



Grouping and classifying electrophysiologically-defined classes of neocortical neurons by single cell, whole-genome expression profiling

Tatiana Subkhankulova¹, Kojiro Yano², Hugh P. C. Robinson² and Frederick J. Livesey^{1*}

¹ Gurdon Institute and Department of Biochemistry, University of Cambridge, Cambridge, UK

² Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge, UK

Edited by:

Jean Rossier,
Ecole Supérieure de Physique et de
Chimie Industrielles, France

Reviewed by:

Quentin Perrenoud,
Ecole Supérieure de Physique et de
Chimie Industrielles ParisTech, France
Andreas H. Burkhalter, Washington
University School of Medicine, USA
Jan Gründemann, University College
London, UK

*Correspondence:

Frederick J. Livesey, Gurdon Institute
and Department of Biochemistry
University of Cambridge, Tennis Court
Road, Cambridge, CB2 1QN, UK.
e-mail: rick@gurdon.cam.ac.uk

The diversity of neuronal cell types and how to classify them are perennial questions in neuroscience. The advent of global gene expression analysis raised the possibility that comprehensive transcription profiling will resolve neuronal cell types into groups that reflect some or all aspects of their phenotype. This approach has been successfully used to compare gene expression between groups of neurons defined by a common property. Here we extend this approach to ask whether single neuron gene expression profiling can prospectively resolve neuronal subtypes into groups, independent of any phenotypic information, and whether those groups reflect meaningful biological properties of those neurons. We applied methods we have developed to compare gene expression among single neural stem cells to study global gene expression in 18 randomly picked neurons from layer II/III of the early postnatal mouse neocortex. Cells were selected by morphology and by firing characteristics and electrical properties, enabling the definition of each cell as either fast- or regular-spiking, corresponding to a class of inhibitory interneurons or excitatory pyramidal cells. Unsupervised clustering of young neurons by global gene expression resolved the cells into two groups and those broadly corresponded with the two groups of fast- and regular-spiking neurons. Clustering of the entire, diverse group of 18 neurons of different developmental stages also successfully grouped neurons in accordance with the electrophysiological phenotypes, but with more cells misassigned among groups. Genes specifically enriched in regular spiking neurons were identified from the young neuron expression dataset. These results provide a proof of principle that single-cell gene expression profiling may be used to group and classify neurons in a manner reflecting their known biological properties and may be used to identify cell-specific transcripts.

Keywords: neocortex, interneuron, pyramidal cell, expression analysis, classification, firing properties

INTRODUCTION

The question of the diversity of neuron types is a challenging contemporary problem in neuroscience (Nelson et al., 2006). It is generally held that a definitive parts list is essential to the accurate description of neuronal circuits and their potential functions (Bota and Swanson, 2007). However, the degree of neuronal diversity and its functional importance are areas of much debate, with a range of views on the degree to which neurons differ and the extent to which they should be grouped (see for example, Masland and Raviola, 2000). Much of the difficulty in defining neuronal cell types has stemmed from the absence of agreement over which neuronal properties can and should be used to define a neuronal cell type. Current and previous strategies for classifying neurons in any region of the central nervous system (CNS) rely on the use of one or more neuronal property, such as morphology, neurotransmitters, passive and active electrical properties, and expression of small numbers of molecular markers, typically proteins (Ascoli et al., 2008).

The use of those approaches is exemplified by the concerted effort by many researchers to define the diversity of GABAergic interneurons in the cerebral cortex (Ascoli et al., 2008). Despite

comprising 20% of the neurons in the mouse cortex (excitatory pyramidal cells constitute the other 80%), GABAergic interneurons are highly diverse in terms of morphology and electrical and molecular properties (Ascoli et al., 2008). Ongoing studies have emphasized the notable diversity of GABAergic interneurons when classifying cells by morphology, projections and electrical properties (Helmstaedt et al., 2009a,b), or by neuropeptide expression (Karagiannis et al., 2009). Classifications of interneurons based on electrical properties do not directly correspond to those based on molecular markers (Markram et al., 2004), although it is possible to classify neurons by single cell RT-PCR analysis of expression of key genes, such as ion channels (Toledo-Rodriguez et al., 2004). Excitatory pyramidal neurons also exhibit considerable diversity in terms of morphology and connectivity, although to a lesser degree than interneurons (Thomson and Lamy, 2007).

