency between samples and low contamination levels, whereas the column-based

and centrifugation-based methods presented greater sample-to-sample variability and elevated contamination levels.

To provide a relevant comparison to our findings, we assessed these quality metrics across three additional datasets, two publicly available column-based samples (10x Genomics) and four samples each internally prepared using the centrifugation-based and machine-assisted methods (Figures S1A–S1L). Notably, metrics from the centrifugation- and column-based methods remained variable across samples, whereas the machine-assisted method produced very consistent results for all metrics. Samples processed with the machine-assisted method averaged the lowest levels of ambient RNA contamination (Rho <0.08) compared to the centrifugation- and column-based methods (Figures S1A, S1E, and S1I). Again, *Malat1* expression was relatively similar across all samples (Figures S1B, S1F, and S1J). The centrifugation-based method demonstrated the highest levels of mitochondrial-derived reads but the lowest levels of ribosomal-derived reads (Figures S1C and S1D). The column-based and machine-assisted methods had comparable percentages of mitochondrial-derived reads, < ~0.05% (Figures S1G and S1K). However, the column-based method displayed the highest level of ribosomal-derived reads (>0.5%) (Figure S1H and S1L). These comparisons reinforce the consistency of our

findings while highlighting the reduced variability of the machine-assisted method.

Method-specific variations in glial gene expression

The accurate identification and characterization of specific cell types, particularly glial cells, is crucial for understanding brain function and pathology in single-cell genomic studies. Different nuclei isolation methods may introduce biases or artifacts that affect the detection of cell type-specific markers, potentially leading to misclassification or incomplete representation of cellular populations. Therefore, we analyzed gene expression profiles across the three isolation methods, with a focus on glial cell types. We compared markers for the microglia population identified by each isolation method with various glial cell type annotations from the Panglao database (Table S2).⁴⁵ To address the specificity of these markers, we assessed the proportion of identified microglia markers by each method that corresponds to different glial cell types. Notably, both the centrifugation-based and machine-assisted methods resulted in significant enrichment ($p_{\text{adj}} < 0.05$) of microglial markers in the Panglao database, with the machine-assisted method yielding the highest proportion of microglia-specific markers (Figure 4A). Further investigation of canonical microglia-specific markers revealed method-specific variations in expression patterns. For example, *C1qa* exhibited the highest expression with the machine-assisted method (Figure 4B). Two other microglia-specific markers, *Hexb* and *Siglech*, showed robust expression across all methods, but expression was greater in the centrifugation-based and machine-assisted methods (Figures 4C and 4D). Interestingly, *Tmem119* expression was elevated specifically in the machine-assisted method (Figure 4E). Despite the enrichment of microglia-specific markers in microglia nuclei, some of them may express markers for other glial cell types, such as oligodendrocytes and satellite glial cells. When examining oligodendrocyte markers within the microglial population, the machine-assisted method showed negligible expression of *Plp1*, whereas both the centrifugation- and column-based methods exhibited some level of *Plp1* expression (Figure 4G). Only the centrifugation-based method demonstrated the expression of *Mbp* (Figure 4H). Similarly, analysis of satellite glial markers within microglia nuclei revealed the highest expression of *Glu1* and *Ptgds* with the centrifugation-based method (Figures 4I and 4J). Taken together, our findings highlight critical method-specific strengths and limitations. For example, machine-assisted isolation maximizes microglia-specific signals while minimizing oligodendrocyte marker contamination. The observed technical variability underscores the importance of isolation method selection in glial cell research. Method-driven differences in marker detection could significantly impact interpretations of cellular identity and function, particularly in studies comparing glial subpopulations or analyzing rare cell types.

Several key glial genes in our analysis have been implicated in various neurodegenerative conditions. For instance, *Tmem119* expression is significantly reduced in activated microglia in AD, whereas *C1qa* is associated with microglial activation after traumatic brain injury.^{49–51} Also, a decrease in astrocytic *Glu1* expression has been associated with epilepsy and subsequent neurodegeneration.^{52,53} Therefore, we explored gene signatures

across various cell states, with UCell summary scoring strategy. This approach uses rank-based enrichment to calculate the relative expression of predefined gene sets within individual cells or nuclei (Table S3).⁵⁴ Compared with the centrifugation-based method, the column-based and machine-assisted methods yielded similar higher signature scores (approximately 0.3) for homeostatic markers in astrocytes (Figure 4J). The disease-associated astrocyte signature generally had low scores, with median scores <0.2 (Figure 4K). In microglia, the machine-assisted method had the highest scores for homeostatic signatures and the lowest scores for disease-associated signatures (Figures 4L and 4M). When homeostatic oligodendrocytes were examined, the column-based and machine-assisted methods demonstrated the highest signature scores, >0.3 , with a notable reduction in the centrifugation-based method score, <0.2 (Figure 4N). Taken together, our comprehensive analyses demonstrate that the choice of nuclear isolation methods significantly affects gene expression patterns and cellular state signatures. Each method varies in its ability to preserve cell type-specific markers and homeostatic states across astrocytes, microglia, and oligodendrocytes, highlighting the importance of method selection in glial cell research.

It is worth noting that previous studies have demonstrated the effects of *ex vivo* microglial activation during single-cell tissue dissociation.¹⁰ Although the isolation methods between nuclei and cells are considerably different, we investigated if we saw similar signatures in our dataset. We performed gene set enrichment analysis (GSEA) with this activated signature, as well as gene sets for microglia identity, disease-associated microglia (DAM), and homeostatic microglia for all nuclei isolation methods (Figures S2A–S2C). The activated microglia gene set consistently demonstrated the lowest enrichment scores (normalized enrichment score: 0.93–1.08) with non-significant adjusted p values. This indicates there is no evidence that the nuclei isolation methods used in this study are preferentially inducing the previously described *ex vivo* microglia activation signature.

Protocol-dependent differences in neuronal subclass capture

Given that the majority of captured nuclei were neurons and that capture efficiency differed among isolation methods, we next analyzed the subclasses of neurons represented. Comparable to our previous microglia analysis, we compared the markers for each neuron cluster with neuron subclass annotations from the Allen Brain Atlas isocortex dataset (Figure 5A). Most clusters showed significant enrichment ($p_{\text{adj}} < 0.05$) of a specific neuron class. To elucidate clusters that did not show significant enrichment, we correlated the top 2,000 variable features of each cluster with the Allen Brain Atlas annotations (Figure 5B). Combining these analyses, we were able to classify each neuron cluster as intratelencephalic-extratelen- cephalic glutamatergic neurons (01 IT-ET Glut), near-projecting corticothalamic glutamatergic neurons (02 NP-CT-L6b Glut), caudal ganglionic eminence gabaergic neurons (06 CTX-CGE GABA), or medial ganglionic eminence gabaergic neurons (07 CTX-MGE GABA). We then compared the proportion of these neuron subclasses captured by each method. Interestingly, in the excitatory neuron population, the

