



Single-neuron RNA-Seq: technical feasibility and reproducibility

Shenfeng Qiu^{1,2}, Shujun Luo³, Oleg Evgrafov¹, Robin Li³, Gary P. Schroth³, Pat Levitt^{1,2}, James A. Knowles^{1,4*}, and Kai Wang^{1,4*}

¹ Zilkha Neurogenetic Institute, University of Southern California, Los Angeles, CA, USA

² Department of Cell and Neurobiology, University of Southern California, Los Angeles, CA, USA

³ Illumina, Inc, Hayward, CA, USA

⁴ Department of Psychiatry, University of Southern California, Los Angeles, CA, USA

Edited by:

Paul Pavlidis, University of British Columbia, Canada

Reviewed by:

Stephen D. Ginsberg, Nathan Kline Institute, USA

Alexandre Kuhn, Agency for Science, Technology and Research (A*STAR), Singapore

*Correspondence:

James A. Knowles and Kai Wang, Department of Psychiatry, University of Southern California, Health Sciences Campus, ZNI 401, M/C 2821, Los Angeles, CA 90089-2821, USA. e-mail: knowles@med.usc.edu; kaiwang@usc.edu

Understanding brain function involves improved knowledge about how the genome specifies such a large diversity of neuronal types. Transcriptome analysis of single neurons has been previously described using gene expression microarrays. Using high-throughput transcriptome sequencing (RNA-Seq), we have developed a method to perform single-neuron RNA-Seq. Following electrophysiology recording from an individual neuron, total RNA was extracted by aspirating the cellular contents into a fine glass electrode tip. The mRNAs were reverse transcribed and amplified to construct a single-neuron cDNA library, and subsequently subjected to high-throughput sequencing. This approach was applied to both individual neurons cultured from embryonic mouse hippocampus, as well as neocortical neurons from live brain slices. We found that the average pairwise Spearman's rank correlation coefficient of gene expression level expressed as RPKM (reads per kilobase of transcript per million mapped reads) was 0.51 between five cultured neuronal cells, whereas the same measure between three cortical layer 5 neurons *in situ* was 0.25. The data suggest that there may be greater heterogeneity of the cortical neurons, as compared to neurons *in vitro*. The results demonstrate the technical feasibility and reproducibility of RNA-Seq in capturing a part of the transcriptome landscape of single neurons, and confirmed that morphologically identical neurons, even from the same region, have distinct gene expression patterns.

Keywords: cell culture, electrophysiology, gene expression, neuron, RNA-Seq, transcriptome

INTRODUCTION

Prior to the advent of high-throughput methods to measure the entire transcriptome, such as microarrays or next-generation sequencing, studies of gene expression and function in the brain were restricted to a relatively small number of genes (Luo and Geschwind, 2001; Zhang et al., 2002). Recently, whole transcriptome sequencing (RNA-Seq) has enabled the measurement of abundance of tens of thousands of RNA species in a given biological sample (Wang et al., 2009; Ozsolak and Milos, 2011). This new generation of high-throughput sequencing technology has delivered on its promise of sequencing DNA, cDNA, and RNA at unprecedented speed and accuracy, thereby providing an increasingly wide-ranging array of data sets that provide insight into biological and disease diversity (Schuster, 2008). Single-cell analysis represents one of the novel areas of application for high-throughput sequencing, which is particularly important for the study of tissues that have a high degree of intrinsic variation, such as the brain. To date, this technology has achieved success in tumor profiling to study somatic DNA mutations in clonal sub-populations: for example, Navin et al. (2011) applied single-nucleus DNA sequencing to investigate tumor population structure and evolution in human breast cancer cases. Single-cell RNA sequencing also

can be used to study differences of individual cells with identical genomes, for example, in a pool of neurons from a single brain.

There is a strong rationale to perform single-neuron RNA expression analysis. The mammalian brain consists of billions of neurons that each typically exhibits over 100,000 macroconnections (Bota et al., 2003; Luo et al., 2008). Individual neurons can be characterized into distinct cell types based on morphology, electrophysiological characteristics, connections, and expressed molecular diversity. Such cell-specific information is diluted when pooling groups of neurons. Gene expression differences occurring in rare cell types may go undetected, because they contribute to only a small fraction of total tissue RNA. Moreover, gene expression may be regulated in opposing directions in different cell types, thereby appearing static in composite data. Cell type-specific transcriptomics will provide a more panoramic view of gene expression, and ultimately networks, rather than from a viewpoint dominated by the effects of single genes. Moreover, correlating gene expression data of individual neurons with single-cell phenotypic data has the potential to refine cell type definitions.

Transcriptome analyses of single neurons have been reported using mostly microarrays. Earlier pioneering studies of single-cell

microarray analysis used a universal PCR amplification procedure, or two rounds of T7 RNA polymerase amplification. For example, Esumi et al. (2008) developed method for single-cell microarray analysis and applied them to gene-expression profiling of GABAergic neuron progenitors. Iscove et al. (2002) demonstrated that the representation is faithfully preserved in global cDNA amplified exponentially from sub-picogram quantities of mRNA. Kamme et al. (2003) examined expression profiling of CA1 neurons in the rat hippocampus using a combination of laser capture, T7 RNA amplification, and cDNA microarray analysis. Single GABAergic neuron progenitors from mouse neocortex were isolated by dissociation and aspiration of green fluorescent protein (GFP)-positive cells, then processed by Super SMART PCR and T7 RNA polymerase amplification. Similarly, Gustincich et al. (2004) combined SMART PCR and T7 RNA polymerase amplification to interrogate the transcriptome in single dopaminergic neurons of the retina in mice. Tietjen et al. (2003) also used similar techniques in single cells from dissociated tissues or collected from intact slices using laser capture. These studies demonstrated that microarray-based analysis of gene expression can be performed on single neuronal cells isolated from defined areas of the brain.