In addition to now classical methods of defining cell types by mRNA and protein expression, there have been several notable studies using global expression profiling of defined groups of neurons to identify novel cortical cell type-specific transcripts (Arlotta et al., 2005; Rossner et al., 2006; Sugino et al., 2006). Retrograde

labelling from targets was used to isolate two classes of layer V neurons, enabling the identification of cell-type specific genes by expression profiling (Arlotta et al., 2005). In a complementary approach, twelve mouse forebrain neuronal populations were isolated by cell-specific transgene expression and expression profiled to identify cell-specific gene expression combinations (Sugino et al., 2006). A taxonomy of cell types based solely on the gene expression data correctly reflected the expected relationships between the different populations, confirming that global expression profiling can correctly identify neuronal subtypes (Sugino et al., 2006).

Here we extend this approach to ask whether single neuron gene expression profiling can prospectively resolve neuronal subtypes into groups, independent of any phenotypic information, and whether those groups reflect meaningful biological properties of those neurons. To do so, we applied methods we developed to compare gene expression among single neural stem cells (Subkhankulova and Livesey, 2006; Subkhankulova et al., 2008) to study global gene expression between single neurons from layer II/III of mouse neocortex (Yano et al., 2006).

In proof of principle experiments, we find that global expression profiling of single cortical neurons can be used prospectively to group and classify neurons in a manner reflecting their physiological properties. The potential advantage of this approach is that it suggests that deep sampling of complex neuronal populations by single cell expression profiling may be a useful approach for defining neuronal complexity in the CNS. However, in this case we used electrophysiology and morphology to restrict the diversity of the population analysed. It will be of interest in future studies to carry out single cell expression studies of unselected neuronal populations to assess whether it is possible to use such data to generate biologically meaningful groupings of neuronal types.

MATERIALS AND METHODS

PATCH-CLAMP RECORDING FROM NEOCORTICAL NEURONS

Forebrains were acutely harvested from postnatal CD1 mice ranging in age from 5 to 20 days sacrificed according to United Kingdom Home Office and local ethical guidelines. 300 μ m thick sagittal slices of cortex containing the somatosensory region were cut with a vibratome (Microslicer DTK-3000, D. S. K., Kyoto, Japan). During slicing, tissue was kept in a low-sodium solution of the following composition (in mM): 254 sucrose, 2.5 KCl, 26 NaHCO₃, 10 glucose, 1.25 NaH₂PO₄, 2 CaCl₂, and 1 MgCl₂. The recording Ringer solution contained (in mM): 125 NaCl, 2.5 KCl, 25 NaHCO₃, 25 glucose, 1.25 NaH₂PO₄, 2 CaCl₂, and 1 MgCl₂. Both slicing and recording solutions were equilibrated with 95% O₂, 5% CO₂ gas to a final pH of 7.4.

Single cortical neurons were identified in layer II/III of the slices using an infrared differential interference contrast microscope (Olympus BX50WI) and their electrophysiological properties were obtained by whole cell patch clamping as described (Tateno et al., 2005), with the addition of recombinant RNase inhibitors to the pipette solution. Somatic patch-pipette recordings were made with a Multiclamp 700-A amplifier (Axon Instruments, Foster City, CA, USA) in current-clamp mode, correcting for prenull liquid junction potential. Whole cell recording pipettes (Clark GC150T-7.5) of resistance 3.9–4.3 M Ω were filled with the following intracellular solution, based on that of Monyer and Jonas (1995): 140 mM KCl, 5 mM EGTA, 3 mM MgCl₂, 5 mM HEPES, pH adjusted to 7.3 with KOH;

