Fixed single-cell transcriptomic characterization of human radial glial diversity

Elliot R Thomsen[1,4], John K Mich[1,4], Zizhen Yao[1,4], Rebecca D Hodge[1], Adele M Doyle[2], Sumin Jang[2], Soraya I Shehata[1], Angelique M Nelson[1], Nadiya V Shapovalova[1], Boaz P Levi[1] & Sharad Ramanathan[1–3]

The diverse progenitors that give rise to the human neocortex have been difficult to characterize because progenitors, particularly radial glia (RG), are rare and are defined by a combination of intracellular markers, position and morphology. To circumvent these problems, we developed Fixed and Recovered Intact Single-cell RNA (FRISCR), a method for profiling the transcriptomes of individual fixed, stained and sorted cells. Using FRISCR, we profiled primary human RG that constitute only 1% of the midgestation cortex and classified them as ventricular zone–enriched RG (vRG) that express ANXA1 and CRYAB, and outer subventricular zone–localized RG (oRG) that express HOPX. Our study identified vRG and oRG markers and molecular profiles, an essential step for understanding human neocortical progenitor development. FRISCR allows targeted single-cell profiling of any tissues that lack live-cell markers.

© 2015 Nature America, Inc. All rights reserved.

Several progenitor cell types underpin human brain development. RG and intermediate progenitor cells (IPCs) are cortical progenitors that reside in the ventricular zone of the cortex (**Fig. 1**)[1–5]. RG are bipolar epithelial cells that extend an apical endfoot to the ventricular surface and a basal process to the pial surface, and give rise to glia and neurons. In contrast, IPCs are strictly neurogenic, lack epithelial morphology, and have a more limited capacity for proliferation and self-renewal[1,3–5]. During a prolonged period of neurogenesis, a region of proliferating progenitors called the outer subventricular zone (oSZ)[2,5,6] expands in the developing brain. The oSZ contains IPCs as well as outer RG (oRG) that express the same canonical transcription factors as RG in the ventricular zone (vRG) but are distinguished by their position, the lack of an apical endfoot and the maintenance of a basal process (**Fig. 1a**)[1,7,8]. oRG are hypothesized to drive the dramatic cortical expansion observed in gyrified brains such as the human brain[3,5,9].

Understanding the molecular diversity of human RG progenitors is an essential first step to determine (i) whether discrete RG populations produce specific mature cell types, and (ii) what molecular events drive the formation of human-specific features such as oRG and the oSZ. Owing to the rarity of RG progenitors, human RG analysis has been restricted to morphology with a few histological markers to confirm cell identity (**Fig. 1b**)[1,7,8], molecular characterization of microdissected tissue containing an unknown variety of cell types[10,11], or live marker–sorted cells whose purity is unknown[12,13]. The lack of markers for RG progenitor subtypes has limited our ability to understand human corticogenesis.

Characterizing the full diversity of rare RG progenitors requires transcriptional profiles of large numbers of single cells, ideally from enriched subpopulations. RG express SOX2 and PAX6, and lack EOMES (also known as TBR2), whereas IPCs can express all three genes[2,4,5]. Sorting cells by these intracellular markers requires fixation, permeabilization and staining; steps that typically lead to highly degraded mRNA and render transcriptomic profiling impossible. Protocols have emerged recently for transcriptional profiling of fixed, stained and sorted cells, but have only been reported for samples of $\geq 10^5$ cells, and never for single cells[14–18].

Here we present FRISCR for RNA isolation and transcriptomic profiling of fixed, permeabilized, stained and sorted single cells. We show that fixation and purification introduce little bias and yield gene expression data similar to that from living cells. Using this technique, we prospectively isolated and profiled single RG from primary human prenatal neocortex. Gene expression analysis identified RG subpopulations that correspond to human oRG and vRG based on position in primary mid-gestation human cortex, and identified new molecular markers that distinguish oRG from vRG cells. FRISCR provides an important new tool for single-cell profiling of human primary tissues, including rare populations in the brain and tissues that lack live-cell markers.

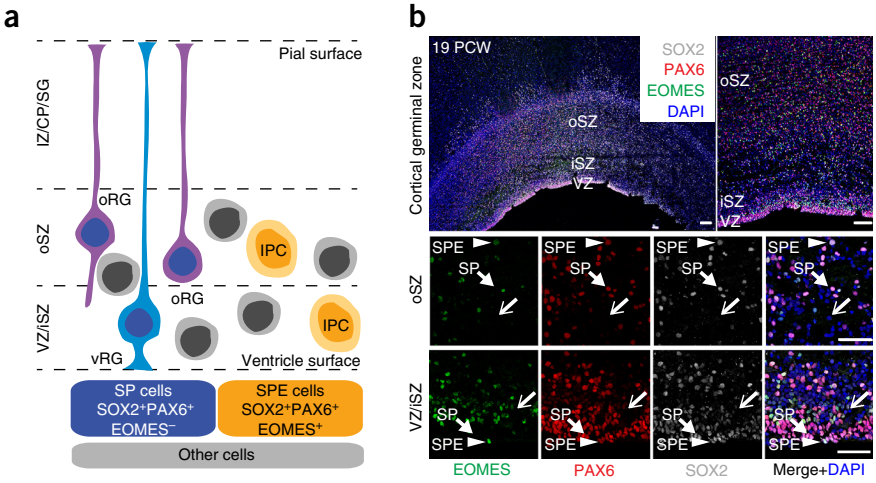
RESULTS

Development of FRISCR

We optimized RNA extraction from fixed cells by always handling cells with RNase-free reagents in the presence of RNase inhibitor and reverse cross-linking RNA at 56 °C for 1 h rather than the standard 80 °C[14]. As comparable methods have only been validated for $>10^5$ cells, we compared fixed human embryonic stem cell (hESC) RNA to that purified from live hESCs using 10,000-, 1,000- and 100-cell samples (**Supplementary Fig. 1a**). Analysis of total RNA showed high RNA integrity numbers from batches of fixed cells using this protocol (**Fig. 2a** and **Supplementary Fig. 1b**) with undetectable RNA losses (**Fig. 2b**). Fixation did not introduce

[1]Allen Institute for Brain Science, Seattle, Washington, USA. [2]Molecular and Cellular Biology, Harvard University, Cambridge, Massachusetts, USA. [3]School of Engineering and Applied Science, Harvard University, Cambridge, Massachusetts, USA. [4]These authors contributed equally to this work. Correspondence should be addressed to B.P.L. (boazl@alleninstitute.org) or S.R. (sharad@cgr.harvard.edu).

RECEIVED 6 MAY; ACCEPTED 23 SEPTEMBER; PUBLISHED ONLINE 16 NOVEMBER 2015; DOI:10.1038/NMETH.3629



c

Cell type	Location	FACS pheno.
Outer radial glia (oRG)	Outer subventricular zone (oSZ)	SP
Intermediate progenitor cell (IPC)	Ventricular zone (VZ), inner subventricular zone (iSZ), oSZ	SPE
Ventricular radial glia (vRG)	VZ	SP

Figure 1 | Human cortical progenitors are diverse and intermixed during development. (**a**) The progenitor compartment includes a mixture of vRG (light blue), oRG (purple), IPCs (orange) and other cortical cell types (gray). RG identified by antibody staining for known markers are SP cells (dark blue nuclei) and IPCs are SPE cells (dark orange nuclei). IZ/CP/SG, intermediate zone/cortical plate/subgranular zone; slashes indicate that in this model, all of those regions are collapsed. (**b**) Stitched low-magnification immunocytochemistry images of 19 PCW germinal zones (top left), or individual micrograph (top right) of the VZ, iSZ and oSZ stained by DAPI, and for EOMES, PAX6 and SOX2. Scale bars, 100 µm. High-magnification micrograph of oSZ region (middle) and VZ/iSZ region (bottom) showing SP (filled arrows) and SPE (arrowheads) cells. Many cells in the VZ, iSZ and oSZ lack progenitor markers and are unknown cell types (open arrows). Scale bars, 50 µm. (**c**) Common human cortical progenitors, including flow cytometry phenotypes (pheno.).

bias in mRNA detection as assessed by quantitative reverse-transcriptase PCR (qRT-PCR); we detected high correlations between live and fixed cells for every gene tested (**Fig. 2c**). We sought to extend this technique to single cells by improving mRNA recovery. Oligo dT25 beads gave better recovery and could be eluted in low volumes after purifying RNA from the reverse cross-linking buffer (**Supplementary Fig. 1c**). The low volume allowed us to apply the entire sample directly to the SmartSeq2 RNA sequencing reaction.