Compared with microarray methods, sequence-based transcriptome profiling has major advantages, such as extended linear range of detection, accuracy, binary low noise reading, and independence from a reference genome. Next-generation sequencing techniques have been applied recently to single-cell transcriptome analysis. For example, Tang et al. (2009, 2010b) used single-cell transcriptome analysis based on RNA-Seq and applied the methods to analyze the developmental program of embryonic stem cells (Tang et al., 2010a). Recently, Eberwine and Bartfai (2011) have examined the single-cell transcriptome in hypothalamic warm sensitive neurons that control core body temperature and fever response. These authors assayed cDNA libraries from single neurons using Affymetrix gene expression arrays, and confirmed the frequency of specific cDNAs by Illumina sequencing. The importance of single-cell transcriptome analysis has been increasingly recognized recently, especially for tracing cell lineages and in diagnostic applications (Stahlberg et al., 2011; Tang et al., 2011). The feasibility and reproducibility of RNA-Seq in single neurons, however, has not been systematically studied, though a protocol to assay single neurons by qPCR was recently published (Citri et al., 2012). Furthermore, experimental variability related to extremely low levels of RNA, and bioinformatics analysis of the RNA-Seq data, can be technically demanding. Our study differs from previous studies in that we developed protocols specifically for assaying neurons, and our method can be directly coupled with electrophysiology studies on individual neurons, thus enabling the correlation of single-cell cellular phenotypes with single-cell expression phenotypes. Below we describe the experimental strategy and application that was designed to test the practicality of single-neuron RNA-Seq using equipment readily available in a typical neurophysiology laboratory. The study demonstrates the ability to generate RNA-Seq data with reasonable reproducibility from individual, electrophysiologically characterized neurons.

MATERIALS AND METHODS

COLLECTION OF SINGLE NEURON CONTENTS

To assess the technical feasibility of RNA-Seq for single-neuron transcriptome profiling, we started with cultured mouse embryonic hippocampal neurons. All procedures used C57Bl6 mice, and were approved by the Institutional Animal Care and Use Committee of the University of Southern California. Animal care and handling conformed to National Institutes of Health guidelines. Hippocampal neuronal cultures were prepared from embryonic day (e) 16–17 mice according to standard protocols (Qiu and Weeber, 2007). Briefly, hippocampi from fetal brains were collected and digested in Hank's balance salt solution containing 0.5 mg/ml of papain at 37°C for 20 min. The hippocampi were then separated into a single-cell solution using aspiration with a glass micropipette. The cells were seeded on poly-D-lysine-coated 12-mm glass coverslips at a density of 10,000 cells/cm² and grown in Neurobasal medium supplemented with 2% B27. Under these growth conditions, most neurons exhibit a glutamatergic phenotype.

At 12 days *in vitro* (DIV), a glass coverslip with cultured neuron was transferred into an electrophysiology recording chamber that was perfused by circulating artificial cerebrospinal fluid (ACSF, containing in mM: 126 NaCl, 2.5 KCl, 26 NaHCO₃, 2 CaCl₂, 2 MgCl₂, 1.25 NaH₂PO₄, and 10 D-glucose). It is critical that ACSF is freshly made with milliQ water and the perfusion system is flushed with ACSF at least twice to maintain a “clean” surrounding for neurons. Neurons were visualized on an Olympus BX51WI microscope equipped with a ×60 water immersion lens and infrared differential interference contrast (DIC) optics, and a focusing step motor that is driven by a Sutter MP285 motion controller. A patch electrode pulled from borosilicate glass was filled with the recording solution: (in mM: 126 K-gluconate, 4 KCl, 10 HEPES, 4 ATP, 0.3 GTP, 10 mM phosphocreatine, and 0.4 Unit/μl RNaseOUT). This patch electrode had an electrical resistance of 2.5–3 MΩ in the open ACSF bath. The electrode was mounted on a micromanipulator (Sutter MP 285) and was placed over the target neuron under visual guidance. Only neurons with a pyramidal-shaped soma were selected for analysis to avoid potential interneurons. The electrode was sealed on the neuron by applying negative pressure. Whole cell configuration was obtained by breaking into the neuron following a gigaOhm seal. Neurons were then voltage clamped at their resting membrane potential with a Multiclamp 700B amplifier controlled by pClamp 10.2 software. Spontaneous neuronal activity was acquired for 30 s followed by a series of current step injections to verify that the neurons were healthy. In cases where there is a large (>100 pA) holding current and truncated action potential firing in response to current injection (neurons being “leaky”), neuronal content was not collected for analysis. Following the brief recording, stronger negative pressure was applied to draw all the soma contents into the electrode. The electrode was rapidly retracted from the bath, and the contents expelled into a thin-wall PCR tube (mounted on a manual manipulator) with 3.5 μl lysis buffer. This lysis buffer was made from nuclease free water with 2% Triton X-100, and 5% RNaseOUT (Invitrogen). The collected single cell contents were then immediately frozen on dry ice and stored at –80°C until further processing.

