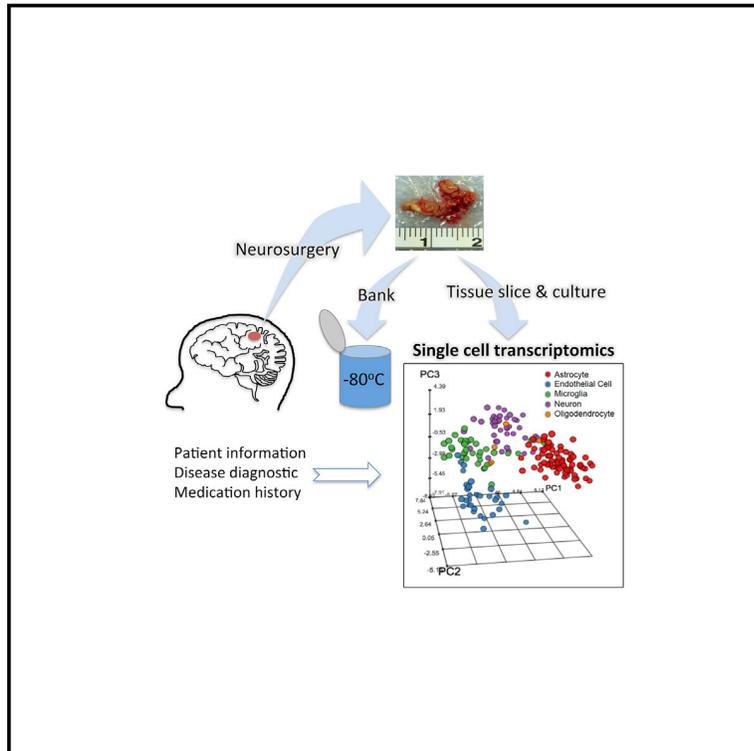# Primary Cell Culture of Live Neurosurgically Resected Aged Adult Human Brain Cells and Single Cell Transcriptomics

Graphical Abstract

Authors

Jennifer M. Spaethling, Young-Ji Na, Jaehee Lee, ..., Jai-Yoon Sul, Junhyong Kim, James H. Eberwine

Correspondence

eberwine@upenn.edu

In Brief

Human CNS disease is best studied in human neuronal cells. Spaethling et al. have succeeded in performing long-term primary cell culture and single-cell transcriptomics of human adult brain cells, including neurons from patients up to 67 years of age.

Highlights

- A system for culturing adult human neuronal cells after neurosurgery is presented

- Among the many adult human brain cells that are long-term cultured are neurons

- Human brain cell-type-enriched pri-miRNAs and lncRNAs are characterized

- Cell-type- and patient-specific transcriptional hierarchies were revealed

CrossMark

CellPress

Primary Cell Culture of Live Neurosurgically Resected Aged Adult Human Brain Cells and Single Cell Transcriptomics

Jennifer M. Spaethling,[1,5] Young-Ji Na,[2,5] Jaehee Lee,[1,5] Alexandra V. Ulyanova,[3] Gordon H. Baltuch,[3] Thomas J. Bell,[1] Steven Brem,[3] H. Isaac Chen,[3] Hannah Dueck,[2] Stephen A. Fisher,[2] Marcela P. Garcia,[1] Mugdha Khaladkar,[2] David K. Kung,[3] Timothy H. Lucas, Jr.,[3] Donald M. O'Rourke,[3] Derek Stefanik,[2] Jinhui Wang,[1] John A. Wolf,[3] Tamas Bartfai,[4] M. Sean Grady,[3] Jai-Yoon Sul,[1] Junhyong Kim,[2] and James H. Eberwine[1,6,*]

[1]Department of Pharmacology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA
[2]Department of Biology, School of Arts and Sciences, University of Pennsylvania, Philadelphia, PA 19104, USA
[3]Department of Neurosurgery, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA
[4]Department of Neurochemistry, Stockholm University, Stockholm 106 91, Sweden
[5]Co-first author
[6]Lead Contact
*Correspondence: eberwine@upenn.edu
 http://dx.doi.org/10.1016/j.celrep.2016.12.066

SUMMARY

Investigation of human CNS disease and drug effects has been hampered by the lack of a system that enables single-cell analysis of live adult patient brain cells. We developed a culturing system, based on a papain-aided procedure, for resected adult human brain tissue removed during neurosurgery. We performed single-cell transcriptomics on over 300 cells, permitting identification of oligodendrocytes, microglia, neurons, endothelial cells, and astrocytes after 3 weeks in culture. Using deep sequencing, we detected over 12,000 expressed genes, including hundreds of cell-type-enriched mRNAs, lncRNAs and pri-miRNAs. We describe cell-type- and patient-specific transcriptional hierarchies. Single-cell transcriptomics on cultured live adult patient derived cells is a prime example of the promise of personalized precision medicine. Because these cells derive from subjects ranging in age into their sixties, this system permits human aging studies previously possible only in rodent systems.

INTRODUCTION

The adult human brain is composed of an intricate network of multiple cell types that interact in direct and indirect ways. Diseases and drugs uniquely and differentially target these various cell types. Single-cell studies allow the highest resolution to assess this variability and cell-type-specific effects. Most past single-cell neuronal cell work has been performed in rodents (Dueck et al., 2015; Miyashiro et al., 1994; Tasic et al., 2016; Zeisel et al., 2015). Cell type studies in humans have been largely limited to postmortem studies (Hawrylycz et al., 2015; Lake et al., 2016), cancer cell lines, and, more recently, acute harvest

of cells from patients (Darmanis et al., 2015; Zhang et al., 2016). Although these studies provide valuable human transcriptomic information, the cells' acute harvest provides no means for morphological or long-term functional investigation other than sequencing. Cell selection methods limit the collection to subpopulations of each cell type, and nucleus sequencing likely results in an incomplete picture of the entire transcriptome. Some studies have focused on human embryonic stem cell (ESC)- and induced pluripotent stem cell (iPSC)-derived neurons to create induced neuron (iN) cells that can produce de novo synaptic connections (Zhang et al., 2013). For studying human CNS disease and drug effects, patient-derived fibroblasts used for iPSCs and stem cells are distinctly affected by disease and drug therapy. Developing and validating a model system that is easily manipulated to investigate the function and responsiveness of a broad range of cell types in the human brain is needed. A culture system that supports long-term survival of multiple adult cell types harvested from the adult human brain would enable an understanding of human cell-type-specific gene regulation without the confounding effects of species differences, cell line effects, or those introduced by *trans*-differentiation.

We have developed a culturing system for healthy adult human brain cells from patient biopsies collected at the time of surgery. These cells were cultured up to 84 days in vitro (DIV) and analyzed with deep sequencing of hundreds of single cells to obtain their individual RNA expression profiles. The single-cell resolution of this study allows us to measure the range and variance of expression of key genes and shows that mouse-derived cell type markers can be inappropriate discriminators of human cell types (Darmanis et al., 2015; Hawrylycz et al., 2015; Zhang et al., 2016). Use of human-sourced enriched gene lists supported by functional pathway analysis resulted in consistent identification of cell types and subtypes using multiple bioinformatic and statistical methods (k-means clustering, gene ontology [GO] annotation enrichment, etc.). We further identified cell-type-enriched primary microRNA (pri-miRNA) and long noncoding RNA (lncRNA) as well as potential transcription factor
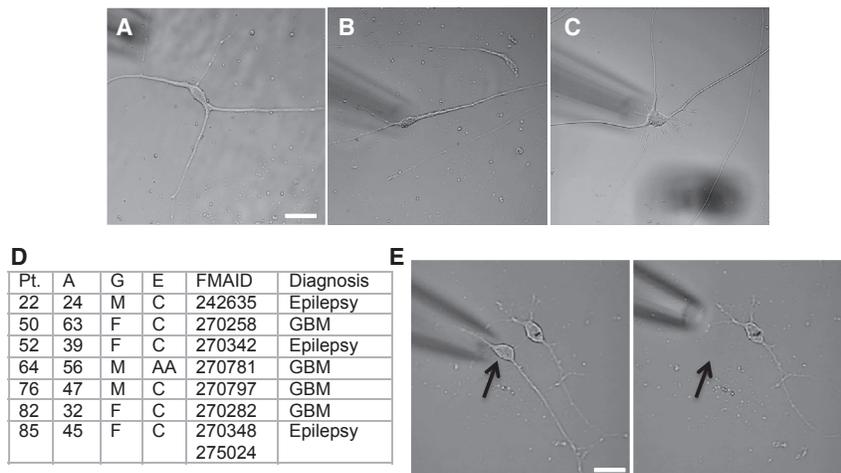
Figure 1. Healthy Long-Term Adult Human Brain Cell Cultures

(A–C) Representative healthy brain cells at 2 (A), 4 (B), and 8 (C) weeks in culture, respectively. Scale bars, 20 μm.

(D) Summary of patient information, including age, gender, ethnicity, FMAID for biopsy location, and diagnosis. M, male; F, female; C, white; AA, African-American; GBM, glioblastoma. Patient 85 also had samples harvested from the hippocampus in addition to the cortical FMAID listed.

(E) Images of culture pre- (left) and post-harvest (right) of a single cell using micropipette aspiration. Scale bar, 20 μm.

See also Figure S1.

Pt.	A	G	E	FMAID	Diagnosis
22	24	M	C	242635	Epilepsy
50	63	F	C	270258	GBM
52	39	F	C	270342	Epilepsy
64	56	M	AA	270781	GBM
76	47	M	C	270797	GBM
82	32	F	C	270282	GBM
85	45	F	C	270348 275024	Epilepsy

control pathways of genes that are candidates for driving the expression of subpopulations of the cell-type-defining genes.