We called this method FRISCR (**Fig. 3a**). To evaluate our approach, we sorted live or fixed H1 hESCs and prepared mRNA by either standard Triton X-100 lysis (TL) or FRISCR. Sequencing library preparation gave comparable amounts of cDNA from individual fixed and live cells (**Fig. 3b** and **Supplementary Fig. 2**), and sequencing (subsampled to five million total reads) produced read alignments that were poor for fixed cells prepared with TL, indicative of much lower mRNA input, or comparable to those with live cells when prepared using FRISCR (**Fig. 3c** and **Supplementary Fig. 2**). The frequency of reads mapping to different transcript classes and number of genes detected per cell was also similar between live and fixed cells prepared using FRISCR (**Fig. 3c** and **Supplementary Fig. 2a**,**c**). Reads across all genes showed a similar 3′ to 5′ bias (**Fig. 3d**); however, fixed cells showed an increased 3′ read bias with longer transcripts (**Supplementary Fig. 2g**). Spearman correlations of the expression of all genes did not discriminate live from fixed cells (**Fig. 3e**), and expression of only two genes in the genome was differentially detected between sets of single cells (**Fig. 3f** and **Supplementary Fig. 3**).

Analysis of External RNA Controls Consortium (ERCC) spike-in mRNAs showed similar linear mRNA amplification by TL and FRISCR (**Supplementary Fig. 2d**), but TL samples showed moderately higher estimated sensitivity compared to FRISCR, recovering ~25% versus ~16% of transcripts, respectively (**Supplementary Fig. 2e**). Nevertheless, FRISCR- and TL-generated single-cell libraries were not separated when clustered by Spearman correlations (**Fig. 3e**), and just over 30 genes were differentially expressed between the two methods, of which many are noncoding or irregularly polyadenylated (**Fig. 3f** and **Supplementary Fig. 3**)[19]. No method showed major changes in G+C coverage, chromosomal transcript representation bias (**Supplementary Fig. 2h**,**i**) or rates of alignment to mRNA (**Fig. 3c**). We detected a subtle increase in the mutation rate in FRISCR-prepared live cells, but preparation using FRISCR substantially reduced the mutation rate of fixed cells compared to preparation with standard TL (**Supplementary Fig. 2f**).

FRISCR profiling of primary RG diversity

We leveraged the ability of FRISCR to profile rare and unmarked cell types by examining the diversity of RG progenitors from primary fetal human cortex. Antibody staining of human cortical

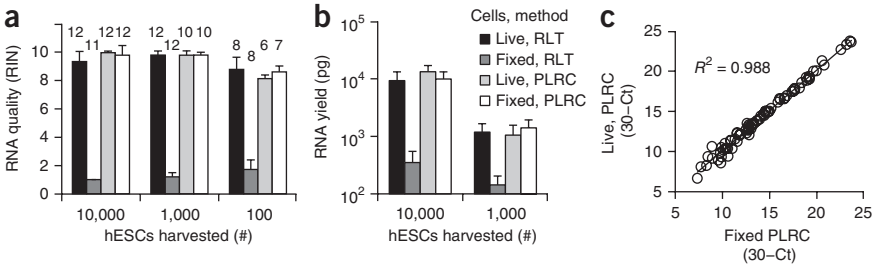
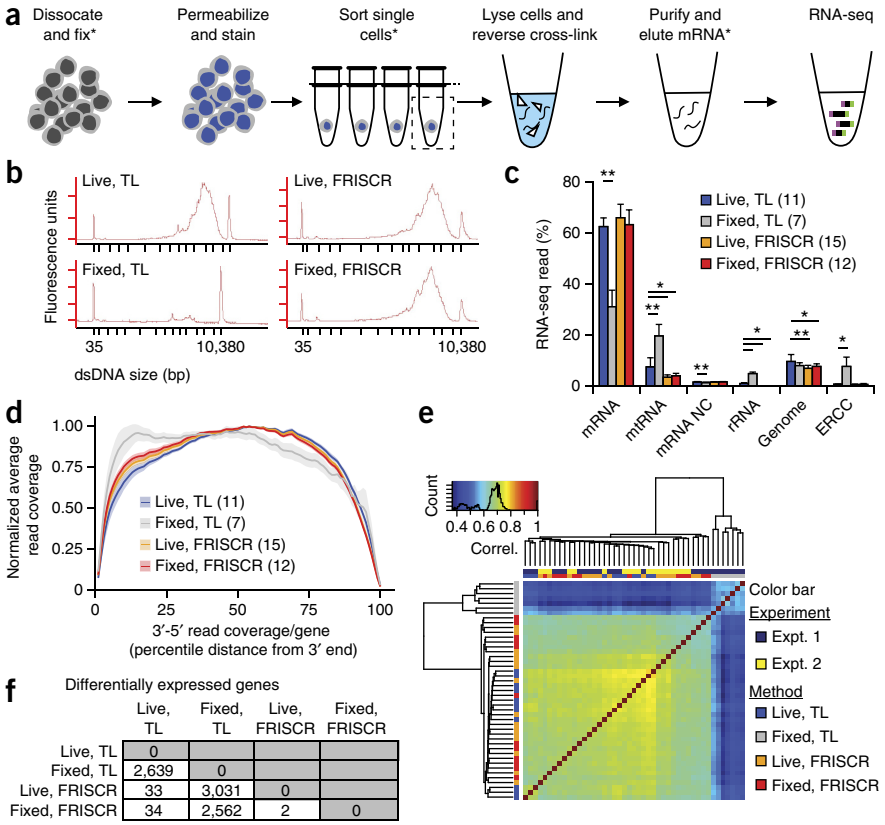


Figure 2 | RNA quality and yield from fixed and sorted hESCs. (**a**) RNA integrity numbers (RINs) of RNA extracted from sorted H1 hESCs. Numbers of biological replicates, run over 2–3 experiments, are indicated above bars. RNA from live or fixed cells was harvested using the Qiagen microRNeasy Kit (RLT), or a modified fixed-cell method with protease lysis and reverse cross-linking (PLRC). (**b**) Yield of column-purified total RNA from H1 hESCs. n = 12 biological replicates, except 1,000-cell PLRC (n = 10 replicates); processed over three independent experiments. Bars (**a**,**b**) show mean ± s.d. (**c**) Expression of 82 genes by qRT-PCR of 10,000 live or fixed hESC samples processed with PLRC. n = 7–8 biological replicates across two independent experiments. Only genes detected in at least three replicates from both conditions are shown.

© 2015 Nature America, Inc. All rights reserved.



	Live, TL	Fixed, TL	Live, FRISCR	Fixed, FRISCR
Live, TL	0			
Fixed, TL	2,639	0		
Live, FRISCR	33	3,031	0	
Fixed, FRISCR	34	2,562	2	0

Figure 3 | Single-cell mRNA purification and amplification from fixed hESCs is similar to that from live cells using FRISCR. (**a**) Steps in the FRISCR method. Asterisks mark potential stopping points. (**b–f**) Analysis of live or fixed single sorted H1 hESCs were processed by Triton-X100 lysis (TL) or FRISCR. Representative Bioanalyzer traces of amplified cDNA (**b**). Alignment of sequencing reads from cells processed by the four experimental conditions (**c**). $*P < 0.05$ and $**P < 0.001$ by one-way analysis of variance (ANOVA) followed by Tukey's post-hoc test. NC, noncoding. Bars show mean ± s.d. Biological replicates in **d–f** are shown in parentheses in the legend for **d**, processed over two independent experiments. Analysis in **d–f** is on 5 million subsampled reads. Read coverage across predicted transcript lengths (**d**). Solid line is the average, and shading illustrates s.d. Hierarchical clustering of single hESCs based on pairwise Spearman correlation of all expressed genes (**e**). (**f**) Number of differentially expressed genes between hESCs from the different conditions calculated by DESeq with adjusted $P < 0.01$.

tissues at 15 to 19 post-conception weeks (15–19 PCW) reproducibly identified early IPCs (SOX2$^+$PAX6$^+$EOMES$^+$; referred to as 'SPE'), and RG (SOX2$^+$PAX6$^+$EOMES$^-$; 'SP') (**Fig. 1**). We used flow cytometry and FRISCR to collect mRNA from single SP cells from four primary tissues (14, 16, 18 and 19 PCW) and single SPE cells from two primary tissues (18 and 19 PCW) (**Fig. 4a**, and **Supplementary Figs. 4** and **5a**). We profiled cells in G_0–G_1 stages to reduce cell cycle–dependent gene expression variation and used DAPI staining, cell size and light scattering properties to exclude debris and doublets (**Supplementary Fig. 4a**). Although RG and IPC progenitors are enriched in the germinal zones, they are rare when the entire cortical thickness is prepared: 3.2 ± 0.9% (mean ± s.d.) of cells were SOX2$^+$PAX6$^+$, and this population was further divided into EOMES$^+$ cells (SPE; 2.1 ± 0.6%) and EOMES$^-$ cells (SP; 1.0 ± 0.3%) (**Fig. 4a** and **Supplementary Fig. 4a**). We then amplified cDNA from 207 SP cells and 48 SPE cells using SmartSeq2 (ref. 20). Primary cells showed more low-molecular-weight cDNA than hESCs (**Fig. 3b** and **Supplementary Fig. 5**), perhaps indicating some RNA degradation due to ≥4 h postmortem intervals in these clinical specimens.