To obtain the cell contents of live single neurons *in situ* from brain tissue, sagittal sections were prepared from postnatal day (P) 21–28 neocortex, similar to that described previously (Qiu et al., 2010, 2011). Briefly, mice were deeply anesthetized by isoflurane and quickly decapitated. Brains were dissected and sectioned on a vibratome in ice-cold, carbogenated choline solution containing (in mM: 110 choline chloride, 25 NaHCO₃, 2.5 KCl, 1.25 NaH₂PO₄, 0.5 CaCl₂, 7 MgSO₄, 25 D-glucose, 11.6 sodium ascorbate, and 3.1 sodium pyruvate). Sagittal slices (350 μm thick) of the right hemisphere containing prefrontal cortex were cut. Slices were then incubated in ACSF for 30 min at 35°C and maintained at 22°C. A single slice was then transferred to the recording chamber with clean ACSF perfusate (see above). Under visual guidance, layer 5 pyramidal neurons in the anterior frontal cortex were targeted by a patch electrode. The patch electrode (3–3.5 MΩ) was filled with the same recording solution as described above. Prior to entering the ACSF, positive pressure in the electrode is applied so that the internal solution is flowing out during the whole process to avoid contamination of the intracellular solution by brain tissues. After the electrode approached the neuron under visual guidance, negative pressure is applied by mouth suction to form a gigaOhm seal. The cell membrane was ruptured by stronger, pulsate suction so that a continuity is formed between the electrode solution and neuronal content. The neurons are then tested for their membrane properties (membrane resistance, capacitance, time constant, resting membrane potential). Current steps were then injected into the neurons to test other electrophysiological properties (action potential threshold, spike frequency adaptation). To aspirate the neuronal contents, strong negative pressure was applied with a syringe, until the soma was completely extracted into the electrode. The complete aspiration of soma content into the patch electrode is visualized under DIC optics and by focusing through various Z plane levels. The neuronal content was collected into a thin-walled PCR tube containing 3.5 μl lysis buffer, flash frozen on dry ice, and stored at –80°C.

RNA-Seq LIBRARY PREPARATION AND DATA GENERATION

RNA-Seq library was generated following manufacturers recommended protocols (SMARTer™ Ultra Low input RNA Kit, Clontech Cat. 634935 and Illumina Paired-end DNA sample preparation kit Cat. PE-102-1001). The SMARTer kit can assay 100 pg of input RNA (well suited for single-cell analysis), and allows the efficient incorporation of known sequences at both ends of cDNA during first-strand synthesis in one step, without the adaptor ligation step. Briefly, the PCR tubes containing the cytoplasmic content of a single cell in 3.5 μl of reaction buffer were thawed on ice, and all the following steps were carried out in a laminar flow hood to reduce the possibility of contamination from environmental DNA sources. First-strand cDNA synthesis was initiated by adding 1 μl of 3' SMART CDS Primer II A (12 μM) at 72°C for 3 min, and then 5.5 μl Mastermix containing SMARTscribe™ Reverse Transcriptase (100 U) was added and the reaction was incubated at 42°C for 90 min, followed by inactivation at 70°C for 10 min. First-strand cDNA was isolated using SPRI Ampure Beads, and then double-strand cDNA were generated by long distance PCR for 18 cycles. Double-stranded DNA (dsDNA)

was purified using SPRI AMPure beads, and the quality was assessed using an Agilent 2100 BioAnalyzer. The double-stranded cDNA fragments were sheared by Covaris, then end-repaired with a combination of T4 DNA polymerase, Klenow fragment polymerase, and poly nucleotide kinase to ensure blunted ends, and then adenylated with a single A-base at the 3'-end of the fragment by the Klenow 3'-5' exo-enzyme. The tail was ligated with the Illumina adapters by T4 DNA ligase. The adapter-ligated sample was size selected by electrophoresis, and amplified by PCR using primers that only anneal to adapter-ligated fragments. The adapter sequences were annealed to the primers on the Illumina flow cell during bridge PCR in Illumina Genome Analyzer IIx sequencer, which generated the clusters necessary to view fluorescence during the sequencing-by-synthesis process.

RNA-Seq DATA ANALYSIS

The FASTQ file and associated base quality scores were generated by factory software associated with the Genome Analyzer IIx sequencer. We used TopHat (Trapnell et al., 2009) version 1.2.0 for aligning the Illumina reads against the reference mouse genome (NCBI build M37) with default parameters. We used Cufflinks (Trapnell et al., 2010) version 1.0.3 to summarize the gene expression values as RPKM (reads per kilobase of transcript per million mapped reads) measures using the –compatible-hits-norm argument, so that only reference transcripts are counted toward the number of mapped hits used in the RPKM denominator. We also used the “–G” argument, which calculates the gene expression levels for all known/annotated transcripts. For the annotated transcript library, we used the latest ENSEMBL (Birney et al., 2004) GTF file, downloaded from ftp://ftp.ensembl.org/pub/release-62/gtf/mus_musculus/ after filtering to remove all mitochondrial and ribosomal transcripts after alignment.

RESULTS AND DISCUSSION

TECHNICAL OVERVIEW OF SINGLE-NEURON RNA-Seq

The analysis of each single-cell transcriptome consists of several independent steps. First, the cell type of interest is identified; this can often be done in neurons through genetic labeling or by characterizing the electrophysiology properties. In many cases, cell-specific promoters can be used to drive the expression of a marker gene such as GFP. Second, RNA from the targeted neuron must be extricated from the surrounding cells and tissues. For instance, this can be done by aspirating the cellular contents into a patch clamp recording electrode under visual guidance. Alternatively, GFP-positive cells can be selected through fluorescence activated cell sorting, or by laser capture microdissection in fixed brain tissues (Lin et al., 2007). Third, because the resulting RNA is typically very low in abundance, it must be amplified and converted to cDNA before being subjected to sequencing or microarray analysis.