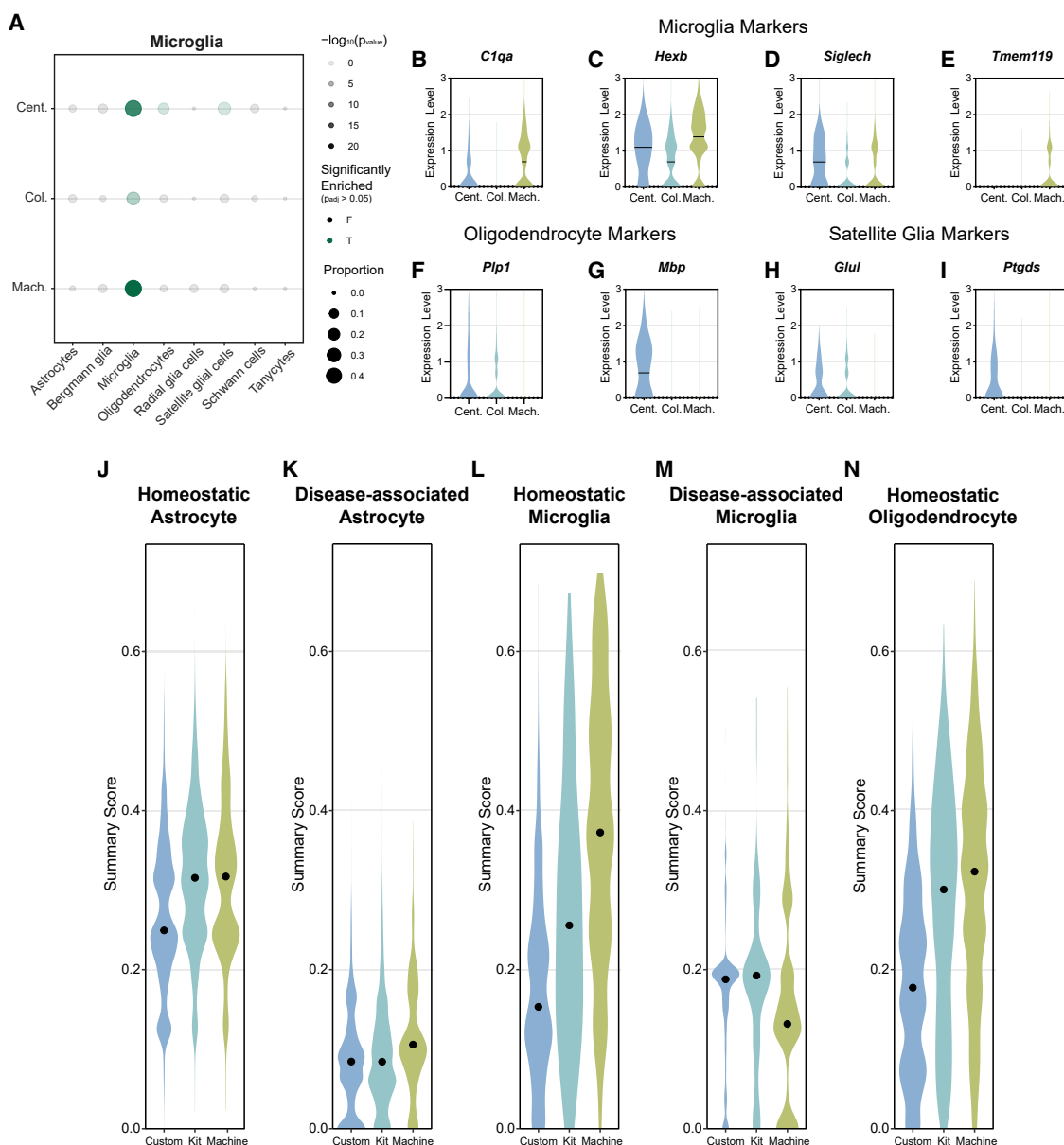


Figure 4. Method-specific variations in glial gene expression

(A) Dotplot of identified microglia marker gene expression compared to glial cell type annotations in the Panglao database. The size of the dots represents the proportion of microglia marker genes expressed within each glial cell type. The shading of dots represents the $-\log_{10}(p_{\text{value}})$ of the marker gene expression in the glial cell type, and the color of dots represents if the expression is significant (green) or not significant (black) determined by Fisher's exact test. Violin plots of expression level of canonical microglia markers (B) *C1qa*, (C) *Hexb*, (D) *Siglech*, and (E) *Tmem119*, as well as oligodendrocyte markers (F) *Plp1* and (G) *Mbp* and satellite glia markers (H) *Glul* and (I) *Ptgs* within the identified microglia population. The line represents the median expression level. UCell summary scoring of cell states, including (J) homeostatic astrocyte, (K) disease-associated astrocyte, (L) homeostatic microglia, (M) disease-associated microglia, and (N) homeostatic oligodendrocyte, within their respective cell type population. Dots represent the median summary score. Color represents the isolation method (centrifugation based, Cent.; column based, Col.; machine assisted, Mach.).

column-based and machine-assisted methods captured similar mean percentages of 01 IT-ET Glut neurons, 59% and 57%, respectively (Figure 5C). However, the column-based method captured the lowest percentage of 02 NP-CT-L6b Glut neurons, ~18% (Figure 5D). Within inhibitory neuron population, the machine-assisted method captured the high-

est proportion of 06 CTX-CGE GABA neurons (~6.5%) (Figure 5E), but the lowest percentage of 07 CTX-MGE neurons (~10.6%) (Figure 5F). The centrifugation- and column-based methods demonstrated similar proportions of both GABA neuron populations (Figures 5E and 5F). Collectively, these results indicate that nuclear isolation method influences