We find that cells maintain their cell type classification throughout their time ex vivo. Morphological analysis of transcriptome-profiled cells suggests that transcriptionally distinct cell types can have a wide range of cell morphology in culture (Zhang et al., 2016) that we extend to other cell types. The human culturing system allows long-term maintenance and characterization of cells derived from a broad range of age groups (the oldest subject assessed was 63 years old). Importantly, such primary cell cultures, by design, will be absent their in vivo cellular connections because the natural microenvironment has been disrupted and, hence, will be somewhat different from their in vivo cellular counterparts. However, the ease of use and decades of fundamental and clinical data resulting from primary cells suggest that cultured adult human brain cells will be useful in understanding the fundamentals of neuronal cell functioning and responsiveness. This adult human primary cell culture resource provides a means for CNS drug testing.

RESULTS

Cortical and hippocampal biopsies were collected from seven patients at the Hospital of the University of Pennsylvania. Three of the patients were diagnosed with epilepsy, and the remainder were diagnosed with a brain tumor; e.g., glioblastoma, World Health Organization (WHO) grade IV, at a distance from the cortical biopsy site (6.825 ± 2.484 mm SD; Figure S1). Four were white females, two white males, and one African-American male, ranging in age from 24–63 years. Tissues were delivered to the laboratory in ice-cold oxygenated artificial cerebrospinal fluid (aCSF) approximately 10 min after excision. The tissue was dissociated, plated, and maintained in an incubator at 37°C, 5% CO_2. The cells in primary cell culture displayed complex morphological characteristics with smooth processes when present and no obvious vacuoles, highlighting their overall health (Chen et al., 2011; Figures 1A–1C). Cells were collected by pipette aspiration between 1 and 84 days after plating (Figure 1E; Table S1). The age of cells that do not divide is the age of the donor plus the

time in culture. For cells such as astrocytes that divide, the cell age is mixed based on the number of cell divisions that occurred in the patient and, subsequently, in culture. Each single cell's (Figure 1D) RNA was aRNA-amplified and deep-sequenced. On average, we obtained approximately 22.5 million (M) unique mapping reads per sample, of which approximately 60% were exonic reads that mapped to an average of 12,000 genes. When compared with publically available genotype tissue expression (GTEx) tissue transcriptome data, these single-cell transcriptomes had the greatest overlap with the whole-brain tissue samples of all brain cell types (Figure S1).

Marker- and Pathway-Based Identification of Human Brain Cell Types

To identify cell classes from the transcriptome data, we first clustered based on cell type marker expression. In our and other groups' experiences, mouse cell type markers sometimes fail to provide strong discrimination of comparable human brain cell types (Darmanis et al., 2015; Zhang et al., 2016). To obtain an initial classification, we used a list of 129 "marker genes" that have been used to successfully discriminate between multiple human brain cell types (Darmanis et al., 2015). As shown in Figure 2, of the 303 adult human brain cells that met our quality standards, we used k-means clustering (k = 8; Experimental Procedures) to group 187 single cells into classes (5 oligodendrocytes; 35 microglia, 42 neurons, 32 endothelial cells, and 73 astrocytes), with the remainder of the cells falling into three unknown classes (49, 45, and 22 cells, respectively; Figure 2A; Table S1). Although many of these unknown cells share some gene expression patterns with astrocytes, they were unable to be confidently classified into the six cell type gene sets. Importantly, the presence of aged adult human cortical neurons in long-term primary cell culture stands in distinction to the mouse system, where it has been difficult to perform primary culture of adult mouse cortical neurons.

To examine the effectiveness of human curated gene lists for revealing mouse cell types and vice-versa, we examined k-means clustering (k = 8) of previously described mouse single-cell data (Tasic et al., 2016). First, we used mouse homologs of a human curated gene list (Darmanis et al., 2015) to cluster the ~1,600 mouse cortical single-cell transcriptomes reported by
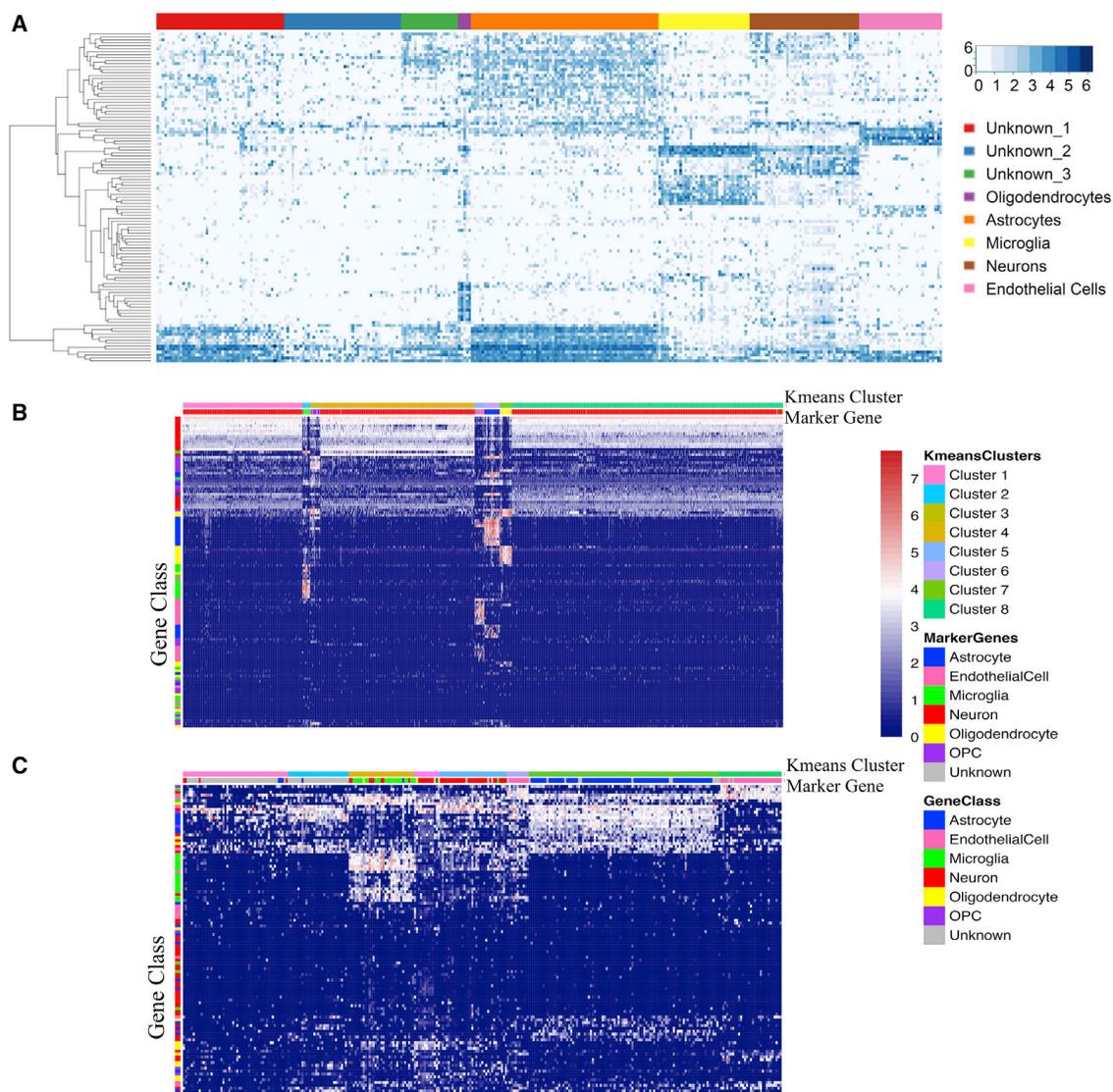
Figure 2. Identification of Cell Types Using Human and Cross-Species Markers

(A) Using human cell type markers, cells fell into eight transcriptional groups representing oligodendrocytes, microglia, neurons, astrocytes, endothelial cells, and unknown.

(B) Mouse homologs of the human curated cell type marker gene list were used to cluster the ∼1,600 mouse cortical single-cell transcriptomes reported by Tasic et al. (2016), showing strong concordance with the original paper's groupings.

(C) We used human homologs of 109 mouse genes described by Tasic et al. (2016) to carry out the same k-means clustering, showing concordance with our groupings but more varied expression and less defined expression patterns for each grouping.

Tasic et al. (2016). There is a high degree of concordance in the original paper's cell type annotations and cluster membership using homologous human curated genes (Figure 2B). We found overlap between the original cell type annotation and new clusters as follows: astrocytes, 100% in cluster 6; neurons, 99.8% in cluster 4, 99.7% in cluster 1, and 99.2% in cluster 8; endothelial cells 74.7% in cluster 5; microglia, 95.6% in cluster 2; and oligodendrocyte precursor cell (OPC), 100% in cluster 7 and 71.4% in cluster 3. We also used human homologs of 109 mouse genes described by Tasic et al. (2016) to carry out the same k-means

clustering on our data (Figure 2C). These mouse markers successfully clustered many brain cell types; however, there were large groups of genes that were expressed in multiple clusters, meaning they were not cell-type-specific in the human.