We sequenced single-cell libraries at low read depth on the MiSeq (Illumina) (**Supplementary Table 1** and **Supplementary Data 1** and **2**). Read alignment proportions were consistent in all but a few cells, which may have displayed poor mapping owing to primer dimers or contaminating dead cells (**Supplementary Fig. 5c**). We included in the analysis libraries with more than 10,000 mapped reads, >20% reads mapping to mRNA and detectable *GAPDH* expression (157 SP and 29 SPE cells; **Supplementary Fig. 5d** and **Supplementary Table 1**). We detected ~3,000–4,000 unique genes per progenitor cell (**Supplementary Fig. 5e**), similar to an independent study of live prenatal human single cells[21]. As expected, *PAX6* and *SOX2* were expressed in most SP and SPE cells, and *EOMES* mRNA was enriched in SPE versus SP cells (86% and 22%, respectively; **Fig. 4b**). Principal-component analysis and hierarchical clustering of genes with variance above technical noise demonstrated SP cells partitioning away from SPE cells independently of read depth, cortical sample or mapping percentage, despite the low read depth and partial sample degradation (**Supplementary Fig. 5f,g**). We did not observe this partition after correlation with all expressed genes or with only ERCC spike-in RNA reads (**Supplementary Fig. 5g**). Lastly, our cells clustered with previously identified RG and IPCs profiled from an independent study on live human cortical cells (**Supplementary Fig. 6a**)[21], and largely lacked expression of common contaminating cell markers (**Supplementary Fig. 6b**). Our data demonstrate that FRISCR can be used to profile rare primary human neuronal progenitors.

The fact that we profiled an order of magnitude more RG and IPCs than had been done in previous transcriptomic studies[21]

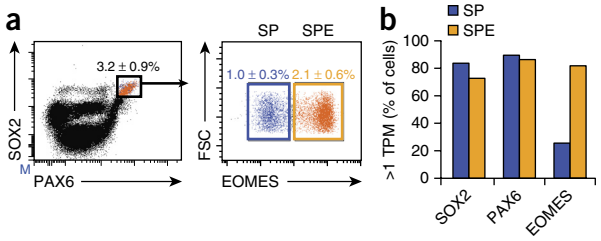


Figure 4 | FRISCR allows profiling of primary human cortical progenitors. (**a**) Frequency of gated cells compared to total cortical cells. Values shown are mean ± s.d.; $n = 3$ independent experiments. Left, only G_0–G_1 phase singlet cortical cells gated with DAPI. Right, only SOX2$^+$PAX6$^+$ cells gated on the left plot. (**b**) Frequency of detection of three genes. TPM, transcripts per million.

© 2015 Nature America, Inc. All rights reserved.

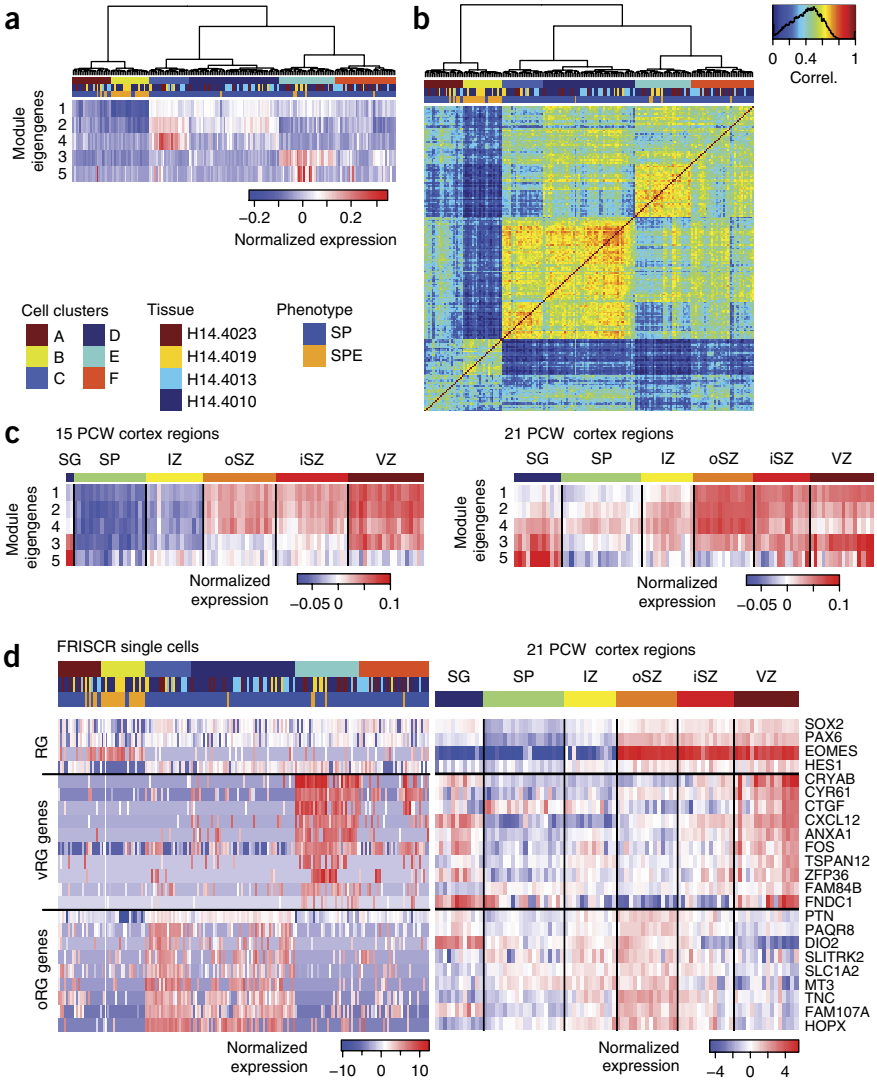


Figure 5 | Identification of human cortical progenitor cell types with FRISCR. (**a**) Module eigengene values derived from WGCNA enrichment analysis. Expression is normalized across all samples. (**b**) Spearman correlations from individual cells by genes identified from the 5 WGCNA modules. (**c**) Correlation of single cell–derived WGCNA module eigengenes with 15 PCW (left) and 21 PCW (right) gene expression data from the BrainSpan Atlas of the Developing Human Brain[11]. Subgranular zone (SG), subplate (SP), intermediate zone (IZ), outer subventricular zone (oSZ), inner subventricular zone (iSZ) and ventricular zone (VZ). (**d**) Genes differentially expressed between cell clusters C and D (oRG) versus cluster E (vRG) from FRISCR-prepared single cells (left), or differentially expressed between oSZ and VZ regions of 21 PCW human tissues (right). Common progenitor markers are included at top for reference. Color bars for **a**,**b** and the FRISCR single cell plot in **d** are indicated at the bottom of **a**.

allowed us to characterize progenitor heterogeneity. We identified five modules of coexpressed genes from high-variance genes by weighted gene coexpression network analysis (WGCNA)[22] (**Fig. 5a** and **Supplementary Fig. 7**). We excluded two modules from clustering that contained many cell cycle–related genes; these modules were enriched in SPE cells, which are twice as likely to be in S–G_2–M phase (**Supplementary Fig. 4b**). We used genes populating the five WGCNA modules to group cells into six clusters labeled A–F (**Fig. 5a**,**b** and **Supplementary Fig. 8**), all but one containing cells from each cortical tissue. Module 1 reflected the division of the majority of RG from IPC cells, and contained canonical markers of RG including *VIM*, and of IPCs including *EOMES*, *HES6* and *NEUROG1* (**Fig. 5a** and **Supplementary Fig. 7**). Modules 2–5 revealed four RG subpopulations as seen by module eigengene values (**Fig. 5a**). Cluster E cells were enriched for module 3 genes including immediate early genes *EGR1* and *FOS*[21], and the genes *CXCL12*, *ANXA1* and *CYR61* (**Supplementary Figs. 7** and **8**). Flow cytometry confirmed that a substantial number of SP cells, but few SPE cells, stained positively for CYR61 (**Supplementary Fig. 9**). A subset of cluster E cells also expressed many genes in module 5 (**Supplementary Figs. 7** and **8**). Cells in clusters C and D expressed higher levels of module 2 and 4 genes than other cell clusters did, but not of module 3 and 5 genes. Module 2 is composed of genes such as *FAM107A*, *HOPX* and *SLCO1C1*; and module 4 genes were enriched specifically in cluster C cells. These analyses revealed molecular diversity within the human RG compartment that had not been previously appreciated to our knowledge.