To demonstrate the feasibility of single-neuron RNA-Seq analysis, we assayed three layer 5 pyramidal neurons from the same brain slice that exhibited typical electrophysiological properties, as well as five randomly selected pyramidal cells from hippocampal neuronal cell culture at 12 DIV (**Table 1**). To harvest cellular contents from live neurons in culture or in brain slices, we used a “direct aspiration” approach that does not require specific

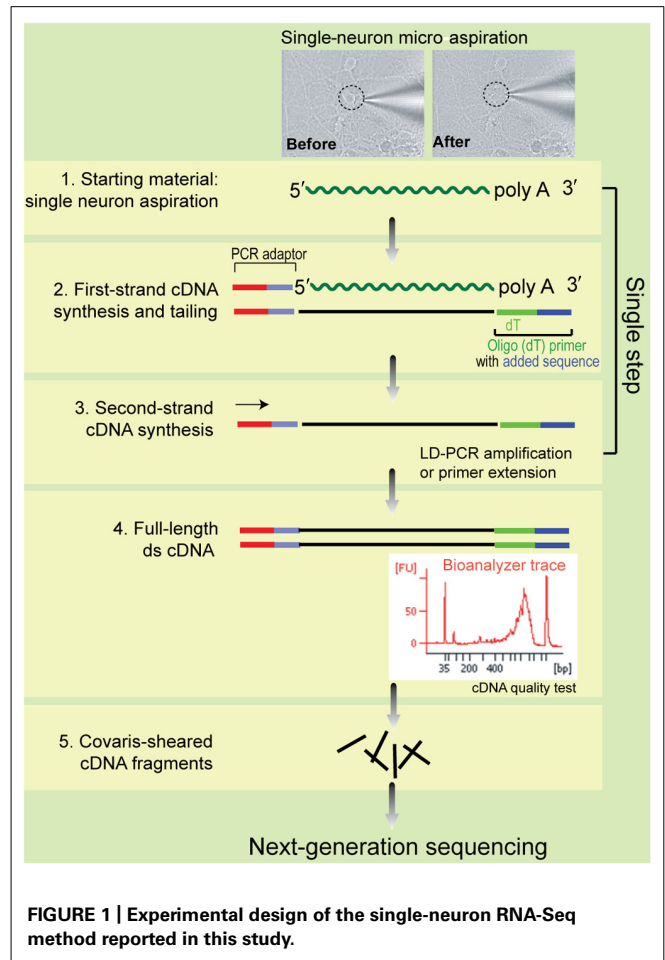
Table 1 | A list of samples used in our RNA-Seq experiment. The live neurons were retrieved from the same brain slice with normal electrophysiological properties.

Sample	ID	Description
HCT20466	Neuron 1	Living neuron from brain slice
HCT20468	Neuron 2	Living neuron from brain slice
HCT20469	Neuron 3	Living neuron from brain slice
HCT20470	4-neuron pool	A pool of 4 neurons from brain slice
HCT20575	Cell 1	Single neuronal cell from DIV12 culture, #18
HCT20576	Cell 2	Single neuronal cell from DIV12 culture, #19
HCT20577	Cell 3	Single neuronal cell from DIV12 culture, #20
HCT20578	Cell 4	Single neuronal cell from DIV12 culture, #21
HCT20579	Cell 5	Single neuronal cell from DIV12 culture, #22
HCT20580	Cell 6	Single neuronal cell from DIV12 culture, #24

instrumentation or labeling by GFP. Double-stranded cDNA was generated from single cell content, and used in standard Illumina library construction and next-generation sequencing with an Illumina Genome Analyzer IIx (Figure 1). On average, ~22 million 50-bp reads were generated from each neuron. The sequence data was converted to FASTQ files, aligned against the mouse genome build NCBI M37 using TopHat software, and then analyzed by the Cufflinks program to quantify the expression levels for each transcript and gene.

QUALITY CONTROL OF SINGLE-NEURON RNA-Seq

In conventional RNA-Seq experiments, the total quantity of RNA samples is typically sufficiently large (>100 ng), so that contamination of foreign RNA is not a major concern. However, in our experiments, given the limited amount of starting materials (typically around ~10 pg of total RNA), careful sample handling techniques become extremely important to reduce the possibility of contamination from multiple sources. In fact, our first set of experiments failed in that the vast majority of reads cannot be aligned and were subsequently traced back to pollen DNA, as the experiments were performed on an open lab bench during spring time in Los Angeles. It is therefore recommended that the retrieval of cell lysate and library preparation be performed under a laminar flow hood with extra caution to avoid any possible contamination. We found two critical things to reduce RNA contamination and degradation: (1) A good DIC optics will reveal clear neuronal boundary from surrounding cells and tissues. In combination with a continuous adjust of Z focus plane and a strong, pulsate suction using a glass syringe (glass plunger lubricated with water), we were able to minimize the suction of surrounding tissues/cell processes into the electrode. Nonetheless, possibility still exists that a small fraction of surrounding tissues and/or incomplete suction of neuronal content could contribute to the heterogeneity in reads (and low Spearman rank). (2) We have optimized the RnaseOUT concentrations in the electrode internal/harvesting solution (0.4 unit/ μ l or 1%) and the reaction



buffer (5%), while other concentrations seemed to result in significant RNA degradation. It should be noted that microaspiration is incapable to capture neuronal contents in distal compartments, such as axons and dendrites, which may contain a small quantity of mRNA (Eberwine et al., 2001; Batish et al., 2012). This caveat notwithstanding, single-neuron RNA-Seq should faithfully capture the majority of transcriptome. A major advantage of this technique is that it can be done in live neurons in brain slices after electrophysiology recording, therefore offers valuable insights on how neuronal gene network determines each neuronal types (for example, neural stem cells differentiation, developmental specification of major neuronal types, such as excitatory projection neurons vs. local inhibitory neurons). Another advantage is that it may be possible to further develop aspiration-based techniques to determine the 3D organization of tissues such as the brain: we can envision a microfabricated array of thousands of micropipettes, under computer control, advancing through tissue, stopping, and aspirating once a cell is detected and then continuing to advance. Data from such an array would then enable 3D reconstruction of the tissue transcriptome. It is also conceivable that future RNA-Seq through the use of laser capture microdissection or other higher throughput microdissection devices may be a viable alternative to the single electrode aspiration approach.