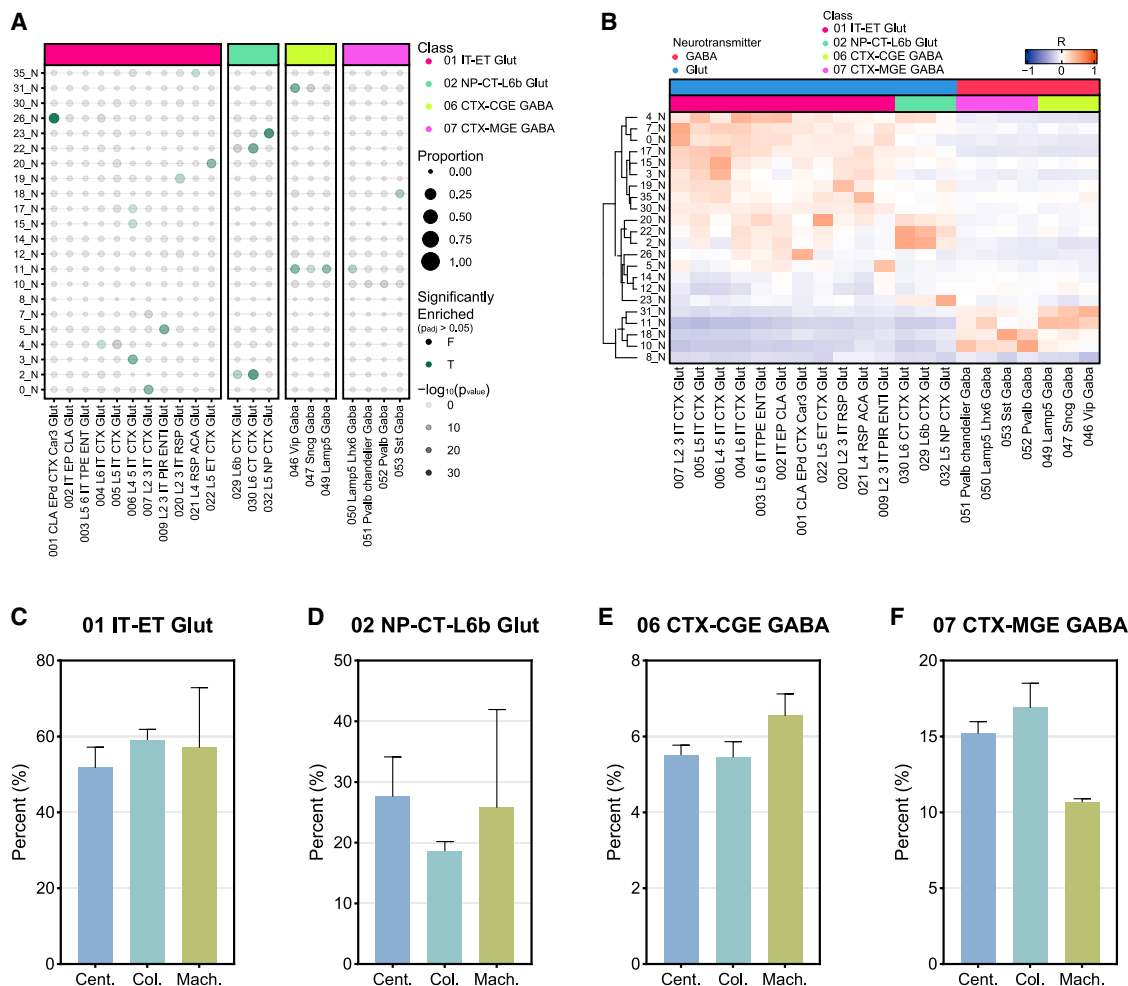


Figure 5. Protocol-dependent differences in neuronal subclass capture

(A) Dotplot of neuronal cluster gene expression compared to neuron subclass annotations from the Allen Brain Atlas (ABA) isocortex database. The size of the dots represents the proportion of marker genes expressed within each subclass. The shading of the dots represents the $-\log_{10}(p_{\text{value}})$ of the marker gene expression in the subclass. The color of the dots represents if the expression is significant (green) or not significant (black) determined by Fisher's exact test. (B) Heatmap of Pearson's correlation coefficients between top 2,000 variable features and ABA annotations. Proportion plots of neurons identified as (C) 01 IT-ET Glut, (D) 02 NP-CT-L6b Glut, (E) 06 CTX-CGE GABA, and (F) 07 CTX-MGE GABA for each isolation method. Error bars represent standard deviation.

neuronal subclass representation, with specific capture biases evident among both excitatory and inhibitory neuron populations.

DISCUSSION

Selecting an appropriate nuclei isolation method is crucial for ensuring reliable and reproducible snRNA-seq data, especially for brain tissue.⁶ The heterogeneous brain composition, comprising various neuronal subtypes, glia, and other specialized cells, presents unique challenges for accurate transcriptomic profiling. Moreover, the risk of isolation-induced artifacts during sample preparation is a significant concern, as the gene expression of microglia and other cell types can be rapidly altered in response to environmental factors.¹⁰ While researchers often select isolation protocols based on ease of use, sample availability, or throughput, there have been few sys-

tematic evaluations of how these choices impact data quality, cell type representation, and gene expression. Differences in nuclei isolation methods can introduce technical artifacts, leading to inconsistencies in data interpretation and limiting cross-study comparisons. Additionally, technical variability between samples can introduce confounding factors, making it difficult to discern true biological differences from methodological artifacts.⁵⁵ These issues also have implications for statistical power because increased technical variability may necessitate larger sample sizes to detect meaningful biological differences. To address these critical issues, we performed a comparative analysis of different isolation protocols and found significant differences in yield, quality, and transcriptional fidelity, all of which have direct implications for downstream analyses. With the growing trend toward expansive, collaborative, and data-driven research in the field, it has become crucial to comprehend the methodological variations

among datasets. This understanding is vital for maintaining reproducibility and ensuring biological accuracy, especially as researchers increasingly rely on large-scale analyses, integrative studies, and published datasets.⁵⁶

Our study highlights substantial variability in transcriptional profiles across different nuclei isolation methods. The centrifugation-based method not only yielded the greatest number of nuclei but also exhibited pronounced transcriptional heterogeneity, likely due to differences in RNA degradation or cytoplasmic contamination between individual preparations (Figures 1 and 2). This suggests that, while maximizing yield, the centrifugation-based method may compromise data integrity by introducing technical noise in a sample-by-sample manner. In contrast, the machine-assisted method resulted in the highest nuclei viability and the lowest levels of mitochondrial and ribosomal RNA contamination (Figures 1 and 3). These advantages are likely due to its automated processing, which reduces the handling time and minimizes the number of technical artifacts. Additionally, the machine-assisted method demonstrated greater consistency across samples in key quality control metrics, including doublet formation and ambient RNA contamination, whereas the centrifugation- and column-based methods exhibited high variability. Such variability can introduce batch effects, leading to inconsistencies in downstream analyses and unreliable biological conclusions.^{55,57} The column-based method, while yielding the lowest number of nuclei, also resulted in the highest levels of ambient RNA contamination. This contamination may distort transcriptomic profiles, leading to the under-representation of specific cell types or misclassification of cellular states.^{34,37} Given the importance of accurately identifying distinct cellular populations in neurodegenerative diseases such as AD, these limitations are particularly concerning. For example, misidentifying DAM, reactive astrocytes, or rare cell subpopulations can hinder efforts to characterize disease progression and cellular dysfunction in AD.⁵⁸ This, in turn, could lead to wasted effort and funding spent pursuing hypotheses based on unreliable data.

Many cell types, such as microglia, are especially sensitive to dissociation-induced activation, which is a known challenge in single-cell studies.^{9,10,59,60} When microglial populations were examined, the machine-assisted method preserved the strongest expression of homeostatic microglial markers and the most accurate representation of their transcriptional signatures (Figures 4 and S2). Because the mice used in this study lacked disease pathology, we expected minimal expression of DAM-related genes. However, both the centrifugation- and column-based methods resulted in elevated DAM signatures, suggesting that technical artifacts, rather than true biological signals, may drive these transcriptional changes. Such artifacts can lead to erroneous classification of microglial states, potentially confounding our understanding of neuroinflammatory processes in AD and other neurodegenerative diseases. Accurate classification of homeostatic versus reactive microglia is critical for studying neurodegeneration, as misinterpretation of transcriptional profiles may obscure disease-specific changes and misdirect therapeutic development.⁶¹ Our profiling experiments emphasize the importance of selecting an isolation method that

minimizes technical distortions to ensure biologically meaningful conclusions.

These findings highlight that nuclear isolation methods may introduce technical artifacts that complicate direct comparisons between studies using different approaches. Each method significantly affects data quality, cell type representation, and transcriptional integrity. The observed variability in centrifugation- and column-based methods highlights the potential risks of relying on high-yield or commercial protocols without fully considering their impact on downstream analyses. This variability not only compromises data quality, cell type representation, and transcriptional integrity but also underscores the need for methodological transparency and rigorous validation. By demonstrating how method selection can alter transcriptional signatures, our study emphasizes the importance of adhering to FAIR guidelines, i.e., making data findable, accessible, interoperable, and reusable (FAIR Principles).⁶² By ensuring that data are both interoperable and reusable, researchers can more easily replicate and validate each other's findings, addressing the reproducibility crisis in scientific research. The ability to integrate and reuse data from diverse sources has accelerated the discovery of new biological insights, ultimately advancing our understanding of cellular diversity and disease mechanisms. The reliability of the machine-assisted method in maintaining low cytoplasmic contamination across samples provides a model for optimizing high-throughput and large-scale studies. In contrast, the inconsistency observed in other methods highlights the need for protocol refinement to ensure accurate biological representation. By demonstrating that method selection can alter transcriptional signatures, our study emphasizes the necessity of methodological transparency and careful validation when datasets from different studies and conditions are compared. Future research should focus on refining isolation strategies across different brain regions, species, and pathological conditions to ensure the most accurate representation of cellular diversity and disease mechanisms at the single-cell level. Optimizing nuclei isolation protocols can improve the accuracy and reproducibility of snRNA-seq studies, ultimately advancing our understanding of cell diversity and disease mechanisms at the single-cell level.