New markers distinguish vRG and oRG cells

To further characterize this diversity, we compared RG gene expression signatures to data in anatomical atlases. The module 2 and 4 eigengenes were enriched in the oSZ relative to the ventricular zone in 21 PCW cortex (unpaired *t*-test, $P = 6.3 \times 10^{-11}$ for module 2 and $P = 1.5 \times 10^{-19}$ for module 4) but not at earlier time points in the BrainSpan Atlas of the Developing Human Brain[11] (**Fig. 5c** and **Supplementary Fig. 10**). The module 3 eigengene was enriched in the ventricular zone at all time points ($P = 3.9 \times 10^{-15}$, $P = 6.2 \times 10^{-12}$, and $P = 7.8 \times 10^{-10}$ for 15, 16, and 21 PCW tissues), whereas module 5 was significantly enriched only in the ventricular zone in 21 PCW tissue ($P = 4.1 \times 10^{-6}$; **Fig. 5c** and **Supplementary Fig. 10a**). Strong *HOPX* expression initiates around embryonic day 70 in macaque according to the NIH Blueprint Non-Human Primate Atlas, coinciding with the emergence of the oSZ[23] (**Supplementary Fig. 10c**). Thus, we hypothesize that vRG express module 3 and module 5 genes, whereas oRG express module 2 and module 4 genes[3,4].

To test the distinction between vRG and oRG cells, we identified genes enriched in modules 2 or 3 that were also differentially expressed between the ventricular zone and the oSZ regions in the 21 PCW BrainSpan Atlas data. We identified ten genes enriched in both the FRISCR-prepared vRG cells and the 21 PCW ventricular zone, and nine genes enriched both in FRISCR-prepared oRG cells and the 21PCW oSZ (**Fig. 5d** and **Supplementary Fig. 10b**). The canonical progenitor markers *SOX2*, *PAX6*, *EOMES* and *HES1* were expressed at similar levels in both germinal zone

© 2015 Nature America, Inc. All rights reserved.

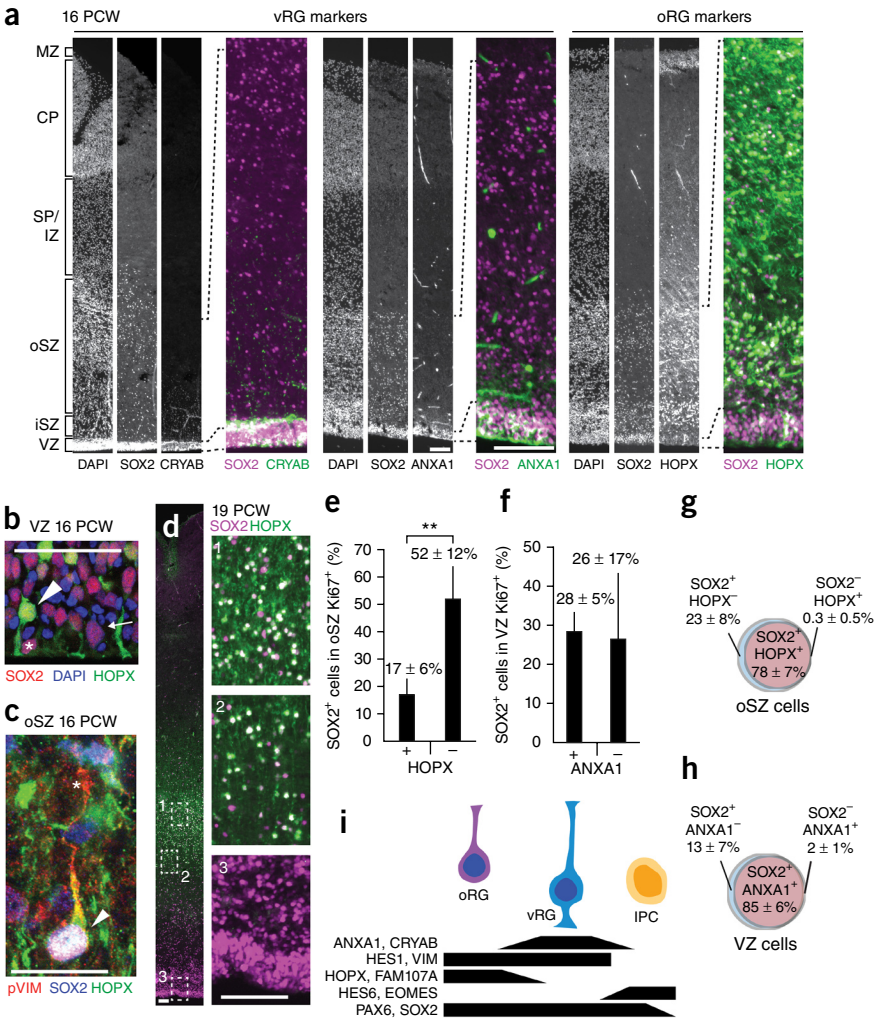


Figure 6 | Confirmation of vRG and oRG markers. (**a**) Grayscale (left) or merged (right) micrographs of 16 PCW cortex germinal zone showing SOX2 (magenta) co-stained with CRYAB, ANXA1 or HOPX (green). Scale bars, 100 µm. Marginal zone (MZ), cortical plate (CP), subplate/intermediate zone (SP/IZ), outer subventricular zone (oSZ), inner subventricular zone (iSZ) and ventricular zone (VZ). (**b**) HOPX expression in the VZ. Scale bar, 50 µm. Co-expression with SOX2 (arrowhead), SOX2-only cells (arrow), and HOPX$^+$ apical endfoot (asterisk) are indicated. (**c**) A mitotic cell co-stained with SOX2, HOPX and phospho-VIM shows basal cell process (asterisk) and cell body (arrowhead). Scale bar, 20 µm. (**d**) Full cortical thickness (left) and insets (right) from germinal zone stained for SOX2 (magenta) and HOPX (green). Scale bar, 100 µm. (**e**) Quantitation of SOX2$^+$HOPX$^+$ or SOX2$^+$HOPX$^-$ cells in oSZ that are Ki67$^+$. (**f**) Quantification of SOX2$^+$ANXA1$^+$ or SOX2$^+$ANXA1$^-$ cells in VZ that are Ki67$^+$. Mean ± s.d. for **e**, $n = 5$; or for **f**, $n = 6$ human tissues at 14–19 PCW; $^{**}P < 0.01$, Wilcoxon rank-sum test. (**g**,**h**) Percentage overlap of SOX2 and HOPX in oSZ cells (**g**), and of SOX2 and ANXA1 in VZ cells (**h**). Data in **g** and **h** are from $n = 9$ and $n = 10$ human tissues between 14 and 19 PCW, respectively. 200 to 500 cells were evaluated per brain. (**i**) Summary of oRG and vRG cell-type markers identified in this study and overlap with existing RG and IPC markers.