REPRODUCIBILITY OF RNA QUANTIFICATION

To evaluate the reproducibility of RNA quantification, we compared gene expression levels (expressed as RPKM) of three individual and one pool of four neighboring mouse layer 5 pyramidal neurons from slices. The correlation of gene expression between single and pooled neurons was compared (Tables 2 and 3, Figure 2). Given the skewed distribution of RPKM values, Spearman's rank correlation was used, rather than Pearson's correlation. On average, the pairwise rank correlation coefficients of the RPKM values between individual pyramidal neurons *in situ* was 0.25, yet the same measure between each individual neuron and a 4-neuron pool was 0.35. In comparison, we also examined six single cultured mouse embryonic hippocampal neurons harvested at 12 DIV. The average pairwise Spearman rank correlation coefficients was 0.51 for six individual neurons grown in cell culture (Figure 3). Although we did not perform technical replicates in RNA-Seq, our previous experience showed that correlations between technical replicates are typically >0.95, so that noises from the sequencing run per se are unlikely to have major influence. Our results suggest that the lower correlation value for different neurons in brain slices may reflect cell type heterogeneity of individual neighboring neurons. Nevertheless, the observed correlations in both experimental paradigms were rather low, implicating either a high level of variability in the expression profile of individual neurons or the presence of substantial noise in the RNA-Seq data. To further explore the latter possibility, we examined the sequence alignments of the housekeeping gene *Actb*

in three neurons from a brain slice (Figure 4). These samples differed in their coverage levels for different exons on the 5'-end of mRNA. This observation is consistent with differential degradation of mRNA or synthesis of cDNA, and may contribute to reduction of correlation between RNA-Seq data from different individual cells.

COMPARISON TO PREVIOUS RNA-Seq AND MICROARRAY STUDIES ON MOUSE TISSUES

To further evaluate the sequencing data with respect to the biology of tissue-specific gene expression, we compared the results to a previous study employing the RNA-Seq technology on mouse tissues (Mortazavi et al., 2008). The previous study examined poly(A)-selected RNA from three mouse tissues (brain, liver, and skeletal muscle), establishing the RPKM as the gold standard for quantifying gene expression levels from RNA-Seq. We found that the Spearman's rank correlation coefficient between Neuron 1 in the present study and the mouse brain in the Mortazavi et al.'s (2008) study was 0.41. In comparison, the same measure was only 0.18 between Neuron 1 and mouse liver in the Mortazavi et al.'s (2008) study, and only 0.21 between Neuron 1 and mouse muscle in the Mortazavi et al.'s study, respectively. This comparative analysis confirmed that RNA-Seq measurements in our study are more similar to gene expression patterns of brain tissues than other tissues.

We next compared the current single-neuron data to a recent RNA-Seq study on mouse neocortical layers (Belgard et al., 2011). This study sequenced transcriptomes from layers 2 to 6b of different areas (primary and secondary) of the adult mouse somatosensory cortex. The Spearman's rank correlation coefficients between neuron 1 and their cortical layer 2/3, 4, 5, 6, and 6b neurons are 0.38, 0.33, 0.34, 0.33, and 0.34, respectively. In comparison, the Spearman's rank correlation coefficients between our cultured Cell 1 and their cortical layer 2/3, 4, 5, 6, and 6b neurons are 0.41, 0.33, 0.36, 0.36, and 0.35, respectively. The slightly increased correlation for cultured neurons as compared to individual neurons may reflect the greater biological variability of neurons from brain slices. The present comparative analysis thus confirmed that data from single-neuron RNA-Seq does indeed recapitulate to expression of genes that are highly expressed in the mouse brain.

To compare our study with previous microarray studies, we further assessed the GNF SymAtlas data set on dozens of mouse tissues, generated on the Affymetrix GNF1M microarray (Su et al., 2004; Lattin et al., 2008). The gene expression values on all tissues were directly taken from the downloaded files. The Spearman's rank correlation coefficient between Neuron 1 in the present study and cerebral cortex, frontal cortex and hippocampus in the GNF data set were 0.23, 0.22, and 0.21, respectively. In comparison, the correlations of RNA-Seq with microarray data from other mouse tissues were much lower. For example, the correlation with retina, heart, lymph node, pancreas, oocyte, and fertilized egg are 0.15, 0.12, 0.11, 0.09, 0.07, and 0.06, respectively. This comparative analysis thus confirmed that results from single-neuron RNA-Seq were more similar to microarray studies on neuronal tissues than other tissues, but the correlations of measurements were generally very poor.

Table 2 | Pairwise Spearman's rank correlation coefficient for between all live neurons from brain slice.

ID	Neuron 1	Neuron 2	Neuron 3	4-neuron pool
Neuron 1	–	1183	1244	1323
Neuron 2	0.18	–	1298	1302
Neuron 3	0.30	0.28	–	1715
4-neuron pool	0.29	0.37	0.40	–

The upper right triangle shows the number of common genes (RPKM > 0 in both samples) used in calculation, whereas the lower left triangle shows the Spearman's rank correlation coefficients.

Table 3 | Pairwise Spearman's rank correlation coefficient for between all neuronal cells from cell cultures.

ID	Cell 1	Cell 2	Cell 3	Cell 4	Cell 5	Cell 6
Cell 1	–	2555	2824	2063	2376	2696
Cell 2	0.52	–	2599	1914	2213	2508
Cell 3	0.57	0.54	–	2350	2665	3099
Cell 4	0.42	0.35	0.39	–	1955	2155
Cell 5	0.60	0.57	0.61	0.35	–	2595
Cell 6	0.58	0.59	0.61	0.43	0.59	–

The upper right triangle shows the number of common genes (RPKM > 0 in both samples) used in calculation, whereas the lower left triangle shows the Spearman's rank correlation coefficients.