Limitations of the study

The effectiveness of a nuclear isolation method for snRNA-seq depends on several factors, including tissue type, processing conditions, and specific research questions. Our study focuses on mouse cortical tissue under normal physiological conditions; thus, the results may differ across other brain regions, tissue types, or species. Different pathological states may also affect isolation efficacy and quality. To ensure consistent comparisons across the tested isolation methods, ancillary protocol choices, such as the nuclei-counting technique used for sample loading, were kept constant. It is possible that groups using alternative experimental workflows and reagents would encounter slightly different results. Six samples were sequenced via three methods to observe the methodological variation. Nuclei preparations were repeated multiple times, in some cases dozens of times, to ensure that the sequenced samples were representative of the expected isolation. However, we have not employed a

prospective statistical power analysis to guide the number of replicates, and statistical power in single-cell experiments remains a topic of active debate in the field. Despite these considerations, the findings of this study provide a solid foundation for method choice and refinement in neuroscience and single-cell genomics.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Jungsuk Kim (jk123@iu.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- The snRNA-seq data generated in this study have been deposited in the GEO database: GSE290858 and are publicly available as of the date of publication.
- The code for the snRNA-seq analysis in this study has been deposited at <https://github.com/jungsukimlab/NucleiIsolation>. An archival version is available at doi:[10.5281/zenodo.18378837](https://doi.org/10.5281/zenodo.18378837).
- Any additional information required to reanalyze the data reported in this study is available from the [lead contact](#) upon request.

ACKNOWLEDGMENTS

This work was supported by National Institutes of Health grants R01 AG077829, R01 AG071281, RF1 AG074543, R21 AG072738, and U01AG076804 (J.K.); T32AG071444 (to D.J.A.); and the Eli Lilly Stark Neuroscience Fellowship (to L.C.D.). The Kim laboratory was also supported by the Strategic Research Initiative (Indiana University), Indiana University Precision Health Initiative, Indiana University Pervasive Technology Institute (supported in part by Lilly Endowment, Inc.), and Shared University Research grants from IBM, Inc., to Indiana University. Single-nucleus RNA libraries were sequenced at the Center for Medical Genomics at Indiana University School of Medicine, which is partially supported by the Indiana University Grand Challenges Precision Health Initiative and the Indiana Genomic Initiative at Indiana University (INGEN); INGEN is supported in part by the Lilly Endowment. The graphical abstract and [Figure 1A](#) were created via paid subscriptions to [BioRender.com](#).

AUTHOR CONTRIBUTIONS

D.J.A., L.C.D., and J.K.: conceptualization and study design; D.J.A. and L.C.D.: data acquisition and curation; D.J.A., H.N.K., K.A.H., R.M., and J.H.P.: data analysis; H.N.K.: data interpretation and manuscript draft; D.J.A., L.C.D., H.N.K., and J.K.: writing, reviewing, and editing the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

This manuscript was developed with the assistance of generative artificial intelligence (AI) tools for writing and editing. Specifically, ChatGPT (4o),⁶³ Gemini (1.5 Flash),⁶⁴ and Perplexity⁶⁵ were employed to enhance the clarity and coherence of the text, refine language usage, and annotate code originally drafted by the authors. Generative AI tools were not used to find citations, perform literature reviews, or draft new analyses/interpretations of data. We, the authors, confirm that all scientific ideas, analyses, and conclusions represent our own original work.

STAR★METHODS

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

- [KEY RESOURCES TABLE](#)
- [EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS](#)
 - Animals
- [METHOD DETAILS](#)
 - Tissue collection
 - Centrifugation-based nucleus isolation
 - Column-based isolation
 - Machine-assisted nuclei isolation
 - Library preparation
- [QUANTIFICATION AND STATISTICAL ANALYSIS](#)
 - snRNA-seq data analysis
 - Statistical procedures

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.crmeth.2026.101337>.

Received: March 23, 2025

Revised: November 15, 2025

Accepted: February 5, 2026

Published: March 23, 2026

REFERENCES

1. Macnair, W., Calini, D., Agirre, E., Bryois, J., Jäkel, S., Smith, R.S., Kulkarni, P., Stokar-Regenscheit, N., Ott, V., Foo, L.C., et al. (2025). snRNA-seq stratifies multiple sclerosis patients into distinct white matter glial responses. *Neuron* 113, 396–410.e9. <https://doi.org/10.1016/j.neuron.2024.11.016>.
2. Zhang, L., Yang, Q., Yuan, R., Li, M., Lv, M., Zhang, L., Xie, X., Liang, W., and Chen, X. (2023). Single-nucleus transcriptomic mapping of blast-induced traumatic brain injury in mice hippocampus. *Sci. Data* 10, 638. <https://doi.org/10.1038/s41597-023-02552-x>.
3. Soelter, T.M., Howton, T.C., Clark, A.D., Oza, V.H., and Lasseigne, B.N. (2024). Altered Glia-Neuron Communication in Alzheimer's Disease Affects WNT, p53, and NFκB Signaling Determined by snRNA-seq. *Cell Commun Signal* 22, 317. <https://doi.org/10.1186/s12964-024-01686-8>.
4. Ma, J.Q., Wang, L., Zhang, Y., Bian, Y.Q., Qu, X.P., Song, L.J., Wang, C., Gao, L., Fang, Q.X., Zhao, D.C., et al. (2024). Single-nucleus RNA sequencing-based construction of a hippocampal neuron atlas in mice with epileptic cognitive impairment. *iScience* 27, 111065. <https://doi.org/10.1016/j.isci.2024.111065>.
5. Bakken, T.E., Hodge, R.D., Miller, J.A., Yao, Z., Nguyen, T.N., Aebermann, 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>.
6. Waag, R., and Bohacek, J. (2023). Single-Nucleus RNA-Sequencing in Brain Tissue. *Curr. Protoc.* 3, e919. <https://doi.org/10.1002/cpz1.919>.
7. 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>.
8. Lacar, B., Linker, S.B., Jaeger, B.N., Krishnaswami, S.R., Barron, J.J., Kelder, M.J.E., Parylak, S.L., Paquola, A.C.M., Venepally, P., Novotny, M., et al. (2016). Nuclear RNA-seq of single neurons reveals molecular signatures of activation. *Nat. Commun.* 7, 11022. <https://doi.org/10.1038/ncomms11022>.