regions (**Fig. 5d**). We confirmed several genes that marked vRG and oRG cells with antibody staining of human cortical sections (**Fig. 6a**). CRYAB and ANXA1 showed signal only in the ventricular zone (except for vascular staining), whereas oRG markers HOPX and F3 (encoded by module 2 genes) showed high expression in the oSZ and iSZ, but faint signal in the ventricular zone of a 15 and 16 PCW cortex (**Fig. 6a** and **Supplementary Figs. 9b** and **11**). We used SOX2 to mark RG, as nearly all SOX2$^+$ cells were PAX6$^+$ in the germinal zones (**Supplementary Fig. 12a**). HOPX$^+$SOX2$^+$ cells in the ventricular zone occasionally exhibited apical process anchored at the ventricular surface at 16 PCW (**Fig. 6b**), and in the oSZ often displayed basally directed processes as assessed by phospho-Vimentin or HOPX staining (**Fig. 6c**,**d** and **Supplementary Fig. 11b**). At later developmental stages such as 19 and 21 PCW, we detected *HOPX* exclusively in the oSZ and almost never in the ventricular zone by gene expression (**Fig. 5d** and **Supplementary Fig. 10b**), and HOPX by antibody staining (**Fig. 6d** and **Supplementary Fig. 11a**). HOPX$^+$ oRG and ANXA1$^+$ vRG were proliferative because 17% of HOPX$^+$SOX2$^+$ cells and 28% of ANXA1$^+$SOX2$^+$ cells were also positive for Ki67 (**Fig. 6e**,**f** and **Supplementary Fig. 11c**). The majority of oRG progenitors were marked by HOPX: more than 78% of SOX2$^+$ cells in the oSZ co-stained with HOPX, and less than 1% of HOPX$^+$ cells in the oSZ were SOX2$^-$ (**Fig. 6g**). In contrast, 85% of SOX2$^+$ cells in the ventricular zone were ANXA1$^+$, and only 2% of ANXA1$^+$ cells in the ventricular zone lacked SOX2 expression (**Fig. 6h**). Module 2 genes such as *HOPX* and module 3 genes such as *ANXA1* are thus robust diagnostic markers for human mid-gestation oRG and vRG. Multiplexing these markers reveals distinct new progenitor populations in the developing human cortex for further analysis (**Fig. 6i** and **Supplementary Fig. 12b–e**). Mouse HOPX expression is restricted to the medial cortex[24], while we detected human HOPX protein and gene expression throughout the human developing cortical germinal zones (**Supplementary Fig. 13**). This suggests HOPX expression has changed through evolution. Study of these progenitors will help reveal molecular changes that underpin cortical evolution.

DISCUSSION

FRISCR is the first method, to our knowledge, that enables targeted mRNA purification and single-cell transcriptomic profiling of fixed cells without compromising data quality compared with live cells. FRISCR has three important advantages over current single-cell profiling methods. First, it allows enrichment of rare cell populations without live-cell markers. Without enrichment, we would have needed to profile 20,000 cortical cells to obtain data from 200 SP cells. Enrichment is particularly valuable for primary human tissues that are genetically inaccessible, or for which live-cell markers are unknown or poorly validated. Second, intracellular markers can be multiplexed with genetic reporters or other marking strategies to prospectively isolate and profile precisely defined cell types. Third, the FRISCR procedure allows

© 2015 Nature America, Inc. All rights reserved.

fixed cells to be stored indefinitely at −80 °C (and easily transported), unlike current protocols that require sorting of cells on the same day as tissue harvesting. Future optimizations will streamline sample processing and could allow both genome and transcriptome sequencing from the same fixed and sorted cell[25] or single-cell profiling of cells marked by FISH probes[15,26].

We demonstrated that FRISCR can be applied to investigate the composition of human neocortical progenitors. Several studies have analyzed differential gene expression or co-expression in microdissected germinal zones including the oSZ[10,11,27]. Although those studies had revealed new RG gene networks, they did not identify distinct oRG markers, likely because the cell type makes up a minority of the sample population. Two recent studies have set out to profile human RG progenitors in single cells. To enrich for cortical progenitors, the authors of the first study had sorted live cells based on CD133 expression and DiI labeling to mark apical (including vRG) and basal progenitors (excluding vRG but including oRG)[12]. ARHGAP11B had been identified and shown to regulate human progenitor dynamics. However, only the apical progenitors expressed *PAX6* and both populations expressed high levels of *EOMES*. Thus, the apical compartment defined by the study likely contained IPCs and RG, whereas the basal compartment likely lacks RG owing to the absence of *PAX6*. We did not detect *ARHGAP11B* expression in RG or IPC cells, indicating that this gene likely has a role in different cell types or stages. The second study used differential CD133, GLAST and CD15 staining to enrich and profiled vRG or oRG[13]. NEUROG2 had been identified as a regulator of oRG cells[13]. High *EOMES* expression had been detected in both vRG and oRG, similar to the first study. We detected only low expression of *NEUROG2* in our RG cells and much higher expression in our IPCs, suggesting that the population of oRG in ref. 13 best corresponds to our IPCs, which also express *HES6* and *EOMES*. Those studies[12,13] did not detect our validated and new vRG- or oRG-specific markers. Together these observations strongly suggest our study is to our knowledge the first to isolate and accurately profile these cells at the transcript and single-cell level. While our paper was in press, a study that profiled single cells from microdissected human germinal zone reported vRG- and oRG-specific gene markers that overlap substantially with our study (including *CRYAB* and *HOPX*, respectively)[28]. This provides independent confirmation that vRG and oRG cells can be distinguished using our new molecular markers.

Our results lay a foundation for the study of human RG diversity, maturation and fate. Using FRISCR, we discovered markers for human vRG and oRG, some of which appear likely to regulate RG self-renewal and differentiation. HOPX, a transcriptional regulator and oRG marker, shows divergent expression between mouse and human developing cortex[24], and may illuminate differences between the two species. HOPX is expressed in adult neurogenic progenitors in the hippocampus as well as self-renewing multipotent progenitors in the intestine, skin and certain cancers[29–33] where it functions by suppressing immediate early response genes[33,34]. Thus, the expansion of the human oSZ compartment in evolution could involve the reuse of existing mechanisms that control stem cell homeostasis in different tissues. Indeed, HOPX was recently shown to mark restricted cardiomyocyte progenitors, and to modulate their fate through the integration of niche BMP and WNT signals[35], pathways that are also thought to underpin corticogenesis[27,36]. Future work can focus on characterizing the origins of mid-gestation vRG and oRG cells, and determining whether they maintain neurogenic potential or have become glial-restricted[8,37,38]. Unraveling the fates of these progenitors and their regulatory mechanisms is central to understanding how the human brain is formed and has evolved.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. Gene Expression Omnibus: GSE71858 (H1 hESC data). dbGaP: phs001016.v1.p1 (human single-cell data). Data from donor H14.4010 is available upon request.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank the Allen Institute founders, P.G. Allen and J. Allen, for their vision, encouragement and support, A. Bernard and E. Lein for assistance with specimen procurement, J. Phillips for technical core support and leadership, A.-R. Krostag for assistance in SmartSeq2 methods, Y. Wang and WiCell, the sources of the H1 hESCs, C. Thompson for program management, B. Tasic, T. Nguyen and V. Menon for technical and scientific assistance, J. Miller for assistance in analysis of BrainSpan Atlas of the Developing Human Brain, and T. Bakken for assistance with the National Institutes of Health (NIH) Blueprint Non-Human Primate Atlas. Human primary samples were received from the "Laboratory of Developmental Biology", supported by NIH award 5R24HD000836 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development. S.J., A.M.D. and S.R. were supported in part by the NIH Directors Pioneer award 5DP1MH099906-03 and National Science Foundation grant PHY-0952766.

AUTHOR CONTRIBUTIONS

E.R.T. established the FRISCR based on prior methods from A.M.D. and S.J., and conducted all FRISCR RNA-seq experiments. J.K.M., R.D.H. and S.I.S. harvested and analyzed primary tissues. A.M.N. cultured and provided H1 hESCs, and N.V.S. did all sorting. Z.Y. conducted all RNA-seq primary analysis and computational analysis of gene expression. B.P.L., E.R.T., J.K.M. and S.R. conceived all experiments, and B.P.L., J.K.M. and S.R. wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Betizeau, M. *et al.* Precursor diversity and complexity of lineage relationships in the outer subventricular zone of the primate. *Neuron* **80**, 442–457 (2013).
2. Borrell, V. & Reillo, I. Emerging roles of neural stem cells in cerebral cortex development and evolution. *Dev. Neurobiol.* **72**, 955–971 (2012).
3. Dehay, C., Kennedy, H. & Kosik, K.S. The outer subventricular zone and primate-specific cortical complexification. *Neuron* **85**, 683–694 (2015).
4. Florio, M. & Huttner, W.B. Neural progenitors, neurogenesis and the evolution of the neocortex. *Development* **141**, 2182–2194 (2014).
5. Lui, J.H., Hansen, D.V. & Kriegstein, A.R. Development and evolution of the human neocortex. *Cell* **146**, 18–36 (2011).
6. Rakic, P. Evolution of the neocortex: a perspective from developmental biology. *Nat. Rev. Neurosci.* **10**, 724–735 (2009).
7. Fietz, S.A. *et al.* OSVZ progenitors of human and ferret neocortex are epithelial-like and expand by integrin signaling. *Nat. Neurosci.* **13**, 690–699 (2010).
8. Hansen, D.V., Lui, J.H., Parker, P.R. & Kriegstein, A.R. Neurogenic radial glia in the outer subventricular zone of human neocortex. *Nature* **464**, 554–561 (2010).
9. Geschwind, D.H. & Rakic, P. Cortical evolution: judge the brain by its cover. *Neuron* **80**, 633–647 (2013).
10. Fietz, S.A. *et al.* Transcriptomes of germinal zones of human and mouse fetal neocortex suggest a role of extracellular matrix in progenitor self-renewal. *Proc. Natl. Acad. Sci. USA* **109**, 11836–11841 (2012).