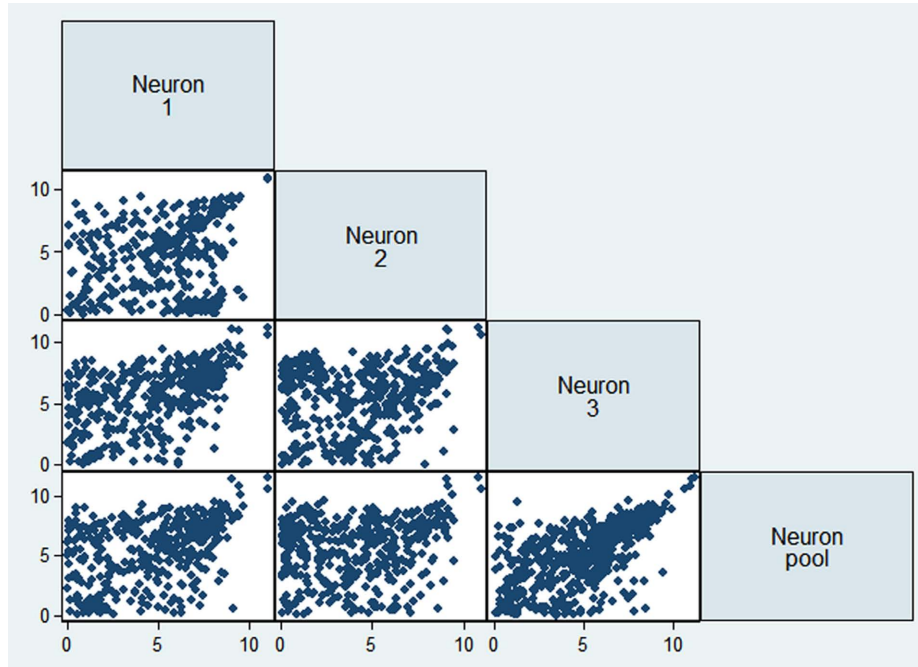


FIGURE 2 | Comparison of gene expression values (as logarithm of RPKM) from three single neurons and the 4-neuronal pool.

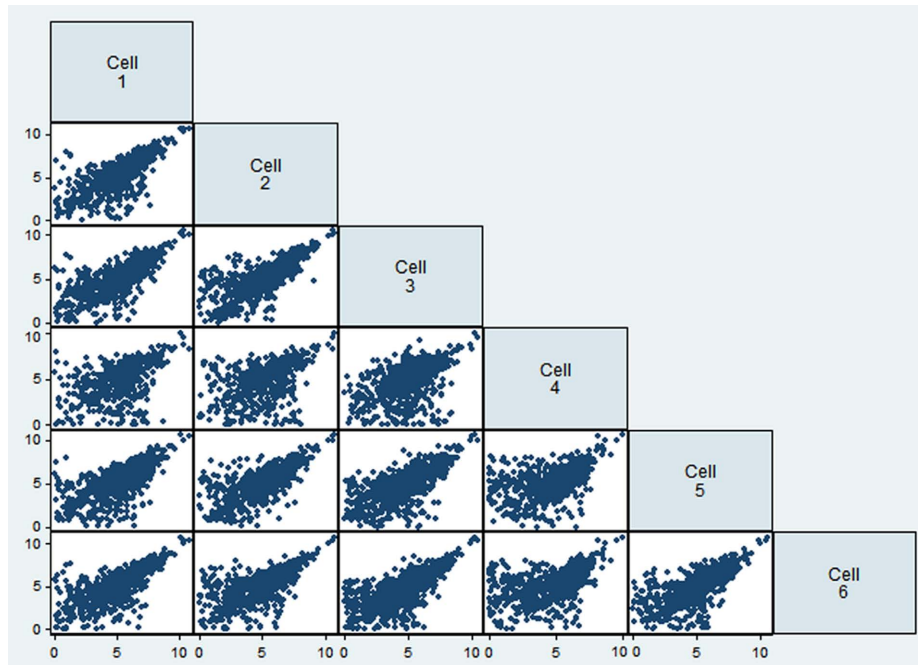
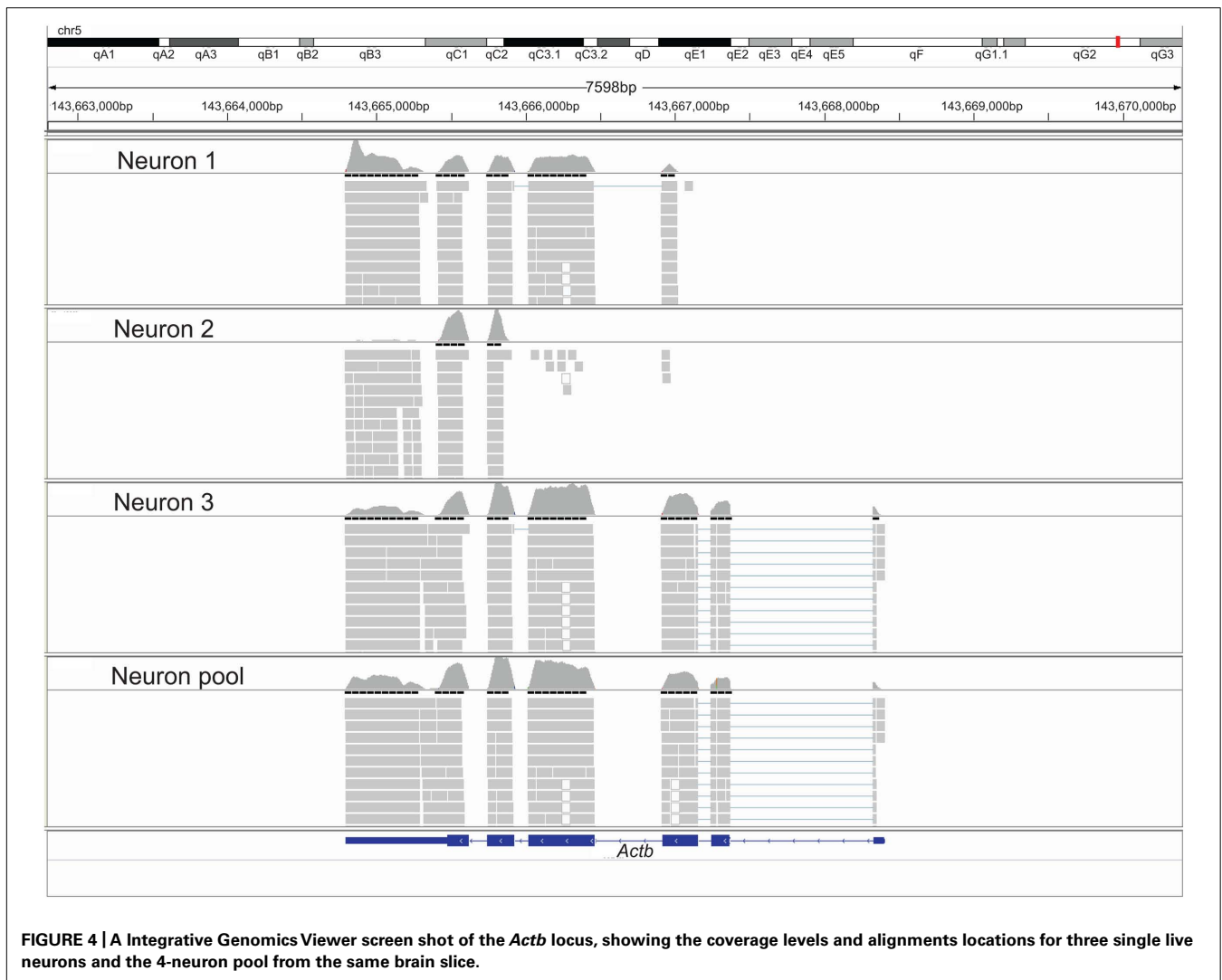