9. van den Brink, S.C., Sage, F., Vértessy, 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>.
10. 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>.
11. 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>.
12. Kim, N., Kang, H., Jo, A., Yoo, S.A., and Lee, H.O. (2023). Perspectives on single-nucleus RNA sequencing in different cell types and tissues. *J. Pathol. Transl. Med.* 57, 52–59. <https://doi.org/10.4132/jptm.2022.12.19>.
13. Denisenko, E., Guo, B.B., Jones, M., Hou, R., de Kock, L., Lassmann, T., Poppe, D., Clément, O., Simmons, R.K., Lister, R., and Forrest, A.R.R. (2020). Systematic assessment of tissue dissociation and storage biases in single-cell and single-nucleus RNA-seq workflows. *Genome Biol.* 21, 130. <https://doi.org/10.1186/s13059-020-02048-6>.
14. Lake, B.B., Ai, R., Kaeser, G.E., Salathia, N.S., Yung, Y.C., Liu, R., Wildberg, A., Gao, D., Fung, H.L., Chen, S., et al. (2016). Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science* 352, 1586–1590. <https://doi.org/10.1126/science.aaf1204>.
15. Tran, M.N., Maynard, K.R., Spangler, A., Huuki, L.A., Montgomery, K.D., Sadasiviah, 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>.
16. Maitra, M., Nagy, C., Chawla, A., Wang, Y.C., Nascimento, C., Suderman, M., Thérout, J.F., Mechawar, N., Ragoussis, J., and Turecki, G. (2021). Extraction of nuclei from archived postmortem tissues for single-nucleus sequencing applications. *Arch. Protoc.* 16, 2788–2801. <https://doi.org/10.1038/s41596-021-00514-4>.
17. Brase, L., You, S.F., D'Oliveira Albanus, R., Del-Aguila, J.L., Dai, Y., Novotny, B.C., Soriano-Tarraga, C., Dykstra, T., Fernandez, M.V., Budde, J.P., et al. (2023). Single-nucleus RNA-sequencing of autosomal dominant Alzheimer disease and risk variant carriers. *Nat. Commun.* 14, 2314. <https://doi.org/10.1038/s41467-023-37437-5>.
18. Regev, A., Teichmann, S.A., Lander, E.S., Amit, I., Benoist, C., Birney, E., Bodenmiller, B., Campbell, P., Carninci, P., Clatworthy, M., et al. (2017). The Human Cell Atlas. *eLife* 6, e27041. <https://doi.org/10.7554/eLife.27041>.
19. Ecker, J.R., Geschwind, D.H., Kriegstein, A.R., Ngai, J., Osten, P., Polio-udakis, D., Regev, A., Sestan, N., Wickersham, I.R., and Zeng, H. (2017). The BRAIN Initiative Cell Census Consortium: Lessons Learned toward Generating a Comprehensive Brain Cell Atlas. *Neuron* 96, 542–557. <https://doi.org/10.1016/j.neuron.2017.10.007>.
20. Yao, Z., van Velthoven, C.T.J., Kunst, M., Zhang, M., McMillen, D., Lee, C., Jung, W., Goldy, J., Abdelhak, A., Aitken, M., et al. (2023). A high-resolution transcriptomic and spatial atlas of cell types in the whole mouse brain. *Nature* 624, 317–332. <https://doi.org/10.1038/s41586-023-06812-z>.
21. Sunkin, S.M., Ng, L., Lau, C., Dolbeare, T., Gilbert, T.L., Thompson, C.L., Hawrylycz, M., and Dang, C. (2013). Allen Brain Atlas: an integrated spatiotemporal portal for exploring the central nervous system. *Nucleic Acids Res.* 41, D996–D1008. <https://doi.org/10.1093/nar/gks1042>.
22. Van de Sande, B., Lee, J.S., Mutasa-Gottgens, E., Naughton, B., Bacon, W., Manning, J., Wang, Y., Pollard, J., Mendez, M., Hill, J., et al. (2023). Applications of single-cell RNA sequencing in drug discovery and development. *Nat. Rev. Drug Discov.* 22, 496–520. <https://doi.org/10.1038/s41573-023-00688-4>.
23. Zhang, W., Maeser, D., Lee, A., Huang, Y., Gruener, R.F., Abdelbar, I.G., Jena, S., Patel, A.G., and Huang, R.S. (2024). Integration of Pan-Cancer Cell Line and Single-Cell Transcriptomic Profiles Enables Inference of Therapeutic Vulnerabilities in Heterogeneous Tumors. *Cancer Res.* 84, 2021–2033. <https://doi.org/10.1158/0008-5472.CAN-23-3005>.
24. Nomura, S. (2021). Single-cell genomics to understand disease pathogenesis. *J. Hum. Genet.* 66, 75–84. <https://doi.org/10.1038/s10038-020-00844-3>.
25. Liu, Z.L., Meng, X.Y., Bao, R.J., Shen, M.Y., Sun, J.J., Chen, W.D., Liu, F., and He, Y. (2024). Single cell deciphering of progression trajectories of the tumor ecosystem in head and neck cancer. *Nat. Commun.* 15, 2595. <https://doi.org/10.1038/s41467-024-46912-6>.
26. Rindler, K., Jonak, C., Alkon, N., Thaler, F.M., Kurz, H., Shaw, L.E., Stingl, G., Weninger, W., Halbritter, F., Bauer, W.M., et al. (2021). Single-cell RNA sequencing reveals markers of disease progression in primary cutaneous T-cell lymphoma. *Mol. Cancer* 20, 124. <https://doi.org/10.1186/s12943-021-01419-2>.
27. Liang, Q., Cheng, X., Wang, J., Owen, L., Shakoar, A., Lillvis, J.L., Zhang, C., Farkas, M., Kim, I.K., Li, Y., et al. (2023). A multi-omics atlas of the human retina at single-cell resolution. *Cell Genom.* 3, 100298. <https://doi.org/10.1016/j.xgen.2023.100298>.
28. Tabula Sapiens Consortium; Jones, R.C., Karkanias, J., Krasnow, M.A., Pisco, A.O., Quake, S.R., Salzman, J., Yosef, N., Bulthaupt, B., Brown, P., et al. (2022). The Tabula Sapiens: A multiple-organ, single-cell transcriptomic atlas of humans. *Science* 376, eabl4896. <https://doi.org/10.1126/science.abl4896>.
29. Tosti, L., Hang, Y., Debnath, O., Tiesmeyer, S., Trefzer, T., Steiger, K., Ten, F.W., Lukassen, S., Ballke, S., Köhl, A.A., et al. (2021). Single-Nucleus and In Situ RNA-Sequencing Reveal Cell Topographies in the Human Pancreas. *Gastroenterology* 160, 1330–1344.e11. <https://doi.org/10.1053/j.gastro.2020.11.010>.
30. Roy, A.L., Conroy, R., Smith, J., Yao, Y., Beckel-Mitchener, A.C., Anderson, J.M., and Wilder, E.L. (2018). Accelerating a paradigm shift: The Common Fund Single Cell Analysis Program. *Sci. Adv.* 4, eaat8573. <https://doi.org/10.1126/sciadv.aat8573>.
31. Spaethling, J.M., Na, Y.J., Lee, J., Ulyanova, A.V., Baltuch, G.H., Bell, T.J., Brem, S., Chen, H.I., Dueck, H., Fisher, S.A., et al. (2017). Primary Cell Culture of Live Neurosurgically Resected Aged Adult Human Brain Cells and Single Cell Transcriptomics. *Cell Rep.* 18, 791–803. <https://doi.org/10.1016/j.celrep.2016.12.066>.
32. Hong, R., Koga, Y., Bandyadka, S., Leshchych, A., Wang, Y., Akavoor, V., Cao, X., Sarfraz, I., Wang, Z., Alabdullatif, S., et al. (2022). Comprehensive generation, visualization, and reporting of quality control metrics for single-cell RNA sequencing data. *Nat. Commun.* 13, 1688. <https://doi.org/10.1038/s41467-022-29212-9>.
33. Machado, L., Relaix, F., and Mourikis, P. (2021). Stress relief: emerging methods to mitigate dissociation-induced artefacts. *Trends Cell Biol.* 31, 888–897. <https://doi.org/10.1016/j.tcb.2021.05.004>.
34. Caglayan, E., Liu, Y., and Konopka, G. (2022). Neuronal ambient RNA contamination causes misinterpreted and masked cell types in brain single-nuclei datasets. *Neuron* 110, 4043–4056.e5. <https://doi.org/10.1016/j.neuron.2022.09.010>.
35. Zhang, Y., Tan, J., Yang, K., Fan, W., Yu, B., and Shi, W. (2023). Ambient RNAs removal of cortex-specific snRNA-seq reveals Apoe(+) microglia/macrophage after deeper cerebral hypoperfusion in mice. *J. Neuroinflammation* 20, 152. <https://doi.org/10.1186/s12974-023-02831-9>.
36. Alvarez, M., Rahmani, E., Jew, B., Garske, K.M., Miao, Z., Benhammou, J.N., Ye, C.J., Pisegna, J.R., Pietiläinen, K.H., Halperin, E., and Pajukanta, P. (2020). Enhancing droplet-based single-nucleus RNA-seq resolution using the semi-supervised machine learning classifier DIEM. *Sci. Rep.* 10, 11019. <https://doi.org/10.1038/s41598-020-67513-5>.