© 2015 Nature America, Inc. All rights reserved.

11. Miller, J.A. *et al.* Transcriptional landscape of the prenatal human brain. *Nature* **508**, 199–206 (2014).
12. Florio, M. *et al.* Human-specific gene ARHGAP11B promotes basal progenitor amplification and neocortex expansion. *Science* **347**, 1465–1470 (2015).
13. Johnson, M.B. *et al.* Single-cell analysis reveals transcriptional heterogeneity of neural progenitors in human cortex. *Nat. Neurosci.* **18**, 637–646 (2015).
14. Hrvatin, S., Deng, F., O'Donnell, C.W., Gifford, D.K. & Melton, D.A. MARIS: method for analyzing RNA following intracellular sorting. *PLoS ONE* **9**, e89459 (2014).
15. Klemm, S. *et al.* Transcriptional profiling of cells sorted by RNA abundance. *Nat. Methods* **11**, 549–551 (2014).
16. Molyneaux, B.J. *et al.* DeCoN: genome-wide analysis of in vivo transcriptional dynamics during pyramidal neuron fate selection in neocortex. *Neuron* **85**, 275–288 (2015).
17. Pan, Y., Ouyang, Z., Wong, W.H. & Baker, J.C. A new FACS approach isolates hESC derived endoderm using transcription factors. *PLoS ONE* **6**, e17536 (2011).
18. Pechhold, S. *et al.* Transcriptional analysis of intracytoplasmically stained, FACS-purified cells by high-throughput, quantitative nuclease protection. *Nat. Biotechnol.* **27**, 1038–1042 (2009).
19. Yang, L., Duff, M.O., Graveley, B.R., Carmichael, G.G. & Chen, L.L. Genomewide characterization of non-polyadenylated RNAs. *Genome Biol.* **12**, R16 (2011).
20. Picelli, S. *et al.* Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat. Methods* **10**, 1096–1098 (2013).
21. Pollen, A.A. *et al.* Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in developing cerebral cortex. *Nat. Biotechnol.* **32**, 1053–1058 (2014).
22. Zhang, B. & Horvath, S. A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.* **4**, 17 (2005).
23. Smart, I.H., Dehay, C., Giroud, P., Berland, M. & Kennedy, H. Unique morphological features of the proliferative zones and postmitotic compartments of the neural epithelium giving rise to striate and extrastriate cortex in the monkey. *Cereb. Cortex* **12**, 37–53 (2002).
24. Mühlfriedel, S., Kirsch, F., Gruss, P., Stoykova, A. & Chowdhury, K. A roof plate-dependent enhancer controls the expression of Homeodomain only protein in the developing cerebral cortex. *Dev. Biol.* **283**, 522–534 (2005).
25. Macaulay, I.C. *et al.* G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. *Nat. Methods* **12**, 519–522 (2015).
26. Bushkin, Y. *et al.* Profiling T cell activation using single-molecule fluorescence in situ hybridization and flow cytometry. *J. Immunol.* **194**, 836–841 (2015).
27. Lui, J.H. *et al.* Radial glia require PDGFD-PDGFRbeta signalling in human but not mouse neocortex. *Nature* **515**, 264–268 (2014).
28. Pollen, A.A. *et al.* Molecular identity of human outer radial glia during cortical development. *Cell* **163**, 55–67 (2015).
29. De Toni, A. *et al.* Regulation of survival in adult hippocampal and glioblastoma stem cell lineages by the homeodomain-only protein HOP. *Neural Dev.* **3**, 13 (2008).
30. Shin, J. *et al.* Single-cell RNA-seq with waterfall reveals molecular cascades underlying adult neurogenesis. *Cell Stem Cell* **17**, 360–372 (2015).
31. Takeda, N. *et al.* Hopx expression defines a subset of multipotent hair follicle stem cells and a progenitor population primed to give rise to K6+ niche cells. *Development* **140**, 1655–1664 (2013).
32. Takeda, N. *et al.* Interconversion between intestinal stem cell populations in distinct niches. *Science* **334**, 1420–1424 (2011).
33. Yamashita, K., Katoh, H. & Watanabe, M. The homeobox only protein homeobox (HOPX) and colorectal cancer. *Int. J. Mol. Sci.* **14**, 23231–23243 (2013).
34. Katoh, H. *et al.* Epigenetic silencing of HOPX promotes cancer progression in colorectal cancer. *Neoplasia* **14**, 559–571 (2012).
35. Jain, R. *et al.* Heart development. Integration of Bmp and Wnt signaling by Hopx specifies commitment of cardiomyoblasts. *Science* **348**, aaa6071 (2015).
36. Chenn, A. & Walsh, C.A. Regulation of cerebral cortical size by control of cell cycle exit in neural precursors. *Science* **297**, 365–369 (2002).
37. Gertz, C.C., Lui, J.H., LaMonica, B.E., Wang, X. & Kriegstein, A.R. Diverse behaviors of outer radial glia in developing ferret and human cortex. *J. Neurosci.* **34**, 2559–2570 (2014).
38. Noctor, S.C., Martinez-Cerdeno, V., Ivic, L. & Kriegstein, A.R. Cortical neurons arise in symmetric and asymmetric division zones and migrate through specific phases. *Nat. Neurosci.* **7**, 136–144 (2004).

© 2015 Nature America, Inc. All rights reserved.

ONLINE METHODS

Cell isolation from fetal cortex. We acquired tissue from the Birth Defects Research Laboratory (BDRL) at the University of Washington, who obtained appropriate written informed consent and provided available nonidentifying information for each sample. The Human Subjects Division at the University of Washington approved tissue acquisition, which was performed according to the requirements of the Uniform Anatomical Gift Act and National Organ Transplant Act for the acquisition of human tissue for biomedical research purposes. Fetal age was determined by foot length and date of last menses.

We divided cortical pieces into one half for fixation, sectioning and immunocytochemistry staining, and the other half for cell isolation. For dissociation, we minced the tissue into small pieces (~0.25–0.5 ml volume) with #5 forceps (Fine Science Tools) in Ca^{2+}-free and Mg^{2+}-free HBSS (14175-095, Life Technologies). We treated minced pieces with 2 ml trypsin solution for 20 min at 37 °C (Ca^{2+}-free and Mg^{2+}-free HBSS, 10 mM HEPES, 2 mM $MgCl_2$, 0.25 mg/ml bovine pancreatic trypsin (EMD Millipore), 10 µg/ml DNase I (Roche), pH 7.6, 100 nM TTX (Tocris, 1069), 20 µM DNQX (Tocris, 0189) and 50 µM DL-AP5 (Tocris, 3693)). We quenched digestion with 6 ml of ice-cold quenching buffer (440 ml Leibovitz L-15 medium, 50 ml water, 5 ml 1M HEPES pH 7.3–7.4, 5 ml 100× penicillin-streptomycin, 2 mg/ml bovine serum albumin (A7030, Sigma), 100 µg/ml trypsin inhibitor (T6522, Sigma), 10 µg/ml DNase I, 100 nM TTX, 20 µM DNQX, and 50 µM DL-AP5). We then pelleted the samples (220*g*, 4 min, 4 °C) and resuspended with 1 ml of quenching buffer and triturated on ice with a 1-ml pipette set to 1 ml, using 25 gentle cycles up and down without forming bubbles. We then diluted the cell suspension to 30 ml in Staining Medium (440 ml Leibovitz L-15 medium, 50 ml water, 5 ml 1 M HEPES pH 7.3–7.4, 5 ml 100× penicillin-streptomycin, 20 ml 77.7 mM EDTA pH 8.0 (prepared from Na_2H_2EDTA), 1 g bovine serum albumin, 100 nM TTX, 20 µM DNQX, and 50 µM DL-AP5), filtered through a 45 µm cell filter, pelleted (220*g*, 10 min, 4 °C), resuspended in 5 ml staining medium, and counted on a hemocytometer (typically ~30–50 million live cells isolated per cortical piece at ~50% viability).