FIGURE 3 | Comparison of gene expression values (as logarithm of RPKM) for six single neuronal cells from cell culture.

CONCLUSION

In this study, we demonstrated the technical feasibility and reproducibility of single-neuron RNA-Seq, as a method that can be applied in standard neuronal cultures or electrophysiology

experiments. The present adaptations of the single-neuron RNA-Seq method do not require sophisticated laboratory equipment, making it technically accessible broadly. Furthermore, we compared the reproducibility of RNA-Seq data between different



neurons in the same brain slice or different cells in the same neuronal culture. These data suggest that a large portion of the observed differences in gene expression are due to cell-to-cell phenotypic variability rather than noise due to technical issues.

REFERENCES

- Batish, M., Van Den Bogaard, P., Kramer, F. R., and Tyagi, S. (2012). Neuronal mRNAs travel singly into dendrites. *Proc. Natl. Acad. Sci. U.S.A.* 109, 4645–4650.
- Belgard, T. G., Marques, A. C., Oliver, P. L., Abaan, H. O., Sirey, T. M., Hoerder-Suabedissen, A., Garcia-Moreno, E., Molnar, Z., Margulies, E. H., and Ponting, C. P. (2011). A transcriptomic atlas of mouse neocortical layers. *Neuron* 71, 605–616.
- Birney, E., Andrews, T. D., Bevan, P., Caccamo, M., Chen, Y., Clarke, L., Coates, G., Cuff, J., Curwen, V., Cutts, T., Down, T., Eyraas, E., Fernandez-Suarez, X. M., Gane, P., Gibbins, B., Gilbert, J., Hammond, M., Hotz, H. R., Iyer, V., Jekosch, K., Kahari, A., Kasprzyk, A., Keefe, D., Keenan, S., Lehvaslaiho, H., Mcvicker, G., Melsopp, C., Meidl, P., Mongin, E., Pettett, R., Potter, S., Proctor, G., Rae, M., Searle, S., Slater, G., Smedley, D., Smith, J., Spooner, W., Stabenau, A., Stalker, J., Storey, R., Ureta-Vidal, A., Woodward, K. C., Cameron, G., Durbin, R., Cox, A., Hubbard, T., and Clamp, M. (2004). An overview of Ensembl. *Genome Res.* 14, 925–928.
- Bota, M., Dong, H. W., and Swanson, L. W. (2003). From gene networks to brain networks. *Nat. Neurosci.* 6, 795–799.
- Citri, A., Pang, Z. P., Sudhof, T. C., Wernig, M., and Malenka, R. C. (2012). Comprehensive qPCR profiling of gene expression in single neuronal cells. *Nat. Protoc.* 7, 118–127.
- Eberwine, J., and Bartfai, T. (2011). Single cell transcriptomics of hypothalamic warm sensitive neurons that control core body temperature and fever response Signaling asymmetry and an extension of chemical neuroanatomy. *Pharmacol. Ther.* 129, 241–259.
- Eberwine, J., Miyashiro, K., Kacharina, J. E., and Job, C. (2001). Local translation of classes of mRNAs that are targeted to neuronal dendrites. *Proc. Natl. Acad. Sci. U.S.A.* 98, 7080–7085.
- Esumi, S., Wu, S. X., Yanagawa, Y., Obata, K., Sugimoto, Y., and Tamamaki, N. (2008). Method for single-cell microarray analysis and application to gene-expression profiling of GABAergic neuron progenitors. *Neurosci. Res.* 60, 439–451.
- Gustincich, S., Contini, M., Gariboldi, M., Puopolo, M., Kadota, K., Bono, H., Lemieux, J., Walsh, P., Carninci, P., Hayashizaki, Y., Okazaki, Y., and Raviola, E. (2004). Gene discovery in genetically labeled single dopaminergic neurons of the retina.

ACKNOWLEDGMENTS

The authors thank members of the Zilkha Neurogenetic Institute at University of Southern California for helpful discussions and comments.