We next augmented the initial classification by assessing the activation of signaling pathways central to each putative cell type's function. We curated a list of publically available pathway databases and scored each pathway's activation levels. The pathways included four neurotransmitter release pathways to represent neuronal function, cell-cycle pathways to represent
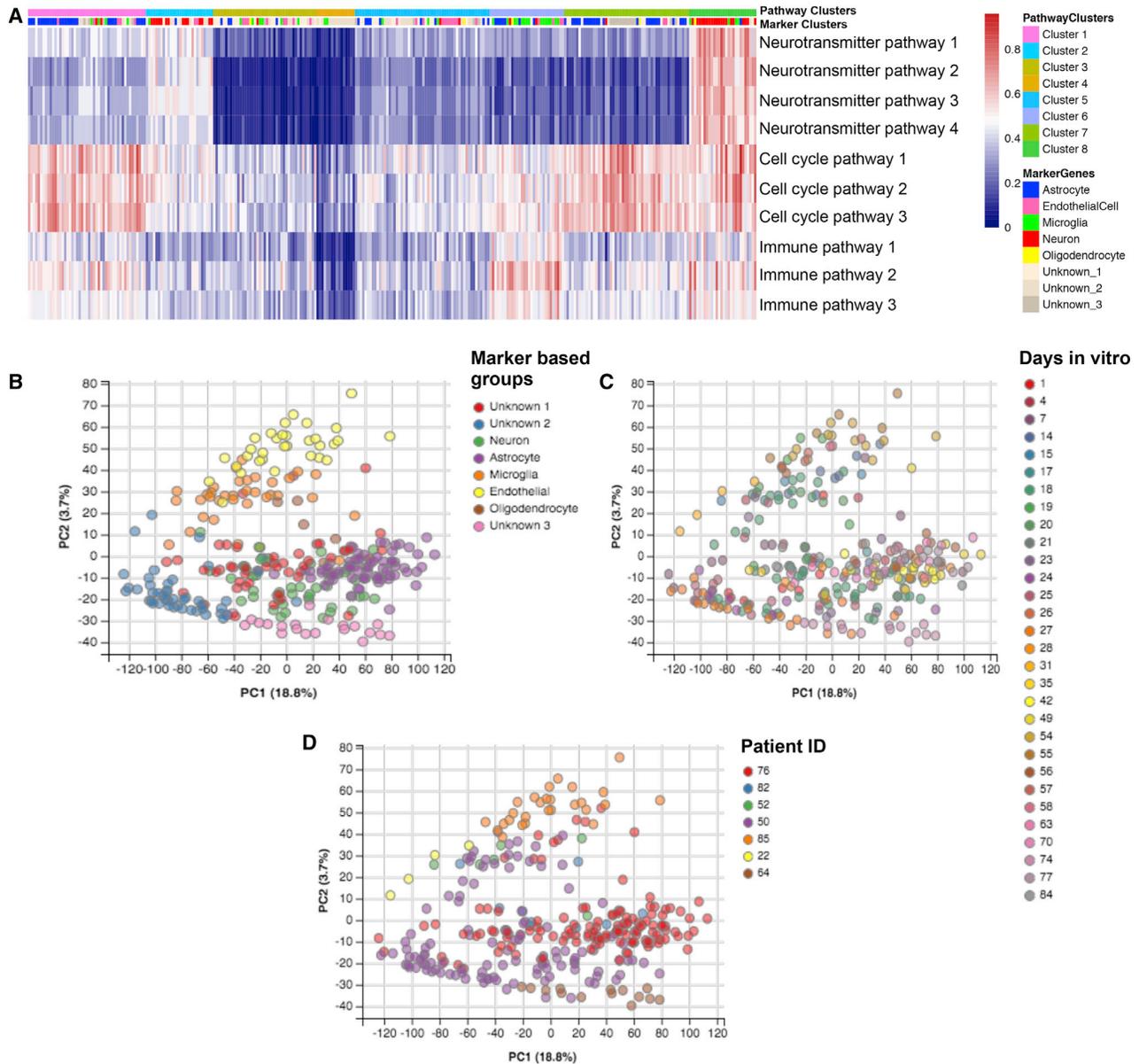
Figure 3. Pathway Activity Based on Unsupervised Clustering of Data

(A) K-means clustering of cells based on pathway activity (see Table S2 for exact pathway names). Top bar, pathway cluster; Bottom bar, marker-based cell type.

(B) PCA analysis of the entire expressed (non-null) transcriptome, showing clusters that correspond well with the cell types determined by k-means clustering using only human cell type marker genes.

(C) Identical PCA analysis color-coded based on DIV, showing DIV spread across clusters.

(D) Identical PCA analysis color-coded based on patient source ID, showing that although there is a primary patient representing the majority of points in each cell type, each cell type cluster does have multiple patients represented.

See also Figure S3.

dividing cells such as astrocytes, and immune response pathways to represent microglial function. K-means clustering (k = 8) was performed on each pathway for all samples (Figure 3A; Table S2). There was substantial overlap between the pathway clustering and the marker based cell type classifications. Astrocytes were the most common type in cluster 1 (55%) and cluster

7 (46%). Cluster 4 mostly consisted of cells from the unknown_2 cell type, whereas microglia were the most common type in cluster 6 (48%), and the neurons made up 71% of cluster 8 (Table S3). Endothelial cells and oligodendrocytes were spread over all of the clusters without a major dominant cluster assignment because there were no suitably unique pathways for their

functions in the pathway databases. Limits in the available curation of pathways in relation to specific cell function constrained the resolution of this analysis, but neurons in marker gene analysis showed the highest degree of neurotransmitter pathway activation, microglial cells showed distinct immune pathway activation, and astrocytes showed the highest degree of cell-cycle pathway activation.

We examined the dispersion patterns for the total transcriptome in contrast to the curated gene sets discussed (above). Figures 3C–3E show 2D principal component analysis (PCA) ordination of the whole transcriptome with covariates (Figure 3B, cell type assignments by curated genes; Figure 3C, culture dates; Figure 3D, patient ID). Although cultures from some patients were enriched in a given cell type, each cell type was seen in multiple patients, highlighting a cell type robustness that extends beyond a single patient's mRNA signature. Additionally, although we were limited by sample availability, cells of varying DIV were represented in each patient and each cell type. Further, PC1 separation was not dominated by any of the covariates, suggesting no dominant batch effects.

Stability of Cell Classification

We sought to characterize moderate- to low-abundance transcripts by high-depth sequencing. Because it is more common to sequence single-cell data to lower depth, we examined the effect of reducing sequencing coverage by creating randomly down-sampled datasets of 1 million, 0.5 million, and 0.1 million reads (Experimental Procedures). The effects of down-sampling depend upon the relative frequency distribution of the transcripts, and, therefore, different cell types and gene classes maybe affected differently. We examined the effect of the number of genes segregated by marker-defined cell type and gene classes (Figure S3). Over all cell types and gene classes, down-sampling resulted in a loss of ~52%, 60%, and 77% of the observed genes for 1M, 0.5M, and 0.1M sequencing depth. The greatest loss occurs for the neuron transcriptomes, with loss of ~63%, 77%, and 85%, respectively, for 1M, 0.5M, and 0.1M sequencing depth. The least loss is seen in the astrocyte cells, with ~36%, 44%, and 65%, respectively, for the same set of sequencing depths. For the different gene classes examined here (transcription factors, lncRNAs, pri-miRNAs and signal pathway genes), no particular class seemed more or less affected than the overall loss rate (other than transcription factors in neurons), suggesting that genes belonging to these classes are evenly distributed within the frequency distribution. Nevertheless, we note that, at a low sequencing depth of 0.1M reads, we fail to see between ~55% to 88% of genes in transcription regulation pathways. Down-sampling had a negligible effect on classification by the penalized linear discriminant analysis (PLDA) function, with less than 1% change in accuracy from the original full dataset for all cell types and all down-sampled treatments. This result is consistent with the idea that shallow-depth sequencing can be useful for identification of major cell types, although lower-abundance genes critical to specific cellular function and specification will be missing. These data also highlight the fact that these cell classifications and some subclassifications could be documented with hundreds of cells compared with the thousands of cells that have been reported elsewhere.

Classifying Brain Cells by Morphology and Non-coding RNA Profiles

For each of the sequenced cells, we imaged the cell prior to harvest to capture its morphology. We estimated a morphological classification tree using six morphological features scored from the images taken prior to RNA collection: gross cell size, cell shape, cell process, process complexity, cell margin, and cytosol/nuclear ratio (Table S1). Multiple morphologies were observable for each cell type (Figure 4). The maximum parsimony method was used (Swofford et al., 1996), and morphological trees were computed using the TBR heuristic search option in the program PAUP* (Swofford, 2003). We identified 31 distinct morphological groups where all cells within each group had the same constellation of morphological features (Figure S2). The maximum parsimony tree split into two major branches, broadly separated into cells that were generally bigger with either simple or complex processes and cells that were generally smaller without processes. Cell boundary states or cytosol/nuclear ratio features did not co-vary with cell subclasses. Particular morphological features showed variable association with the transcriptional cell types. For example, there were putative astrocytes with both complex and simple processes (Figures 4A and 4B), putative microglia with cell bodies that were very similarly sized or much larger than their nuclei (Figures 4C and 4D), large and small putative neurons (Figures 4E and 4F), and putative endothelial cells with and without processes (Figures 4G and 4H). As suggested by the representative images of the range of morphologies (Figure 4), examining cell morphology within each transcriptional cell type showed variable distribution of morphological traits for many cell types. We found that the vast majority of astrocytes were both large (98%) and had processes (99%), endothelial cells were mostly large (86%) but only 45% had processes, 67% of microglia were large and 58% had processes, and 65% of neurons were large and 78% had processes. The oligodendrocytes were divided with regard to size and process presence, but the small sample size in that class made it difficult to characterize (Table S1). To match the threshold used to select the cell type markers, we used 60-fold enrichment to identify genes associated with process-bearing cells. These genes were significantly enriched in GO terms such as synapse, neuron projection, cell junction, and axon (Bonferroni-corrected p = 0.003, 0.004, 0.02, and 0.05, respectively), thus suggesting the possibility that higher expression of these genes either results from or is responsible for the presence of processes.