Cell isolation from culture. We maintained H1 hESCs (WiCell) on Matrigel (Corning) in mTESR1 media (StemCell Technologies), authenticated them by karyotype analysis (Cell Line Genetics), and periodically tested them for sterility and mycoplasma contamination (IDEXX BioResearch). We dissociated adherent cultures with StemPro Accutase Cell Dissociation Reagent (Life Technologies), and then centrifuged the cells (220*g*, 3 min) and the dissociation solution was removed.

FRISCR. *Cell wash.* We washed then resuspended the cells in RNase-free staining buffer (SB) (1× PBS pH 7.4, 1% RNase-free BSA (Gemini Bioproducts), and 0.0025% RNasin Plus (Promega)). We placed cells on ice until fixation or sorting.

Fixation. We fixed the single cell suspension with 4% PFA (Electron Microscopy Sciences) in PBS on ice for 15 min, then pelleted them (335*g*, 3 min, 4 °C), washed them once with 1 ml SB, then resuspended the cells in SB at 10 million cells/ml, and froze them at −80 °C in aliquots.

Permeabilization and staining. We thawed and permeabilized the cells by resuspending and incubating for 10 min on ice in 1× PBS, 0.1% Triton X-100 (Sigma), 1% BSA, 0.0025% RNasin Plus. We incubated one million cells in SB (1× PBS, 1% BSA, 0.0025% RNasin Plus) with primary antibodies for 30 min at 4 °C. Antibodies used were: Alexa Fluor 488– or PE-conjugated anti-PAX6 (O18-1330; BD Biosciences), PE-conjugated anti-DCX (30/Doublecortin; BD Biosciences), PerCP-Cy5.5-conjugated anti-SOX2 (O30-678; BD Biosciences), eFluor660-conjugated anti-EOMES (WD1928; eBioscience), rabbit anti-CYR61 (D4H5D; Cell Signaling Technology) followed by PE-conjugated goat anti-rabbit IgG (Life Technologies). We washed cells in SB, resuspended them in SB containing 1 µg/ml DAPI (Life Technologies), and filtered before sorting.

Single-cell sorting. We carried out cell sorting on a BD FACS ARIA-II SORP (BD Biosciences) using a 130 µm nozzle. We sorted single cells into strip tubes containing 5 µl of PKD buffer (Qiagen) with 1:16 proteinase K solution (Qiagen) and ERCC spike-in synthetic RNAs (Life Technologies).

Cell lysis, reverse crosslinking and RNA purification. To lyse cells and purify RNA with dT25 beads, we thawed samples at 25 °C, room temperature (RT), mixed them, and then incubated at 56 °C for 1 h in a thermal cycler with the lid set at 66 °C. We vortexed cells for 10 s, spun them down, and placed them on ice. We prepared oligo dT25 magnetic beads (Life Technologies) with three washes of 1× hybridization buffer (2 × SSPE, 0.05% Tween 20, 0.0025% RNasin Plus) and then resuspended in half of the original volume of 2× hybridization buffer. 5 µl of washed dT25 beads (0.05 mg of beads) were used per reaction.

We added beads to reverse crosslinked samples and then heated to 56 °C for 1 min, incubated at RT 10 min to allow mRNA hybridization, and then placed mixtures on ice. We washed beads two times in 100 µl of ice-cold hybridization buffer, followed by a subsequent wash using ice-cold 1× PBS, 0.0025% RNasin Plus. We removed PBS and added 2.8 µl of RNase-free water, then resuspended the beads and incubated the mixture at 80 °C for 2 min to elute mRNA, then immediately pelleted on a room temperature magnet. We rapidly removed the supernatant containing mRNA, transferred it to a new tube and stored it at −80 °C.

RNA extraction from populations of cells. *Sorting.* We sorted cells as described above. We sorted populations up to 1,000 cells into 1.5 ml tubes with 100 µl of PKD solution. We sorted 10,000 cell samples into 1.5 ml tubes containing 500 µl of SB, pelleted them, and resuspended the cells in 100 µl of PKD solution.

Cell lysis, reverse cross-linking and RNA purification. We purified total RNA using either the standard live-cell methods from the RNeasy Micro Kit (RLT method) or a modified protocol for fixed cells using the miRNeasy FFPE Kit (Qiagen) (protease lysis, reverse cross-linking method (PLRC)). For the PLRC method, we thawed cells in 100 µl of PKD with Proteinase K solution at RT, mixed the samples, and incubated them at 56 °C for 1 h. Samples were then centrifuged at 20,000*g* for 20 min. We transferred the supernatant to a new tube and added 10 µl of DNase booster buffer (a component of the Qiagen RNeasy kit) and 10 µl of DNase I stock solution and mixed by inversion. Then we incubated samples 15 min at RT then added 320 µl of RBC buffer and 1,120 µl of ethanol and mixed. Finally, we applied the samples to MinElute columns, washed twice with 500 µl RPE Buffer, and eluted RNA into 15 µl of water. We evaluated

© 2015 Nature America, Inc. All rights reserved.

doi:10.1038/nmeth.3629

and quantified RNA on a Bioanalyzer 2100 (Agilent) using the RNA Pico Kit. To normalize input, we used 1 µl directly from 1,000 cells, 1 µl of tenfold diluted RNA from 10,000 cells, or 1 µl of tenfold concentrated RNA from 100 cells (Savant DNA 120 SpeedVac).

SmartSeq2. We prepared sequencing libraries as previously reported[20]. After reverse transcription and template switching, we amplified cDNA with KAPA HotStart HIFI 2× ReadyMix (Kapa Biosystems) for 19 or 22 cycles for RNA from single hESC or cortical progenitor cells, respectively. We purified PCR products using Ampure XP beads (Beckman Coulter). We quantified cDNA using a High Sensitivity DNA Chip (Agilent) on a Bioanalyzer 2100, or with the Quant-iT PicoGreen dsDNA Assay Kit (Life Technologies) on an Enspire plate reader (PerkinElmer). We used 1 ng of cDNA to generate RNA-seq libraries using the Nextera XT library prep system (Illumina). Single cells from primary tissue contained a lower amount of mRNA compared to H1 hESCs, requiring a reduction in ERCC spike-in RNAs by tenfold and addition of three extra PCR cycles. We carried out sequencing of human cortical progenitors on Illumina MiSeq using 31-base paired-end reads. We carried out sequencing of hESCs on the HiSeq using 50-base paired-end reads. We saw few global differences in read statistics between MiSeq runs and HiSeq runs when samples were assessed on both instruments.

RNA-seq data analysis. We aligned raw read data to GRCh37 (hg19) using the RefSeq annotation gff file downloaded on 23 April, 2013. We performed transcriptome alignment first using RSEM[39], then we aligned unmapped reads to hg19 using Bowtie[40], and then we aligned remaining unmapped reads to the ERCC sequences. Using a custom script we calculated the read mapping % of each cell to mRNA (RefSeq), mitochondrial RNA (mtRNA), noncoding RNA (mRNA NC), ribosomal RNA (rRNA), genome and ERCC RNA. We performed principal-component and clustering analysis using transcripts per million (TPM) values ($\log_2$-transformed). We used only high-variance genes with adjusted $P < 0.05$ and those that were present (mapped reads > 0) in more than ten cells for analysis. RSEM-generated gene count and TPM data from primary fetal human tissue is supplied as **Supplementary Data 1** and **2**, and raw data will be provided upon request.

Computational analysis. *Normalization.* All gene expression heatmaps show data normalized across all samples for each gene or eigengene. Single-cell gene expression data and eigengene values are TPM, whereas microarray data from BrainSpan Atlas of the Developing Human Brain[11] is shown as 'expression value'. Heatmap values show $\log_2$(TPM or expression value + 1) – average TPM or expression value.

WGCNA analysis. We included only high variance genes based on DESeq2 with an adjusted $P < 0.01$ and expressed in at least five cells. We performed WGCNA clustering using soft power of four, cut height of 0.995, and minimum module size of ten. To filter gene modules, we used genes of each module to cluster cells into two clusters, and then evaluated to determine which of the module genes are differentially expressed between the two clusters. We removed a gene module if one of the two clusters contained less than five cells, or if the total differential score ($-\log_{10}$ adjusted P value of differentially expressed genes) was less than 40. This filtering criterion eliminates gene modules driven by very small numbers of cells, and gene modules with little discriminating power. We also eliminated two gene modules corresponding to cell cycle states. We clustered cells using hclust based on filtered module genes and Ward's distance measure. After the initial clusters were determined, we detected differentially expressed genes between every pair of cell clusters using the Bioconductor limma package. The union of all such genes was defined as cluster-specific markers. We then reclustered the cells based on the marker genes using hclust to determine the final clusters.