- Proc. Natl. Acad. Sci. U.S.A.* 101, 5069–5074.
- Iscove, N. N., Barbara, M., Gu, M., Gibson, M., Modi, C., and Winegarden, N. (2002). Representation is faithfully preserved in global cDNA amplified exponentially from sub-picomogram quantities of mRNA. *Nat. Biotechnol.* 20, 940–943.
- Kamme, F., Salunga, R., Yu, J., Tran, D. T., Zhu, J., Luo, L., Bittner, A., Guo, H. Q., Miller, N., Wan, J., and Erlander, M. (2003). Single-cell microarray analysis in hippocampus CA1: demonstration and validation of cellular heterogeneity. *J. Neurosci.* 23, 3607–3615.
- Lattin, J. E., Schroder, K., Su, A. I., Walker, J. R., Zhang, J., Wiltshire, T., Saijo, K., Glass, C. K., Hume, D. A., Kellie, S., and Sweet, M. J. (2008). Expression analysis of G protein-coupled receptors in mouse macrophages. *Immunome Res.* 4, 5.
- Lin, D. M., Loveall, B., Ewer, J., Deitcher, D. L., and Sucher, N. J. (2007). Characterization of mRNA expression in single neurons. *Methods Mol. Biol.* 399, 133–152.
- Luo, L., Callaway, E. M., and Svoboda, K. (2008). Genetic dissection of neural circuits. *Neuron* 57, 634–660.
- Luo, Z., and Geschwind, D. H. (2001). Microarray applications in neuroscience. *Neurobiol. Dis.* 8, 183–193.
- Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* 5, 621–628.
- Navin, N., Kendall, J., Troge, J., Andrews, P., Rodgers, L., McIndoo, J., Cook, K., Stepansky, A., Levy, D., Esposito, D., Muthuswamy, L., Krasnitz, A., McCombie, W. R., Hicks, J., and Wigler, M. (2011). Tumour evolution inferred by single-cell sequencing. *Nature* 472, 90–94.
- Ozsolak, F., and Milos, P. M. (2011). RNA sequencing: advances, challenges and opportunities. *Nat. Rev. Genet.* 12, 87–98.
- Qiu, S., Anderson, C. T., Levitt, P., and Shepherd, G. M. (2011). Circuit-specific intracortical hyperconnectivity in mice with deletion of the autism-associated Met receptor tyrosine kinase. *J. Neurosci.* 31, 5855–5864.
- Qiu, S., Champagne, D. L., Peters, M., Catania, E. H., Weeber, E. J., Levitt, P., and Pimenta, A. F. (2010). Loss of limbic system-associated membrane protein leads to reduced hippocampal mineralocorticoid receptor expression, impaired synaptic plasticity, and spatial memory deficit. *Biol. Psychiatry* 68, 197–204.
- Qiu, S., and Weeber, E. J. (2007). Reelin signaling facilitates maturation of CA1 glutamatergic synapses. *J. Neurophysiol.* 97, 2312–2321.
- Schuster, S. C. (2008). Next-generation sequencing transforms today's biology. *Nat. Methods* 5, 16–18.
- Stahlberg, A., Kubista, M., and Aman, P. (2011). Single-cell gene-expression profiling and its potential diagnostic applications. *Expert Rev. Mol. Diagn.* 11, 735–740.
- Su, A. I., Wiltshire, T., Batalov, S., Lapp, H., Ching, K. A., Block, D., Zhang, J., Soden, R., Hayakawa, M., Kreiman, G., Cooke, M. P., Walker, J. R., and Hogenesch, J. B. (2004). A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc. Natl. Acad. Sci. U.S.A.* 101, 6062–6067.
- Tang, F., Barbacioru, C., Bao, S., Lee, C., Nordman, E., Wang, X., Lao, K., and Surani, M. A. (2010a). Tracing the derivation of embryonic stem cells from the inner cell mass by single-cell RNA-Seq analysis. *Cell Stem Cell* 6, 468–478.
- Tang, F., Barbacioru, C., Nordman, E., Li, B., Xu, N., Bashkurov, V. I., Lao, K., and Surani, M. A. (2010b). RNA-Seq analysis to capture the transcriptome landscape of a single cell. *Nat. Protoc.* 5, 516–535.
- Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B. B., Siddiqui, A., Lao, K., and Surani, M. A. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. *Nat. Methods* 6, 377–382.
- Tang, F., Lao, K., and Surani, M. A. (2011). Development and applications of single-cell transcriptome analysis. *Nat. Methods* 8, S6–S11.
- Tietjen, I., Rihel, J. M., Cao, Y., Koentges, G., Zakhary, L., and Dulac, C. (2003). Single-cell transcriptional analysis of neuronal progenitors. *Neuron* 38, 161–175.
- Trapnell, C., Pachter, L., and Salzberg, S. L. (2009). TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25, 1105–1111.
- Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., Van Baren, M. J., Salzberg, S. L., Wold, B. J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28, 511–515.
- Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* 10, 57–63.
- Zhang, W., Wang, H., Song, S. W., and Fuller, G. N. (2002). Insulin-like growth factor binding protein 2: gene expression microarrays and the hypothesis-generation paradigm. *Brain Pathol.* 12, 87–94.

Conflict of Interest Statement: Gary P. Schroth, Robin Li, and Shujun Luo are employees of Illumina, Inc. The authors declare no other conflict of interests.

Received: 14 April 2012; accepted: 19 June 2012; published online: 06 July 2012.

Citation: Qiu S, Luo S, Evgrafov O, Li R, Schroth GP, Levitt P, Knowles JA and Wang K, (2012) Single-neuron RNA-Seq: technical feasibility and reproducibility. *Front. Gene.* 3:124. doi: 10.3389/fgene.2012.00124

This article was submitted to *Frontiers in Neurogenomics*, a specialty of *Frontiers in Genetics*.

Copyright © 2012 Qiu, Luo, Evgrafov, Li, Schroth, Levitt, Knowles and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.