We next assessed non-coding RNA expression, lncRNA and pri-miRNA, as possible discriminators of cell types because of their role in simultaneously modulating multiple genes. Using a curated list of 935 lncRNAs (Amaral et al., 2011) to perform a k-means cluster analysis (k = 8) of all of our samples, we found that there were lncRNA clusters, each strongly associated with marker-based cell types: cluster 2 with 69% of endothelial cells, cluster 6 with 71% of astrocytes, cluster 1 with 55% of neurons, cluster 5 with 80% of microglia, and cluster 8 with 80% of oligodendrocytes (Figure 5A; Table S1). A similar analysis of the 30 pri-miRNAs in our data showed that clusters found using only pri-miRNAs was generally disconcordant with marker gene-based clusters, although there was one cluster with a high degree of overlap with microglia.
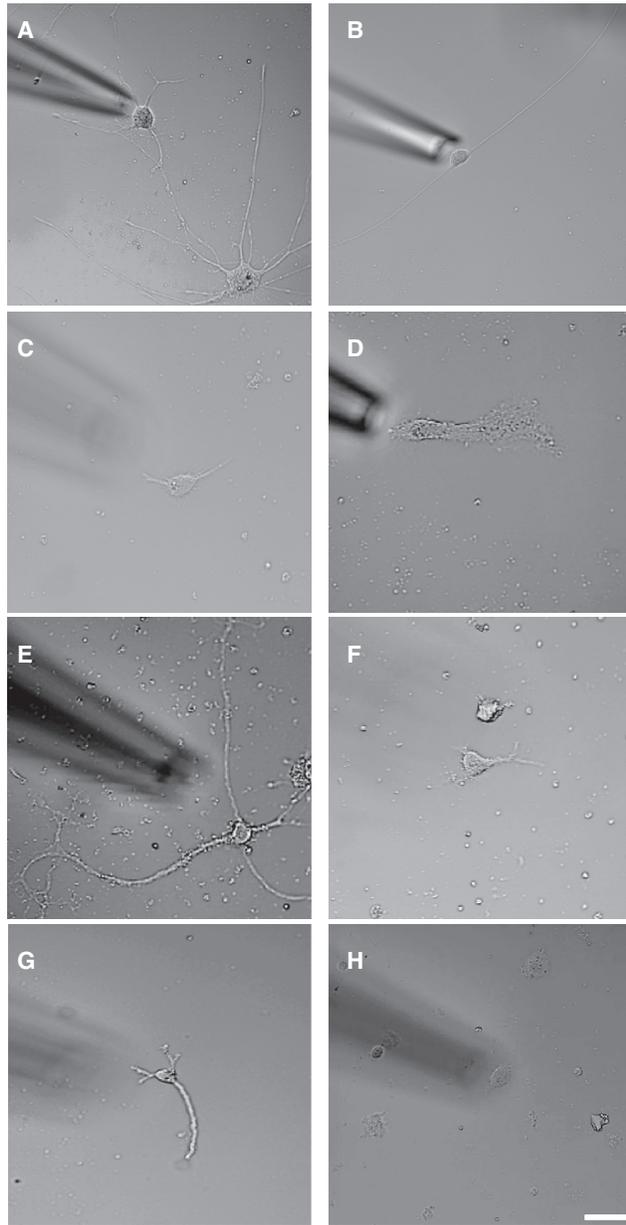
Figure 4. Range of Cell Morphology

(A–D) Astrocytes were observed with both complex (A) and simple processes (B), and microglia had cell bodies that were very similarly sized (C) or much larger than their nuclei (D).

(E–H) We observed both large (E) and small (F) neurons and endothelial cells with (G) and without processes (H).

Scale bar, 20 μm. See also Table S1 and Figure S2.

Cultured Adult Human Brain Cells Lacked Stem Cell Signatures

Although the cells analyzed were cultured from normal tissue harvested during the course of surgery, we wanted to evaluate the possible expression of proto-oncogene mRNA expression in cells from cancer patients (Vogelstein et al., 2013) to rule out the possibility that we had cultured oncogenic cells. We found enrichment of only one proto-oncogene (EGFR) in the cells from cancer patients over those from epilepsy patients. In addition, the tumor suppressor p53 is also expressed at similar or elevated levels in these cells from patients across all diseases, suggesting the normal anti-oncogenic function of p53 is intact in these cells. Because many of these samples are sourced from patients with glioblastoma, we are cognizant of the phenomenon of brain cells de-differentiating to a stem cell nature during tumorigenesis, indicated by a loss of mature cell type markers and increased expression of Oct4, Sox2, c-Myc, and Klf4 (Friedmann-Morvinski et al., 2012; Li et al., 2011). Expression levels of these genes across diseases in our sample set did not show broad expression of these markers, except for Sox2 expression, which was present across cells of all diseases and not restricted to glioblastoma samples. Furthermore, Sox2 is known to be expressed by astrocytes and is, therefore, alone not a good measure of the stem-like phenotype (Xia and Zhu, 2015). There was one cell of 300 that did have expression of the four stem cell genes, but it fell into one of the unknown cell type categories. This suggests that we are able to detect stem cells but that the vast majority of our cells are adult, mature, and differentiated cells.

Cell Type-Enriched Markers for Human Brain Cells

Using our marker gene-based classification, we examined cell-type-specific gene expression patterns to detect new markers. Again, to match the threshold used to identify cell type markers, we first identified genes that were enriched >60-fold in each cell type compared with every other cell type in a pairwise fashion. This analysis produced hundreds of additional candidate cell type markers for human cells (Table S4). We performed a GO term enrichment analysis for these genes and found significant enrichment of immune response genes in microglia, neuron ensheathment annotations in oligodendrocytes, neurogenesis and projection development in neurons, blood vessel and extracellular matrix organization genes in endothelial cells, and astrocyte differentiation and synapse support genes in astrocytes (Figure 6A).

In addition to these individual genes, we used PLDA to find a weighted linear combination of genes that would have high utility for identifying each cell type (Witten and Tibshirani, 2011). We extracted a discriminant axis for each cell type against every other cell type (Figure 6B). The PLDA functions resulted in a cross-validation (1:1) accuracy of 97%, 97%, 99%, 89%, and 99%, respectively, for marker-defined astrocytes, neurons, endothelial cells, microglia, and oligodendrocytes. The genes with large absolute value loading coefficients in the PLDA axis overlapped significantly with the original 129 marker genes (Table S5). The PLDA generated highly accurate classifiers using the whole transcriptome, with the penalization constraint to sparsely utilize the gene features. We examined the distribution of absolute value of the coefficient loadings for each cell type (Figure S4), which showed that most of the information is concentrated within the top 20 genes. It is desirable to reduce the feature set as much as possible to guard against over-fitting. Therefore, we constructed new PLDA classifiers based on the top 20 genes for each cell type. The cross-validation (1:1) accuracy ranged
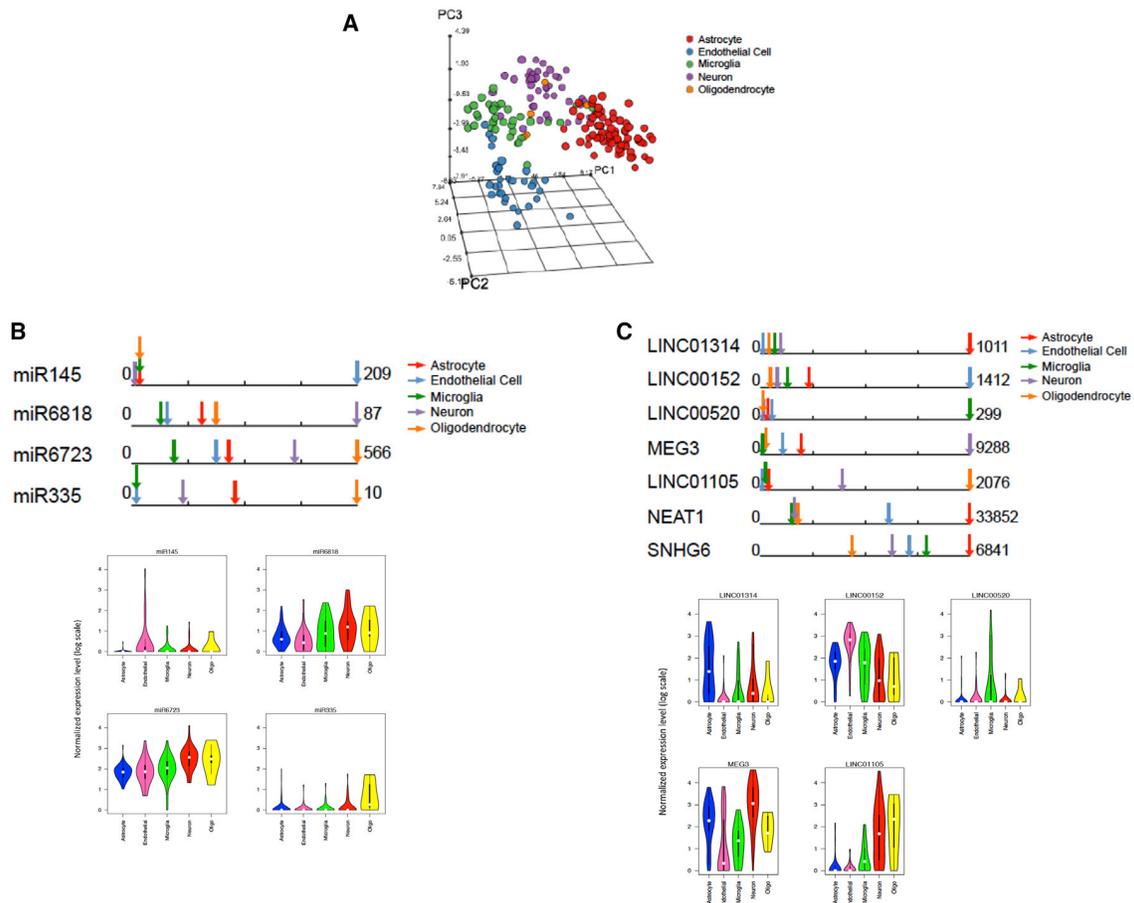
Figure 5. Noncoding RNAs Are Enriched in Cell Types and Can Discriminate Cell Types

(A) 3D PCA analysis of normalized lncRNA expression shows dispersion patterns that are generally consistent with marker-based cell types.

(B) Average normalized pri-miRNA expression in each cell type for significantly differentially expressed pri-miRNAs and the distribution of expression, highlighted by the violin plot, which shows the probability density for each pri-miRNA.

(C) Normalized lncRNA expression levels highlighting lncRNAs enriched in a single cell type (first five lncRNAs), enriched in multiple cell types (NEAT1), and SNGH6, which is highly expressed in all cell types (Figure S5), and violin plot showing the probability density for each lnc RNA.

from 94%–99%, showing the utility of this reduced discriminant function (Table S6). We used the reduced PLDA function to analyze the cells in the three previously unknown clusters (Figure 2A) and were able to assign identities to 45 of 116 unknown cell types (Table S7). We propose that the PLDA functions will have utility for identification of human cell types and for deconvolving cell mixtures into single-cell frequency counts (by expressing a tissue expression value as linear combinations of these discriminant axes).