37. Young, M.D., and Behjati, S. (2020). SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. *GigaScience* 9, gaa151. <https://doi.org/10.1093/gigascience/gaa151>.
38. McGinnis, C.S., Murrow, L.M., and Gartner, Z.J. (2019). DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst.* 8, 329–337.e4. <https://doi.org/10.1016/j.cels.2019.03.003>.
39. Fleming, S.J., Chaffin, M.D., Arduini, A., Akkad, A.D., Banks, E., Marioni, J.C., Philippakis, A.A., Ellinor, P.T., and Babadi, M. (2023). Unsupervised removal of systematic background noise from droplet-based single-cell experiments using CellBender. *Nat. Methods* 20, 1323–1335. <https://doi.org/10.1038/s41592-023-01943-7>.
40. Hafemeister, C., and Satija, R. (2019). Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol.* 20, 296. <https://doi.org/10.1186/s13059-019-1874-1>.
41. Martelotto, L. (2020). Frankenstein' protocol for nuclei isolation from fresh and frozen tissue for snRNAseq. <https://doi.org/10.17504/protocols.io.3eqgjdww>.
42. xGenomics Chromium Nuclei Isolation Kit. (2022). Document Number CG000505 Rev A. <https://doi.org/10.17504/protocols.io.3eqgjdww>. https://cdn.10xgenomics.com/image/upload/v1660261285/support-documents/CG000505_Chromium_Nuclei_Isolation_Kit_UG_RevA.pdf.
43. S2Genomics Nuclei Isolation from Frozen Mouse Brain Tissue for Single Nuclei Sequencing Applications P/N 100-272-080. <https://info.s2genomics.com/hubfs/2024%20Website%20Update%20Resources/s2-genomics-demonstrated-protocol-nuclei-isolation-and-cleanup-frozen-mouse-brain-v241018.pdf>.
44. Han, X., Wang, R., Zhou, Y., Fei, L., Sun, H., Lai, S., Saadatpour, A., Zhou, Z., Chen, H., Ye, F., et al. (2018). Mapping the Mouse Cell Atlas by Microwell-Seq. *Cell* 172, 1091–1107.e17. <https://doi.org/10.1016/j.cell.2018.02.001>.
45. Franzen, O., Gan, L.M., and Björkgrén, J.L.M. (2019). PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data. *Database* 2019, baz046. <https://doi.org/10.1093/database/baz046>.
46. Liu, B., Li, C., Li, Z., Wang, D., Ren, X., and Zhang, Z. (2020). An entropy-based metric for assessing the purity of single cell populations. *Nat. Commun.* 11, 3155. <https://doi.org/10.1038/s41467-020-16904-3>.
47. Clarke, Z.A., and Bader, G.D. (2024). MALAT1 expression indicates cell quality in single-cell RNA sequencing data. Preprint at bioRxiv. <https://doi.org/10.1101/2024.07.14.603469>.
48. Montserrat-Ayuso, T., and Esteve-Codina, A. (2024). High content of nuclei-free low-quality cells in reference single-cell atlases: a call for more stringent quality control using nuclear fraction. *BMC Genom.* 25, 1124. <https://doi.org/10.1186/s12864-024-11015-5>.
49. Holden, S.S., Grandi, F.C., Aboubakr, O., Higashikubo, B., Cho, F.S., Chang, A.H., Forero, A.O., Morningstar, A.R., Mathur, V., Kuhn, L.J., et al. (2021). Complement factor C1q mediates sleep spindle loss and epileptic spikes after mild brain injury. *Science* 373, eabj2685. <https://doi.org/10.1126/science.abj2685>.
50. Yan, X.J., Li, Y.B., Liu, W., Wu, H.Y., and Yu, G.F. (2022). Elevated Serum Complement C1q Levels After Traumatic Brain Injury and Its Association with Poor Prognosis. *Neuropsychiatr. Dis. Treat.* 18, 47–55. <https://doi.org/10.2147/NDT.S348682>.
51. Ma, W., Oswald, J., Rios Angulo, A., and Chen, Q. (2024). Tmem119 expression is downregulated in a subset of brain metastasis-associated microglia. *BMC Neurosci.* 25, 6. <https://doi.org/10.1186/s12868-024-00846-3>.
52. Eid, T., Lee, T.S.W., Patrylo, P., and Zaveri, H.P. (2019). Astrocytes and Glutamine Synthetase in Epileptogenesis. *J. Neurosci. Res.* 97, 1345–1362. <https://doi.org/10.1002/jnr.24267>.
53. Sandhu, M.R.S., Gruenbaum, B.F., Gruenbaum, S.E., Dhaher, R., Deshpande, K., Funaro, M.C., Lee, T.S.W., Zaveri, H.P., and Eid, T. (2021). Astroglial Glutamine Synthetase and the Pathogenesis of Mesial Temporal Lobe Epilepsy. *Front. Neurol.* 12, 665334. <https://doi.org/10.3389/fneur.2021.665334>.
54. Andreatta, M., and Carmona, S.J. (2021). UCell: Robust and scalable single-cell gene signature scoring. *Comput. Struct. Biotechnol. J.* 19, 3796–3798. <https://doi.org/10.1016/j.csbj.2021.06.043>.
55. Leigh, D.M., Lischer, H.E.L., Grossen, C., and Keller, L.F. (2018). Batch effects in a multiyear sequencing study: False biological trends due to changes in read lengths. *Mol. Ecol. Resour.* 18, 778–788. <https://doi.org/10.1111/1755-0998.12779>.