3′ to 5′ bias analysis. To assess 5′ to 3′ read coverage bias, transcripts not alternatively spliced, and with TPM > 1 in at least 30 samples were selected. Each transcript was divided into 100 bins uniformly distributed across the transcript, and average read coverage within each bin was determined from only uniquely aligned reads (RSEM ZW score > 0.5). Data are shown as normalized average read coverage per bin for all genes or genes of a predicted length.

Mutational analysis. To detect mutations in each sample, we filtered alignment bam files to exclude reads with more than two errors or ambiguous mapping (RSEM ZW score < 0.95). Variance calling was performed by comparing reads to the reference genome using samtools mpileup followed by bcftools call[41]. We filtered detected mutations to exclude mutations within three base pairs of insertion or deletion mutations, with a quality score of < 10, depth of < 5 or an alignment mapping quality (MQ) score of < 50, using bcftools. We defined common variants as nucleotide changes where the variant is shared by more than 10 samples, and by more than two thirds of all samples that pass the quality filter threshold. The mutation rate for each sample is defined as non-common variant mutations divided by the number of genomic positions that pass the above-mentioned filter.

G+C and chromosomal bias analysis. To assess the G+C bias between TL- and FRISCR-prepared live and fixed single hESCs, we pooled genes into ten equally sized bins based on percentage G+C content. For each bin, we calculated the fraction of the genes detected in each sample out of all annotated genes. To test for any chromosomal bias, we binned genes by their chromosomal locations. We calculated the fraction of the genes detected in each sample out of all annotated genes for each chromosome.

Human Atlas comparisons. For the BrainSpan Atlas of the Developing Human Brain[11] laser capture microdissection (LCM) microarray data sets eigengene expression was determined using the same gene modules described above. In both data sets we compared side-by-side a subset of genes differentially expressed between ventricular zone (VZ), iSZ and oSZ and differentially expressed between cell clusters C and D (oRG cells) versus cell cluster E (vRG cells) (**Fig. 5d**). We determined the statistical significance of eigengene enrichment in human LCM data by unpaired t test comparing the VZ and the oSZ.

Tissue immunocytochemistry. We fixed cortical tissue sections by immersion in 4% paraformaldehyde in PBS for 24 h at 4 °C. We washed tissue in PBS, then transferred to and stored in 30% sucrose in PBS (Sigma) for 2 days (or until tissue sank). We embedded tissue in Tissue-Tek O.C.T (Sakura Finetek) and

© 2015 Nature America, Inc. All rights reserved.

stored at −80 °C. We cut 25 µm coronal cryosections and stored at −80 °C. For staining we thawed slides and dried for 15 min at RT, washed briefly with PBS to remove OCT, and blocked in 5% goat or donkey serum (Jackson Immunoresearch), 0.1% Triton X-100 in PBS for 1 h at RT. We diluted primary antibodies in blocking solution and applied to slides 4–12 h rocking at RT. We next washed slides in PBS and incubated in secondary antibodies diluted 1:1,000 plus 0.5 µg/ml DAPI (Life Technologies) in blocking solution for 1–3 h rocking at RT, then applied coverslips with Prolong Gold mounting media (Life Technologies). We used primary antibodies: goat anti-SOX2 (Santa Cruz Biotechnology; SC17320), mouse anti-SOX2 (BD Biosciences, 245610), rabbit anti-PAX6 (BioLegend; PRB-278P); mouse anti-EOMES (eBioscience; WD1928); rabbit anti-HOPX (Santa Cruz Biotechnology; FL-73), mouse anti-CRYAB (Abcam; 1B6.1-3G4); Rabbit anti-c-JUN (Cell Signaling Technology; 60A8), Goat anti-human Coagulation Factor III/Tissue Factor (F3) (R&D Systems; AF2339), mouse anti-Ki67 (BD Biosciences, B56), rabbit anti-Ki67 (Abcam, ab15580), mouse anti-phospho-vimentin (Enzo; 4A4), and mouse anti-Lipocortin-1 (ANXA1) (BioLegend; 74/3). Alexa 488-, 555-, and 647-conjugated secondary antibodies were from Life Technologies.

We acquired high-resolution confocal images on a Leica TCS SP8 confocal microscope. Representative maximum intensity projections were created from two to three adjacent optical sections from the middle of a *z*-stack. We acquired epifluorescence images of the full cortical thickness on a Nikon Eclipse Ti with a motorized stage using a 10X objective, and we stitched together images using MetaMorph software. Brightness and contrast adjustments were made using Adobe Photoshop CS6 and we assembled figures using Adobe Illustrator CS6.

Statistics. *Sample inclusion/exclusion.* For analysis of RNA from populations of hESCs (**Fig. 2** and **Supplementary Fig. 1a**), we sorted four cell samples per condition/number on three independent days (12 total samples per condition/number) for the 10,000- and 1,000-cell experiments, and we conducted two independent experiments for the 100-cell experiment (eight total samples per condition/number). Two outliers from both 1,000 hESC Live/PLRC and Fixed/PLRC samples were excluded from analysis because of process failure and no observed RNA. The Bioanalyzer was unable to call an RNA integrity number for one 10,000-cell Fixed/RLT sample, and one 100-cell Live/PLRC sample; and one 100-cell Live/PLRC, and Fixed/PLRC samples were lost due to mechanical failure.

For sequencing single hESCs (**Fig. 3**, and **Supplementary Figs. 2** and **3**), we sorted eight individual hESCs on two days of experiments for each condition (16 cells total per condition). Library generation failed for some cells, as judged by insufficient cDNA for tagmentation (0.1 ng/µl) yielding $n = 12–15$ cells per condition (**Supplementary Fig. 2b**). We processed cells with sufficient cDNA into libraries for sequencing by tagmentation, however, samples from only one experiment of Fixed/TL samples were processed for sequencing, and we used the entire sample for library generation due to low cDNA yield. We included all sequenced single hESCs for analysis after subsampling to five million reads per cell. Importantly, the failure rate of live cells verses fixed cells is comparable.

For sequencing single human cortex progenitors (**Figs. 4** and **5**, and **Supplementary Figs. 5–8**), we sorted 207 total SP cells from across four brains and 48 single SPE cells from across two brains for library preparation and all were sequenced. Of these, cells with more than 10,000 mapped reads, >20% reads mapping to mRNA, and detectable *GAPDH* expression were included in further analysis (157 SP and 29 SPE cells analyzed), all reads are included in analysis.

For analysis of human cortex by ICC, in **Figure 6e** we stained six brains and excluded one because of poor staining ($n = 5$ total analyzed), in **Figure 6f** we stained six brains and analyzed all, in **Figure 6g** we stained ten brains and excluded one due to poor staining ($n = 9$ total analyzed), in **Figure 6h** we stained nine brains and analyzed all, in **Supplementary Figure 12a** we stained four brains and excluded one due to poor staining ($n = 3$ total analyzed), and in **Supplementary Figure 12b**,**c** we stained four brains and analyzed all.

Power analysis and statistical tests. In **Figure 2b**,**c** prospective power for the comparisons was 1.0 given effect size = 5 × s.d. (about twofold change), $n = 28$ total samples, and $\alpha = 0.05$.

In **Figure 3d** we used one-way ANOVA followed by Tukey's post-hoc test for the indicated comparisons. Data within each group followed normal distributions as indicated by Shapiro-Wilk test. Prospective power for the comparisons were 0.99-1.0 given effect size = 1–9 × s.d. (about twofold changes), $n = 44$ total samples, and $\alpha = 0.05$.

In **Figure 6e**, we used Wilcoxon rank sum test because $SOX2^{+}HOPX^{-}$ data did not follow a normal distribution by the Shapiro-Wilk test. Prospective power for this comparison was 0.98 given effect size = 3 × s.d. (about twofold change), $n = 5$ brains per group, and $\alpha = 0.05$.

For **Figure 6f** prospective power for the comparisons was 1.0 given effect size = 5 × s.d. (about twofold change), $n = 5$ brains per group, and $\alpha = 0.05$. We conducted all power analyses using G*Power[42].

39. Li, B. & Dewey, C.N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
40. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
41. Li, H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987–2993 (2011).
42. Faul, F., Erdfelder, E., Lang, A.G. & Buchner, A.G. *Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* **39**, 175–191 (2007).

© 2015 Nature America, Inc. All rights reserved.

doi:10.1038/nmeth.3629