Cell Type-Specific Transcription Factor Binding Sites, pri-miRNAs, and lncRNAs in Human Brain Cells

Using the genes enriched in each cell type, we asked whether there was potentially shared control of the genes by cell-type-specific transcription factor and/or miRNA binding motifs enriched in those genes using the ToppGene suite (Chen et al., 2009). Although there was some overlap between oligodendrocytes and neurons, every cell type had unique transcription fac-

tors that potentially controlled their enriched genes' expression (Table S7). For example, oligodendrocyte-enriched genes had enrichment in AP4 and MyoD (likely because of homology with OLIG1/2, (Hernandez and Casaccia, 2015)) transcription factor binding sites, whereas many microglial genes have potential to be controlled by PEA3. Neuronal genes had enrichment of multiple transcription factor and miRNA biding sites.

The presence of pri-miRNAs in our transcriptome was expected because pri-miRNAs have poly-A tails that can be amplified in our procedure. We performed differential expression analysis to identify cell type-enriched pri-miRNAs (Figure 5B). We found that MIR6723 is more highly expressed in neurons than in astrocytes, endothelial cells, and microglia, whereas MIR6818 is more highly expressed in neurons than in astrocytes and endothelial cells. MIR1199 and MIR335 are more highly expressed in oligodendrocytes than in any other cells. MIR145 is more highly expressed in endothelial cells than in astrocytes, microglia, and neurons. It has been shown previously that
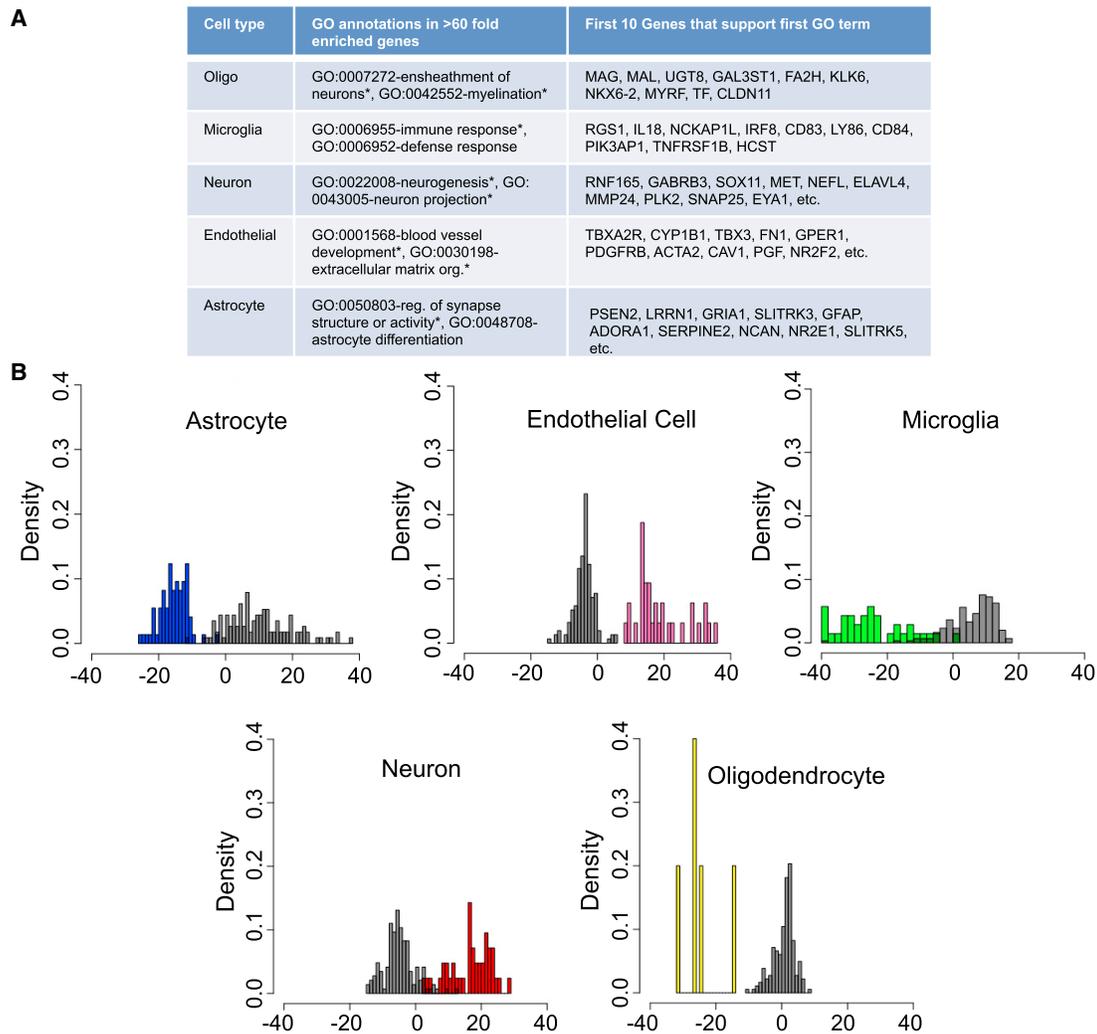
A

Cell type	GO annotations in >60 fold enriched genes	First 10 Genes that support first GO term
Oligo	GO:0007272-ensheathment of neurons*, GO:0042552-myelination*	MAG, MAL, UGT8, GAL3ST1, FA2H, KLK6, NKX6-2, MYRF, TF, CLDN11
Microglia	GO:0006955-immune response*, GO:0006952-defense response	RGS1, IL18, NCKAP1L, IRF8, CD83, LY86, CD84, PIK3AP1, TNFRSF1B, HCST
Neuron	GO:0022008-neurogenesis*, GO: 0043005-neuron projection*	RNF165, GABRB3, SOX11, MET, NEFL, ELAVL4, MMP24, PLK2, SNAP25, EYA1, etc.
Endothelial	GO:0001568-blood vessel development*, GO:0030198-extracellular matrix org.*	TBXA2R, CYP1B1, TBX3, FN1, GPER1, PDGFRB, ACTA2, CAV1, PGF, NR2F2, etc.
Astrocyte	GO:0050803-reg. of synapse structure or activity*, GO:0048708-astrocyte differentiation	PSEN2, LRRN1, GRIA1, SLITRK3, GFAP, ADORA1, SERPINE2, NCAN, NR2E1, SLITRK5, etc.

B

Figure 6. Cell Type-Specific Analysis
(A) Summary of enrichment of GO terms for each cell type. * indicates GO terms significantly enriched < 0.05 with Bonferroni correction; others listed are < 0.05 without correction.
(B) Histogram of PLDA score distribution for each of the five different cell types. The x axis of each plot shows the PLDA scores, whereas the histogram shows the scatter of the cells. The top colored panels show the cell type of interest, whereas the black bar panels show the other cell types.
See also Tables S4, S5–S7, S8 and Figure S3.

endothelial cells express MIR145 for critical regulation of smooth muscle cells in the peripheral vasculature (Hergenreider et al., 2012). Our study suggests that endothelial cells in the brain also express MIR145, which targets genes related to adherens junction and tight junction pathways. MIR335, which is highly expressed in oligodendrocytes, targets genes related to the neurotrophin signaling pathway (Vlachos et al., 2012).

To identify cell-type-dependent long noncoding RNAs, we performed differential expression analysis on lncRNAs by cell type. Among 935 lncRNAs, 113 lncRNAs are differentially expressed across cell types (Figures 5C; Figure S5). There was a range of expression from strong single cell type enrichment (i.e., LINC01314 astrocytes, LINC00152 endothelial cells, LINC00520 microglia, MEG3 neurons, and LINC01105 oligoden-

drocytes) to enrichment in a subset of cell types (i.e., NEAT1 high in astrocytes and endothelial cells) to shared expression across all cell types (i.e., SNHG6; Figure 5C). Meg3, which we found to be enriched in neurons, has been found previously in GABAergic neurons (Mercer et al., 2010).

mRNA Profiles Predict Neuronal Subclasses
Some of the CNS cell types have well known subclasses. For example, neurons are often categorized as excitatory or inhibitory; other researchers have argued for 7 to 16 neuronal subtypes in humans and at least 28 in mice (Darmanis et al., 2015; Lake et al., 2016; Zeisel et al., 2015). In an attempt to subdivide adult human neurons into finer subclasses, we used a curated list of possible excitatory or inhibitory markers (Darmanis et al.,
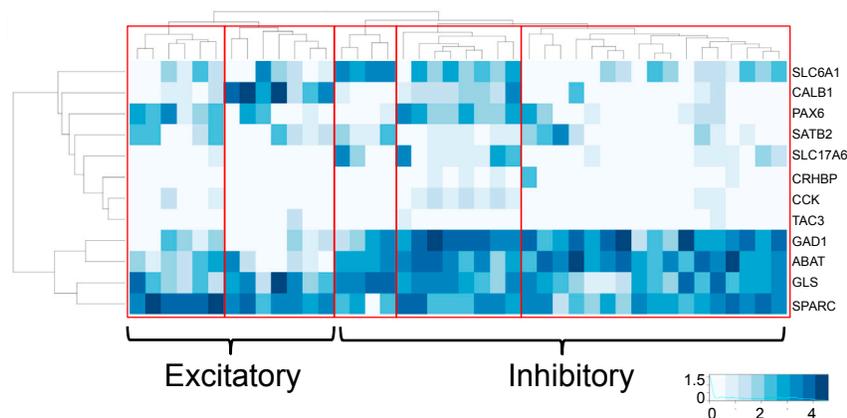
2015; Fish et al., 2011; Takamori, 2006). Marker expression suggested a mixed biology, with many cells expressing both excitatory and inhibitory markers, that has been observed previously in other species. We hierarchically clustered the neurons using this gene list and found that the cells fell into five clusters, and many displayed the possibility of co-release of neurotransmitters (Figure 7). Cluster 1 had relatively high expression of SPARC, suggesting cells undergoing significant remodeling and neurite outgrowth. Cluster 2 had high expression of calbindin 1 (CALB1) and almost no expression of glutamate decarboxylase 1/2 (GAD1/2, often used as markers of inhibitory neurons), suggesting that this cluster represents a group of CALB1+ excitatory neurons (Chard et al., 1995). Groups 3 and 5, with their expression of GAD1, are likely both inhibitory, although finer distinction was not obvious. Group 4 has mid-level expression for many excitatory and inhibitory linked genes but stronger expression of inhibitory markers, highlighting an example of cells exhibiting co-release probability. These findings show the importance of single-cell resolution phenotyping of neuronal subclasses to accurately assess the range of unique possible cell types that may result in mixed neurobiology.