56. Murphy, A.E., Fancy, N., and Skene, N. (2023). Avoiding false discoveries in single-cell RNA-seq by revisiting the first Alzheimer's disease dataset. *eLife* 12. <https://doi.org/10.7554/eLife.90214>.
57. Leek, J.T., Scharpf, R.B., Bravo, H.C., Simcha, D., Langmead, B., Johnson, W.E., Geman, D., Baggerly, K., and Irizarry, R.A. (2010). Tackling the widespread and critical impact of batch effects in high-throughput data. *Nat. Rev. Genet.* 11, 733–739. <https://doi.org/10.1038/nrg2825>.
58. Deng, Q., Wu, C., Parker, E., Liu, T.C.Y., Duan, R., and Yang, L. (2024). Microglia and Astrocytes in Alzheimer's Disease: Significance and Summary of Recent Advances. *Aging Dis.* 15, 1537–1564. <https://doi.org/10.14338/AD.2023.0907>.
59. Haimon, Z., Volaski, A., Orthgiess, J., Boura-Halfon, S., Varol, D., Shemer, A., Yona, S., Zuckerman, B., David, E., Chappell-Maor, L., et al. (2018). Re-evaluating microglia expression profiles using RiboTag and cell isolation strategies. *Nat. Immunol.* 19, 636–644. <https://doi.org/10.1038/s41590-018-0110-6>.
60. Kang, S.S., Ebbert, M.T.W., Baker, K.E., Cook, C., Wang, X., Sens, J.P., Kocher, J.P., Petrucelli, L., and Fryer, J.D. (2018). Microglial translational profiling reveals a convergent APOE pathway from aging, amyloid, and tau. *J. Exp. Med.* 215, 2235–2245. <https://doi.org/10.1084/jem.20180653>.
61. Gao, C., Jiang, J., Tan, Y., and Chen, S. (2023). Microglia in neurodegenerative diseases: mechanism and potential therapeutic targets. *Signal Transduct. Target. Ther.* 8, 359. <https://doi.org/10.1038/s41392-023-01588-0>.
62. Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., da Silva Santos, L.B., Bourne, P.E., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* 3, 160018. <https://doi.org/10.1038/sdata.2016.18>.
63. OpenAI (2024). *ChatGPT* [Large language model].
64. Google (2024). *Gemini* [Large language model].
65. AI, P. (2024). *Perplexity* [AI search engine].
66. 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., et al. (2017). Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* 8, 14049. <https://doi.org/10.1038/ncomms14049>.
67. Hao, Y., Stuart, T., Kowalski, M.H., Choudhary, S., Hoffman, P., Hartman, A., Srivastava, A., Molla, G., Madad, S., Fernandez-Granda, C., and Satija, R. (2024). Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat. Biotechnol.* 42, 293–304. <https://doi.org/10.1038/s41587-023-01767-y>.
68. Aciri, D.J., Mustaklem, R., Horan-Portelance, L., and Mesecar, M. (2024). neurogenetics/SAHA: ALL VERSIONS. Zenodo. <https://doi.org/10.5281/zenodo.14040699>.
69. Lein, E.S., Hawrylycz, M.J., Ao, N., Ayres, M., Bensinger, A., Bernard, A., Boe, A.F., Boguski, M.S., Brockway, K.S., Byrnes, E.J., et al. (2007). Genome-wide atlas of gene expression in the adult mouse brain. *Nature* 445, 168–176. <https://doi.org/10.1038/nature05453>.
70. Phipson, B., Sim, C.B., Porrello, E.R., Hewitt, A.W., Powell, J., and Oshlack, A. (2022). propeller: testing for differences in cell type proportions in single cell data. *Bioinformatics* 38, 4720–4726. <https://doi.org/10.1093/bioinformatics/btac582>.
71. Korotkevich, G., Sukhov, V., Budin, N., Shpak, B., Artyomov, M.N., and Sergushichev, A. (2021). Fast gene set enrichment analysis. Preprint at bioRxiv, 060012. <https://doi.org/10.1101/060012>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
EZ nuclei suspension buffer	Sigma	EZ PREP NUC-101
RNase Inhibitor	Millipore Sigma	3335399001
Nuclei PURE 2M Sucrose Cushion Solution	Sigma	S9308
Nuclei PURE Sucrose Cushion Buffer	Sigma	S9058
Nuclei Debris Removal Stock Solution	S2 Genomics	#100-253-628
Nuclear Storage Reagent	S2 Genomics	#100-063-623
Critical commercial assays		
Chromium Nuclei Isolation Kit	10X Genomics	PN-1000493
Deposited data		
snRNA-seq data	GEO database	GSE290858
Experimental models: Organisms/strains		
C57BL/6J	The Jackson Laboratory	000664; RRID: IMSR_JAX:000664
Software and algorithms		
Cell Ranger	10x Genomics	N/A
SoupX	https://doi.org/10.1093/gigascience/giaa151	N/A
DoubletFinder	https://doi.org/10.1016/j.cels.2019.03.003	N/A
Seurat	https://doi.org/10.1038/s41587-023-01767-y	N/A
SCTransform	https://doi.org/10.1186/s13059-019-1874-1	N/A
scMCA	https://doi.org/10.1007/978-1-4939-9057-3_6	N/A
PanglaoDB	https://doi.org/10.1093/database/baz046	N/A
ROGUE	https://doi.org/10.1038/s41467-020-16904-3	N/A
SAHA	https://doi.org/10.5281/zenodo.14040698	N/A
UCell	https://doi.org/10.1016/j.csbj.2021.06.043	N/A
Other		
Code	This paper	https://github.com/jungsukimlab/NucleiIsolation doi:10.5281/zenodo.18378837