Identification of Patient-Associated miRNAs and Transcription Factors

Although these cells were from multiple patients with varying diseases and disparate ages and whose cells were cultured for different lengths of time, there remained underlying patient-based transcriptional patterns that were shared between cells. By grouping the samples by patient, we identified a differentially expressed gene list defined as expressed >60-fold over at least four other patients and performed enrichment analysis for many factors using ToppGene (Chen et al., 2009). For example, there were 351 genes that were enriched in samples from patient 52 (Table S8), excluding those that were specifically linked to a cell type (as listed in Table S1). This gene set showed significant enrichment in being controlled by the ETS2, ELF1, and PEA3 transcription factors, whereas a similar analysis of patient 64 showed regulation by a number of transcription factors, including E2F, E2F1, and many others but also showed strong regulation by miRNAs (for example, 519a, 500, 24, 1284, and 29c; p < 0.05, Bonferroni-corrected). Similar analyses were completed for all

patients. They highlight a variety of shared and distinct control patterns of genes across patients, including precursor microRNA (miR) expression. Interestingly, patients 64 and 85 have very little in common, including exhibiting different diagnoses, genders, and ethnicities. However, both show enrichment of miR29 a, b, and c binding sites in their enriched gene set, suggesting that there is some other environmental or genetic commonality resulting in the need for similar regulation of miR29-regulated genes.

Single-Cell Drug-Induced Transcriptome Modifications

We also attempted to dissect the patterns of gene expression associated with patient drug treatment history. To identify these genes, we compared expression patterns within each drug-treated sample type to non-drug-treated samples using a recently developed single-cell differential expression (SCDE) package (Kharchenko et al., 2014). Because of the inevitable uneven sampling of drug treatments between patients, many of the treatment factors were compounded, and power for distinguishing effects was low. After correction for multiple testing, in endothelial cells, we found association in Keppra and dexamethasone treatment (compounded together) for the genes (adjusted p value in parentheses): GFAP (0.002), CD14(0.041), HTRA3 (0.041), and PDPN (0.041). Other suggestive cell type/ drug/gene differential test results are listed in Table S9. These data will benefit from additional patient samples from similarly drug-treated individuals.

DISCUSSION

Investigating human disease and cellular response to drug therapy is most appropriately performed in humans. Such experiments are most conveniently performed in a culture system for ease of access, solution perfusion, manipulation, and imaging. Until now, there has not been an adult human brain cell culturing system for such studies. Here we report the ability to culture healthy adult primary brain cells for months. Although adult rodent brain cell cultures are difficult to generate, human brain cell cultures did not suffer the same attrition. The cells exhibited healthy morphologies and maintained strong cell type marker RNA expression, suggesting that these cells thrive as their true cell types, not as a shadow of their original robustness. This system allows for the investigation of human disease and drug effects at the single-cell level with the added capability of performing testing in a cell-type-specific manner.

In our desire to analyze single human brain cell transcriptomes and cellular variability in expression profiles, we have deep-sequenced over 300 single cells from adult human brain primary cell cultures, enabling us to identify various cell types. Conventional wisdom would suggest that each type of cell has a set of known markers whose abundant presence consistently defines a cell phenotype (Redwine and Evans, 2002). Recent work in mice using Cre-mediated fluorescent tagging of marker positive cells has begun to highlight the range of cell types in mice expressing each presumptively canonical cell-specific marker (Tasic et al., 2016). Similarly, our study has shown that adult human brain cells also exhibit non-cell-type-restricted expression of many such "canonical" markers; however, other RNAs can serve as appropriate markers. .

This cell type subgroup analysis led us to three conclusions. First, although we have samples from multiple ages, diseases, and drug treatments, we found that the cell type expression profiles were robust enough to stand out among this biological variation. Second, we have a better idea of the variability in gene expression that results in subclasses of cells in the human brain. For example, markers found using populations of cells would suggest a simplicity and almost bimodal (on/off) expression of these genes between cell types that is generally not observed for mRNA in the human system. Often marker genes derived from mouse studies are not highly expressed and are sometimes absent in human cells in the cell type they are presumed to define. A network of genes relevant to the cell's function whose expression determines the cell's fate, including lncRNAs and pri-miRNAs, appears to be more relevant to cell identity rather than the high expression of any one given mouse cell type derived marker. Last, in part by comparing expression patterns in our cultured cells to those harvested acutely from patient brains, we find that there is a resiliency to the cell type expression profiles (Darmanis et al., 2015) and in vivo morphology (Darmanis et al., 2015; Fields, 2013) that is maintained in our cultures that makes this system useful for studying primary oligodendrocytes, endothelial cells, astrocytes, microglia, and multiple classes of interneurons and excitatory neurons, just to name a few.

Morphology alone was not the best differentiator of these human cultured cells. In rodents, by contrast, a trained researcher can easily differentiate between a neuron and astrocyte based on classical shapes. Expanding upon recent findings (Zhang et al., 2016), we found that the morphology of cultured human astrocytes is less distinct and more complex than that of rodent cultured astrocytes. There seems to be a larger number of cell-type-associated morphologies than in rodent cultured cells, possibly to meet the uniquely complex system demands on each cell in the human system. Just as our preconceived notions about marker gene expression based upon rodent studies are more complex in the human system, so too are the expected responsivities of human cells. The transcriptomes of adult human cells suggest a broader range of expression of surface channels and receptors than in the mouse, with a likely broader range of functional responses to stimuli. This is predicted by the higher complexity of networks required to perform higher-level functioning in humans.

In addition to morphological differences between human brain cells and those from lower species, the defining genes that are used as markers as well as those that drive that cell's functional phenotype can be distinct. Although mouse-derived oligodendrocyte markers were the most successful at distinguishing human oligodendrocytes from other cell types (high expression of ~60% of mouse oligodendrocytes markers in the majority of human oligodendrocytes), mouse-derived neuronal markers were less successful (high expression of ~3% of expressed mouse neuronal markers in the majority of human neurons; Cahoy et al., 2008). This may, in part, be due to the greater genetic diversity of human patients in comparison with that of inbred mouse strains. These realities highlight the necessity of using human cells for human disease and drug studies.

In clustering these neurons into neuronal subtypes, we found that the expression profiles were more complex than expected. Interestingly, although cluster 1, shown in Figure 7, had the highest level of SPARC, all other neuronal clusters also expressed the gene, suggesting a continual need and capacity for neural remodeling (Andres et al., 2011). Cluster 1's low level expression of many genes is in line with the possibility that they are highly plastic cells able to adjust as the network needs require. We hypothesize that stem cells have a low level of expression of a wide range of genes, allowing for more rapid differentiation into a mature cell as the need arises, and that these cells clusters exhibit this profile feature. Similarly, cluster 4's mid-level expression of many genes suggests a progenitor cell classification, whereas its strong expression of GAD1 and ABAT suggests that they are inhibitory. Together, this suggests a lack of full commitment to a single cell type, perhaps enabling the plasticity of the human brain to be directed by specific microenvironmental signaling cues. Cluster 2's high expression of calbindin 1 (CALB1) which is involved in synaptic plasticity in excitatory neurons (Chard et al., 1995), and near absence of GAD1 and GAD2 suggest that this group is excitatory, although CALB1 has also been found in GABAergic neurons of the human cortex (del Río and DeFelipe, 1996). Groups 3 and 5 have strong expression of inhibitory genes and low expression of many excitatory genes, suggesting that they are inhibitory neurons; however, there was a small but noticeable amount of expression of traditionally excitatory neuron-linked genes, highlighting the likelihood of cells that co-release multiple neurotransmitters.

This mixed expression again suggests the functional plasticity required by and built into the human nervous system. Some of these cell groups contrast those observed by others (Darmanis et al., 2015), in which five communities of interneurons were composed of cells co-expressing PAX6/RELN, CPLX3/SPARC/SV2C, and a distinct group of PVALB+ cells, among other distinctions. In our adult cells although we found that the majority of PVALB+ cells fell in a single class, it was the class that had low to mid-level expression of many genes. We did not find a strong correlation between PAX6 and RELN, although they were co-expressed in some cases or within the CPLX3/SPARC/SV2C gene combination. In fact, in contrast to SPARC, CPLX3 and SV2C had low and sporadic expression in our analysis. The transcriptional complexity we have observed has been functionally shown by others, where neurons may not express single neurotransmitters. For example, a neuron can express

both glutamate and gamma-aminobutyric acid (GABA) and be responsible for both types of signaling. This co-expression of conventional excitatory and inhibitory markers is abundant in our adult cultured human neurons and suggests a dramatically altered cellular ipseity. The complexity of these data highlight how difficult it is to determine a finite number of transcription-based cellular subclasses that meaningfully exist. At the extreme, each cell is a transcriptional and physiological "unicorn" exhibiting unique transcriptional profiles and physiological interactions with the transcriptome, providing a "hypothesis as to presumptive cell function." The cellular environment plays upon the cell's transcriptome to produce cells of needed function, hence the observed transcriptional plasticity.

Even with the variability of adult human neuronal cell transcriptomes we describe, there appear to be overriding cellular regulatory hierarchies, with the general cell type, such as neuron or endothelial cell, being consistent across patients and diseases, whereas there is also a patient-specific transcriptome that can define an overall patient expression pattern. Because there are many distinguishable cell types, this set of discriminators must exert their function on top of the patient-specific discriminators. These transcriptional networks likely result from hierarchical epigenetic modification of cellular genomes, highlighting the need for a robust single-cell epigenomic platform.

Although our drug-effect analyses were underpowered, we were able to predict some long-lasting (6 weeks post-removal from on-board patient drug therapy) drug-induced changes in cell-type-specific gene expression. This long-lasting effect suggests an epigenetic effect of these drugs on the patient. With more patient-derived samples, this approach may prove to be informative in drug efficacy studies as well as in assessing potential adverse side effects. The techniques in this study are potentially useful as an approach to personalized precision medicine.

We have successfully cultured six major classes of brain cells, identified enriched RNAs in each class, found systems that likely control this gene enrichment, and have predicted alterations in gene expression that are the result of in vivo therapeutic interventions in multiple cell types. Further, it is clear that many RNAs should be designated as cell-type-selective rather than cell-type-specific because the marker will likely be greatly enriched in one cell type but also be expressed in other cell types in a non-defining manner. Primary culturing of adult brain cells from human biopsies has allowed us to capture the range of expression for each cell type and to glimpse the plasticity built into the human system. In addition to providing a model system to study human disease and drug treatments, these data have provided insight into the plasticity and range of phenotypes inherent in human brain cells that are necessary for the proper functioning of the human brain.

EXPERIMENTAL PROCEDURES

Neurosurgery Harvest

Adult human brain tissue was collected at the Hospital of University of Pennsylvania (Institutional Review Board [IRB] #816223) using standard operating procedures for enrollment and consent of patients,. Human brain tissue collection, handling, and de-identification of patient clinical data also followed standard operating procedures. The approximate region of cortex from which the specimen was collected was identified using Brodmann area maps that were then linked to the publicly available Foundational Model of Anatomy Ontology (http://bioportal.bioontology.org/ontologies/FMA; Figure 1A; Figure S1A). Using cortex or hippocampal tissue that was resected as part of a neurosurgical procedure for the treatment of epilepsy or brain tumors, we collected a 5 × 5 × 5 mm block of tissue. This tissue was immediately transferred to a sterile container with ice-cold sucrose aCSF solution (2 mM $CaCl_2$-$2H_2O$, 10 mM glucose, 3 mM KCl, 26 mM $NaHCO_3$, 2.5 mM NaH_2PO_4, 1 mM $MgCl_2$-$6H_2O$, and 202 mM sucrose, with 5% CO_2 and a 95% O_2 gas mixture) for transfer to the laboratory. Sucrose aCSF was oxygenated for at least 1 hr before the scheduled surgery to keep the brain tissue alive during transport. Tissues arrived in the laboratory ~10 min after excision.

Culturing

Brain tissue was digested using papain (20 U, Worthington Biochemical) and incubated for 10–15 min at 37°C, followed by Leupeptin (a papain inhibitor, 100 μM, Sigma-Aldrich) to stop the reaction. After enzymatic dissociation, centrifugation (1,500 rpm for 3 min) followed by gentle mechanical dissociation were performed with a fire-polished glass Pasteur pipette. The cells were counted in an Autocounter (Invitrogen) using trypan blue (1%, Sigma-Aldrich) to exclude dead cells. Cells were plated on poly-L-lysine-coated (0.1 mg/ml, Sigma-Aldrich) 12-mm coverslips at a density of 3×10^4 cells/coverslip. Cultures were incubated at 37°C, 95% humidity, and 5% CO_2 in medium (Neurobasal B, 27%–1%; penicillin/streptomycin, 1%; Thermo-Fisher Scientific). The medium was changed by replacing 50% with fresh medium every 3 days.

Morphology

Images were taken before harvest and analyzed. Six features were scored: gross cell size with three states (0 = small, 1 = medium, 2 = large); cell shape with three states (0 = radial and round, 1 = rod, 2 = radial and amorphous); cell process with three states (0 = no processes, 1 = uni-directional processes, 2 = multi-directional processes); process complexity with three states (0 = no process, 1 = simple processes, 2 = complex processes); cell margin with two states (0 = smooth, 1 = rough); and cytosol/nuclear ratio with two states (0 = small, 1 = large).

Amplification, Library, Sequencing, and Read Processing

Single cells were harvested using a microcapillary pipette (Morris et al., 2011). Samples were snap-frozen until processing. Three rounds of standard aRNA amplification were completed, followed by Truseq stranded library generation as outlined by Illumina without the initial fragmentation step (Eberwine et al., 1992; Hashimshony et al., 2012). Samples were sequenced using either an Illumina Hiseq 2500 or Nextseq 500. After sequencing, reads were de-multiplexed with the CASAVA software package, version 1.8.2 (Illumina). Reads were processed with the PennSCAP-T Pipeline (https://github.com/safisher/ngs). Sequence alignments were performed with Spliced Transcripts Alignment to Reference (STAR) (Dobin et al., 2013). Exonic reads that uniquely mapped to the Genome Reference Consortium Human Reference 38 (GRCh38/Hg38) were processed with VERSE (BioRXiv) using a hierarchical assignment scheme and the GENCODE 21 transcriptome. Samples with more than two million uniquely mapping reads and more than 20% exonic mapping were included in the analysis. Reads were normalized for differences in sequencing depth across samples prior to all further analyses by scaling raw read counts by sample-specific size factors as estimated in DESeq (Anders and Huber, 2010).

Computational Analyses
GTEx

Similarity in gene expression between single cells in this study and multiple tissues in the Genotype-Tissue Expression (GTEx) dataset (GTEx Consortium, 2013), a large-scale public resource for human gene expression across tissues, was determined by digitizing the read counts of top-ranked highly expressed genes for each cell in this study according to several thresholds (from 0 to 500, interval 10). We calculated the median across samples within each tissue and digitized the read counts of top genes based on the median. We computed the Jaccard similarity coefficient of the digitized read counts between each single cell and each tissue.

PCA

Principal component analysis was performed using the R package in IDV (http://kim.bio.upenn.edu/software/idv.shtml). Clustering was performed on the full transcriptome, excluding genes with zero expression in all cells to eliminate clustering based on null expression. The plot was then color-coded to show the covariates for each comparison (marker-based cell type, DIV, patient ID).

K-Means

K-means clustering of markers, pri-miRNAs, and lncRNAs was performed using the R package in IDV (http://kim.bio.upenn.edu/software/idv.shtml). To assess the appropriate number of clusters for the k-means algorithm, we carried out 2-fold and 5-fold cross-validation experiments for k = 2–15. For each cross-validation experiment, we randomly split the data into training and test and computed the k-means centroids on the training data. The test data membership was fit to the closest centroid from the training data, and the sum-within-cluster distance of the test data were defined as prediction error. Randomized cross-validation experiments were repeated ten times for each k. Table S4 shows the prediction error and its SE as a function of k. We established an elbow in the reduction of prediction error between k = 6 and k = 8 (cf. Miligan and Cooper, 1985). We expected seven different types of cells (excitatory neurons, inhibitory neurons, astrocytes, microglia, oligodendrocytes, oligodendrocyte precursor cells, and endothelial cells). To allow for additional cell types, we chose k = 8 as for the cluster model. To identify the best k-means clusters, we tuned the nstart option of the kmeans function in the R package (nstart = 1,000).

Gene Annotations

The list of lncRNAs was the union of the human lncRNAs found in the lncRNA database (Amaral et al., 2011) and genes beginning with "LINC" annotated in gencode.v21 as Level-1,2 exons. Significant enrichment of miRNA and transcription factor binding sites in the enriched genes for each patient and cell type was determined using the ToppGene suite (Chen et al., 2009). Significant enrichment of GO terms in genes enriched in each cell class was determined using the GO Enrichment Analysis Suite as curated by the Gene Ontology Consortium (Gene Ontology Consortium, 2015).

Pathway Analysis

To find pathway-derived subgroups, we used neuron-related functions, immune-related functions, and cell cycle- and signaling-related pathways from the Molecular Signatures Database (MSigDB) v5.1 (Liberzon et al., 2011). We focused on curated pathway gene sets from online pathway databases such as BioCarta, Kyoto Encyclopedia of Genes and Genomes (KEGG), and Reactome. Pathway activity is defined as the percentage of genes with more than ten normalized reads in a pathway. The k-means method was used for clustering pathway activity to identify pathway-dependent cell types.

PLDA Expression Analysis

We identified the genes that best discriminate between each cell type and other cell types using PLDA (Witten and Tibshirani, 2011). The PLDA is a classification technique that, by adopting sparsity constraints, achieves feature selection. When the standard estimate for the within-class covariance matrix is singular, the usual discriminant rule cannot be applied. By considering L1 and fused lasso penalties on the discriminant vectors, we can solve the problem efficiently. In this study, the PLDA was implemented in the R package "penalizedLDA." The lasso penalty tuning parameter lambda = 1e−04 was used from several trials, and the number of discriminant vectors was equal to the number of classes minus 1 because it must be no greater than the number (classes − 1).

Random Sampling and Sequencing Depth

To assess the effect of a reduced sequencing depth, we used "sample" (https://travis-ci.org/alexpreynolds/sample), a function based on the algorithm developed by Jeffrey S. Vitter that randomly samples read pairs from a SAM file. We generated a series of down-sampled datasets where subsets of 1 million, 0.5 million, and 0.1 million mapped reads were randomly sampled from the original dataset in this study, repeating 25 times for each down-sampled level. To see how many genes are expressed in the original data and down-sampled datasets, we compared the number of detected genes between the sets. Detected genes are defined as genes with a raw read > 0 in at least 70% of samples for each cell type. To examine whether reduced sequencing depth affects prediction accuracy, we applied the PLDA function generated by the original dataset to the down-sampled dataset.

Differential Expression

For differential expression analysis of pri-miRNA and lncRNAs, we carried out one-way ANOVA using cell type as a factor. For multiple corrections, we applied the Benjamini-Hochberg method.