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Animals

All animal experiments were approved and performed in compliance with the guidelines of the Institutional Animal Care and Use Committee at Indiana University. Male C57BL/6J mice (B6; The Jackson Laboratory, 000664) were maintained under a 12-h/12-h light/dark cycle in a temperature-controlled room with free access to food and water until they reached 6 months of age.

METHOD DETAILS

Tissue collection

At 6 months, B6 mice were anesthetized with Avertin (250 mg/kg, intraperitoneal) and transcardially perfused with cold phosphate-buffered saline (1x PBS). The brains were immediately removed. After the cerebellum and olfactory bulb were removed, the right hemisphere was dissected into cortical and hippocampal regions and immediately frozen in dry ice. The samples were stored at -80°C until further processing.

Centrifugation-based nucleus isolation

Frozen anterior cortex tissue was added to a 2 mL tube with 1 mL of ice-cold EZ nuclei suspension buffer (Sigma, EZ PREP NUC-101). The tissue was dissociated via gentle, manual pipetting. The dissociated tissue was transferred into a 15 mL conical tube, and an additional 6 mL of ice-cold EZ nuclei suspension buffer was added. The mixture was incubated at 4°C for 10 minutes while rotating. After rotation, the sample was centrifuged at 500×g for 5 minutes at 4°C. The supernatant was removed, and the pellet was resuspended in 1 mL of ice-cold EZ nuclei suspension buffer with gentle pipetting. An additional 6 mL of EZ nuclei resuspension buffer was added. The mixture was incubated at 4°C for 10 minutes while rotating. After rotation, the sample was centrifuged at 500×g for 5 minutes at 4°C. The supernatant was removed. One milliliter of Nuclei Wash and Resuspension Buffer (NWR Buffer) (1× PBS, 2% bovine serum albumin (BSA), and 2 U/μl RNase inhibitor (Millipore Sigma, 3335399001)) was added, and the mixture was gently pipetted. An additional 6 mL of NWR Buffer was added. The sample was subsequently centrifuged at 500×g for 5 minutes at 4°C. This wash step with NWR Buffer was repeated four times. After the last wash, the nuclei were resuspended in 1 mL NWR buffer, and then another 6 mL was added. A 70-μm and 30-μm cell strainer was prewetted with LS Column Calibration Buffer (1× PBS, 0.5% BSA). The nuclei were filtered through a 70 μm strainer and then through a 30 μm strainer to remove clumps. The remaining solution was centrifuged at 500×g for 5 minutes at 4°C. The supernatant was removed, and the pellet was resuspended in 1 mL of NWR Buffer. 500 μL of the resuspended nuclei were added to an Eppendorf tube containing 900 μL of sucrose cushion buffer (Nuclei Pure 2 M sucrose cushion solution (Sigma, S9308) and 10% nuclei PURE sucrose cushion buffer (Sigma, S9058)). The mixture was carefully pipetted together. In a separate Eppendorf tube, 500 μL of sucrose cushion buffer was added. To create the sucrose gradient, 1400 μL of the nuclear suspension in sucrose cushion buffer was added carefully to the top of an Eppendorf tube containing 500 μL of sucrose cushion buffer without mixing. The sucrose gradient was centrifuged at 13,000×g for 45 minutes at 4°C. After centrifugation, the supernatant was almost completely removed, leaving approximately 100 μL in the tube. The pellet was resuspended in 1 mL of NWR Buffer with gentle pipette mixing. The solution was filtered through a 40 μm cell strainer to remove any remaining debris.

Column-based isolation

Nuclei suspensions were prepared according to the 10x Chromium Nuclei Isolation Kit instructions, with all steps performed on ice.⁴² Frozen anterior cortex tissue was added to a prechilled sample dissociation tube (10x Genomics, 2000564). 200 μL of lysis buffer (10x Genomics, 2000558) was added to the tube, and the tissue was dissociated with a pestle. Then, 300 μL of additional lysis buffer was added, and the sample was incubated on ice for 10 minutes. The solution was then added to a Nuclei Isolation Column (10x Genomics, 2000562) in a collection tube (10x Genomics, 2000563). The column was subsequently centrifuged at 13,000×g for 20 seconds at 4°C. The flowthrough was quickly vortexed. The mixture was then centrifuged at 500×g for 3 minutes at 4°C. The supernatant was removed, and the pellet was resuspended in 500 μL of Debris Removal Buffer (10x Genomics, 2000560). The suspension was subsequently centrifuged at 700×g for 10 minutes at 4°C. The supernatant was removed and resuspended in 1 mL of wash and resuspension buffer (1× PBS, 10% BSA, 2.5% RNase Inhibitor (10x Genomics, 2000565)) and then centrifuged at 500×g for 5 minutes at 4°C. The washing step was repeated once more. The supernatant was removed, and the nuclear suspension was resuspended in 50 μL of wash and resuspension buffer by gentle pipette mixing.

Machine-assisted nuclei isolation

Nuclei were isolated via the Singulator 100 system, with all steps performed on ice.⁴³ Before isolation, nuclei isolation cartridges (S2 Genomics, #100-063-287) were precooled to -25°C, and the simulator was prechilled to ensure optimal processing conditions. For each sample, the frozen anterior cortex was placed into a precooled nucleus isolation cartridge along with 75 μL of RNase inhibitor (Millipore Sigma, 3335399001). The samples were processed via the Standard Nuclei Isolation Protocol. The resulting 3 mL of the nuclear suspension was transferred to a 15 mL conical tube and centrifuged at 500×g for 5 minutes at 4°C to pellet the nuclei. The supernatant was aspirated, and the pellet was resuspended in 3 mL of 20% Nuclei Debris Removal Stock Solution (S2 Genomics, #100-253-628) in Nuclear Storage Reagent (S2 Genomics, #100-063-623). The suspension was centrifuged for 8 minutes at 700×g and 4°C with a decreased brake setting (3 out of 10, Sorvall Legend X1R Centrifuge (Thermo Fisher)). The tubes were carefully removed without disturbing the gradient, and the gradient was aspirated from just below the meniscus via a P1000 pipette to remove the myelin debris, which was visible as a haze or flake at the top. Finally, the nuclei were resuspended in 3 mL of resuspension buffer (2% BSA in Nuclear Storage Reagent).

Library preparation

Representative images of nuclei isolation were obtained using an EVOS XL Core microscope. Nuclei concentration was quantified via trypan blue staining with a Countess Automated Cell Counter. All the nuclear suspensions were processed with 10x Chromium. Each nuclear suspension was counted and diluted to 1,000 nuclei per μL then loaded into a single-cell chip G and run on the Chromium Controller for GEM generation and barcoding. Sample processing and library preparation were performed according to the manufacturer's instructions via the Chromium Next GEM Single Cell 3' v3.1 dual index kit (10x Genomics) and SPRIselect paramagnetic bead-based chemistry (Beckman Coulter Life Sciences). The cDNA and library quality were assessed via a 2100 Bioanalyzer and a high-sensitivity DNA kit (Agilent Technologies). The final library concentration was determined via a QuBit fluorometer and a dsDNA HS assay kit (Thermo Fisher Scientific). Sequencing was performed on a NovaSeq 6000 (v1.5 S2; Illumina) with a 28-10-10-91 read setup to a mean sequencing depth of 437.9 ± 23.9 million reads per sample.

QUANTIFICATION AND STATISTICAL ANALYSIS

snRNA-seq data analysis

The sequencing data were processed with the Cell Ranger pipeline (v7.0.1, 10x Genomics) and aligned to GRCm 38 (cellranger reference genome: gex-mm10-2020-A).⁶⁶ The filtered feature-cell barcode matrices (including the hashtag count matrix) generated by CellRanger were loaded into SoupX (v1.6.2) in RStudio (v1.4.1717) running R (v4.3.1).³⁷ SoupX was used to quantify ambient RNA using default parameters. Doublets were identified with DoubletFinder (v2.0.4).³⁸ The data were then loaded into Seurat (v5.0.1).⁶⁷ The data were normalized via SCTransform (v0.4.1).⁴⁰ Nuclei were clustered via the first 35 principal components based on an elbow plot. Cluster marker genes were identified via the *FindAllMarkers* function. Clusters were manually annotated to the broad cell type level (i.e. Excitatory Neurons, Astrocytes, etc.) via scMCA (v0.2.0), which is provided by the Mouse Cell Atlas, and with PanglaoDB.^{44,45} To further refine annotations, cluster marker and gene expression profile analysis for cell type and subclass annotations was performed with SAHA (pre-release)⁶⁸ with PanglaoDB and Allen Brain Atlas.⁶⁹ The cell type proportions were quantified using speckle (v0.0.3).⁷⁰ Population homogeneity was estimated with ROGUE (v1.0).⁴⁶ The mitochondria-derived reads were identified by the "mt" prefix, and the ribosomal reads were identified by the "Rp[sl]" prefix. Summary scoring was performed with UCell (v2.2.0).⁵⁴ Gene set enrichment analysis (GSEA) was performed using fgsea (v1.24.0).⁷¹

Statistical procedures

All tests were conducted using packages' default settings unless specifically mentioned otherwise. Error bar definitions and statistical tests are defined in the figure legends.