Morphology Analysis

For morphology grouping and covariance analysis, cell size, cell process, and process complexity were treated as ordered states, whereas the remaining features were treated as unordered states. Cells with identical feature scores were grouped as a single unit, resulting in a total of 31 morphologically distinct groups. Morphological trees were computed with the maximum parsimony method (Swofford et al., 1996) using the TBR heuristic search option in the program PAUP* (Swofford, 2003).

Drug-Mediated Differential Expression

SCDE analysis was used to find differences in mean gene expression between drug-treated cells and drug-untreated cells. Cells were given a specific cell type using the Benjamini-Hochberg method for multiple testing correction (Kharchenko et al., 2014).

ACCESSION NUMBERS

The accession number for the RNA sequencing data reported in this paper is dbGaP: phs000833.v5.p1.

SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures and nine tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2016.12.066.

AUTHOR CONTRIBUTIONS

ACKNOWLEDGMENTS

REFERENCES

Amaral, P.P., Clark, M.B., Gascoigne, D.K., Dinger, M.E., and Mattick, J.S. (2011). lncRNAdb: a reference database for long noncoding RNAs. Nucleic Acids Res. 39, D146–D151.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. 11, R106.

Andres, R.H., Horie, N., Slikker, W., Keren-Gill, H., Zhan, K., Sun, G., Manley, N.C., Pereira, M.P., Sheikh, L.A., McMillan, E.L., et al. (2011). Human neural

stem cells enhance structural plasticity and axonal transport in the ischaemic brain. Brain 134, 1777–1789.

Cahoy, J.D., Emery, B., Kaushal, A., Foo, L.C., Zamanian, J.L., Christopherson, K.S., Xing, Y., Lubischer, J.L., Krieg, P.A., Krupenko, S.A., et al. (2008). A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. J. Neurosci. 28, 264–278.

Chard, P.S., Jordán, J., Marcuccilli, C.J., Miller, R.J., Leiden, J.M., Roos, R.P., and Ghadge, G.D. (1995). Regulation of excitatory transmission at hippocampal synapses by calbindin D28k. Proc. Natl. Acad. Sci. USA 92, 5144–5148.

Chen, J., Bardes, E.E., Aronow, B.J., and Jegga, A.G. (2009). ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. Nucleic Acids Res. 37, W305-11.

Chen, S., Tran, S., Sigler, A., and Murphy, T.H. (2011). Automated and quantitative image analysis of ischemic dendritic blebbing using in vivo 2-photon microscopy data. J. Neurosci. Methods 195, 222–231.

Darmanis, S., Sloan, S.A., Zhang, Y., Enge, M., Caneda, C., Shuer, L.M., Hayden Gephart, M.G., Barres, B.A., and Quake, S.R. (2015). A survey of human brain transcriptome diversity at the single cell level. Proc. Natl. Acad. Sci. USA 112, 7285–7290.

del Río, M.R., and DeFelipe, J. (1996). Colocalization of calbindin D-28k, calretinin, and GABA immunoreactivities in neurons of the human temporal cortex. J. Comp. Neurol. 369, 472–482.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21.

Dueck, H., Khaladkar, M., Kim, T.K., Spaethling, J.M., Francis, C., Suresh, S., Fisher, S.A., Seale, P., Beck, S.G., Bartfai, T., et al. (2015). Deep sequencing reveals cell-type-specific patterns of single-cell transcriptome variation. Genome Biol. 16, 122.

Eberwine, J., Yeh, H., Miyashiro, K., Cao, Y., Nair, S., Finnell, R., Zettel, M., and Coleman, P. (1992). Analysis of gene expression in single live neurons. Proc. Natl. Acad. Sci. USA 89, 3010–3014.

Fields, R.D. (2013). Neuroscience: Map the other brain. Nature 501, 25–27.

Fish, K.N., Sweet, R.A., and Lewis, D.A. (2011). Differential distribution of proteins regulating GABA synthesis and reuptake in axon boutons of subpopulations of cortical interneurons. Cereb. Cortex 21, 2450–2460.

Friedmann-Morvinski, D., Bushong, E.A., Ke, E., Soda, Y., Marumoto, T., Singer, O., Ellisman, M.H., and Verma, I.M. (2012). Dedifferentiation of neurons and astrocytes by oncogenes can induce gliomas in mice. Science 338, 1080–1084.

Gene Ontology Consortium (2015). Gene Ontology Consortium: going forward. Nucleic Acids Res. 43, D1049–D1056.

GTEx Consortium (2013). The Genotype-Tissue Expression (GTEx) project. Nat. Genet. 45, 580–585.

Hashimshony, T., Wagner, F., Sher, N., and Yanai, I. (2012). CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. Cell Rep. 2, 666–673.

Hawrylycz, M., Miller, J.A., Menon, V., Feng, D., Dolbeare, T., Guillozet-Bongaarts, A.L., Jegga, A.G., Aronow, B.J., Lee, C.K., Bernard, A., et al. (2015). Canonical genetic signatures of the adult human brain. Nat. Neurosci. 18, 1832–1844.

Hergenreider, E., Heydt, S., Tréguer, K., Boettger, T., Horrevoets, A.J., Zeiher, A.M., Scheffer, M.P., Frangakis, A.S., Yin, X., Mayr, M., et al. (2012). Atheroprotective communication between endothelial cells and smooth muscle cells through miRNAs. Nat. Cell Biol. 14, 249–256.

Hernandez, M., and Casaccia, P. (2015). Interplay between transcriptional control and chromatin regulation in the oligodendrocyte lineage. Glia 63, 1357–1375.

Kharchenko, P.V., Silberstein, L., and Scadden, D.T. (2014). Bayesian approach to single-cell differential expression analysis. Nat. Methods 11, 740–742.

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.

Li, Y., Li, A., Glas, M., Lal, B., Ying, M., Sang, Y., Xia, S., Trageser, D., Guerrero-Cázares, H., Eberhart, C.G., et al. (2011). c-Met signaling induces a reprogramming network and supports the glioblastoma stem-like phenotype. Proc. Natl. Acad. Sci. USA 108, 9951–9956.

Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P., and Mesirov, J.P. (2011). Molecular signatures database (MSigDB) 3.0. Bioinformatics 27, 1739–1740.

Mercer, T.R., Qureshi, I.A., Gokhan, S., Dinger, M.E., Li, G., Mattick, J.S., and Mehler, M.F. (2010). Long noncoding RNAs in neuronal-glial fate specification and oligodendrocyte lineage maturation. BMC Neurosci. 11, 14.

Miligan, G.W., and Cooper, M.C. (1985). An examination of procedures for determining the number of clusters in a data set. Psychometrika 50, 159–179.

Miyashiro, K., Dichter, M., and Eberwine, J. (1994). On the nature and differential distribution of mRNAs in hippocampal neurites: implications for neuronal functioning. Proc. Natl. Acad. Sci. USA 91, 10800–10804.

Morris, J., Singh, J.M., and Eberwine, J.H. (2011). Transcriptome analysis of single cells. J. Vis. Exp. 50, 2634.

Redwine, J.M., and Evans, C.F. (2002). Markers of central nervous system glia and neurons in vivo during normal and pathological conditions. Curr. Top. Microbiol. Immunol. 265, 119–140.

Swofford, D.L. (2003). PAUP* Phylogenetic Analysis Using Parsimony (*and Other Methods) Version 4.0b10 for UNIX (Sunderland, MA: Sinauer Associates).

Swofford, D.L., Olsen, G.J., Waddell, P.J., and Hillis, D.M. (1996). Phylogenetic inference. In Molecular Systematics, D.M. Hillis, C. Moritz, and B.K. Mable, eds. (Sunderland: Sinauer Associates), pp. 407–514.

Takamori, S. (2006). VGLUTs: 'exciting' times for glutamatergic research? Neurosci. Res. 55, 343–351.

Tasic, B., Menon, V., Nguyen, T.N., Kim, T.K., Jarsky, T., Yao, Z., Levi, B., Gray, L.T., Sorensen, S.A., Dolbeare, T., et al. (2016). Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. Nat. Neurosci. 19, 335–346.

Vlachos, I.S., Kostoulas, N., Vergoulis, T., Georgakilas, G., Reczko, M., Maragkakis, M., Paraskevopoulou, M.D., Prionidis, K., Dalamagas, T., and Hatzigeorgiou, A.G. (2012). DIANA miRPath v.2.0: investigating the combinatorial effect of microRNAs in pathways. Nucleic Acids Res. 40, W498-504.

Vogelstein, B., Papadopoulos, N., Velculescu, V.E., Zhou, S., Diaz, L.A., Jr., and Kinzler, K.W. (2013). Cancer genome landscapes. Science 339, 1546–1558.

Witten, D.M., and Tibshirani, R. (2011). Penalized classification using Fisher's linear discriminant. J. R. Stat. Soc. Series B Stat. Methodol. 73, 753–772.

Xia, M., and Zhu, Y. (2015). The regulation of Sox2 and Sox9 stimulated by ATP in spinal cord astrocytes. J. Mol. Neurosci. 55, 131–140.

Zeisel, A., Muñoz-Manchado, A.B., Codeluppi, S., Lönnerberg, P., La Manno, G., Juréus, A., Marques, S., Munguba, H., He, L., Betsholtz, C., et al. (2015). Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. Science 347, 1138–1142.

Zhang, Y., Pak, C., Han, Y., Ahlenius, H., Zhang, Z., Chanda, S., Marro, S., Patzke, C., Acuna, C., Covy, J., et al. (2013). Rapid single-step induction of functional neurons from human pluripotent stem cells. Neuron 78, 785–798.

Zhang, Y., Sloan, S.A., Clarke, L.E., Caneda, C., Plaza, C.A., Blumenthal, P.D., Vogel, H., Steinberg, G.K., Edwards, M.S., Li, G., et al. (2016). Purification and Characterization of Progenitor and Mature Human Astrocytes Reveals Transcriptional and Functional Differences with Mouse. Neuron 89, 37–53.