in this case 200 units/ml of SUPERase-In RNase inhibitor (Ambion, Austin Texas, USA) were added and KCl concentration reduced to 125–130 mM to achieve the same final osmolarity as the Monyer and Jonas solution. Neurons for RNA harvesting were selected by morphology and on their firing properties in order to restrict the range of cell types sampled for gene expression profiling. Recordings were carried out at room temperature to reduce RNase activity. Signals were low-pass filtered at 5 kHz (4-pole Bessel filter) and sampled at 20 kHz with 12-bit resolution, using a National Instruments analog card, and custom software (MatDAQ, Hugh Robinson, 1995–2010), written in Matlab and C. Stimulus generation was also at a 20 kHz rate, with 12-bit resolution. All analysis was performed in Matlab (Mathworks, Natick, MA, USA). Classification of cells as regular spiking (RS) or fast spiking (FS) was based quantitatively on their clearly different spike shapes, and agreed with other features such as maximum spike rate and the qualitative pattern of spike firing at threshold (Tateno et al., 2004), and the pyramidal (RS) and nonpyramidal (FS) morphology under infrared differential interference contrast optics. Action potentials from responses just above threshold were convolved with the first derivative of a Gaussian function: $G'(t) = -t \exp[-t^2/(2\sigma^2)]$, where $\sigma = 1.3$ ms, to produce a smoothed first-derivative of the action potential voltage trajectory. This was then aligned by the first negative-going zero-crossing (corresponding to the AP peak), and the interval to the second, positive-going zero-crossing (corresponding to the interspike membrane potential minimum) was measured. FS cells are defined as showing intervals less than 10 ms, while RS cells had intervals of 20 ms or greater.

SINGLE CELL TOTAL mRNA AMPLIFICATION AND EXPRESSION PROFILING

The cytoplasmic contents of each neuron were harvested after recording by aspiration into the patch pipette and deposited in a total volume of 5 μ l. Total RNA was isolated from the cytoplasmic contents by the acid phenol-guanidinium extraction method (Trizol, Sigma; (Chomczynski and Sacchi, 1987). Total RNA was precipitated in ethanol in the presence of 100 pg polyinosinic acid (Sigma) and 1 μ l of linear polyacrylamide (Ambion) as carriers to reduce the loss of low copy number mRNAs. Total mRNA from each of 18 neuronal cells was globally amplified using a PCR-based technique (Brady and Iscove, 1993; Brady et al., 1995; Iscove et al., 2002) with minor variations (Subkhankulova and Livesey, 2006). PCR products were purified with the CyScribe GFX Purification kit (Amersham Bioscience) and labeled with Cy3/Cy5-modified dCTP using Klenow DNA polymerase (BD Bioscience). The labeled samples were hybridized against a reference of amplified cDNA pooled from all of the neurons studied. Gene expression was studied using mouse oligonucleotide arrays printed in-house (65-mer, 23232 oligonucleotides representing ~20,000 genes; Compugen oligonucleotide library) in pairs of dye-swapped hybridisations, as previously described (Subkhankulova and Livesey, 2006; Subkhankulova et al., 2008). Images of hybridised arrays were gathered on a microarray scanner (Axon Instruments GenePix). Data from single microarrays were extracted using the GenePix analysis package. The mean intensity of each feature (spot) and the local background at 532 nm (Cy3) and 635 nm (Cy5) were extracted. The background-corrected 635 nm/532 nm ratio was calculated for each feature as the initial measurement of relative mRNA levels for each gene represented on the array.

SINGLE CELL qRT-PCR

Single cell qRT-PCR was carried out using the globally amplified single cell cDNA as a template. Each real time PCR mix contained was carried out in a 10 μ l final volume, containing 0.2 μ l of each primer (10 mM) and 5 μ l 2 \times Master mix (DyNAmo Capillary SYBR Green qPCR Kit, Finnzymes). Intron-spanning primers were designed to Emx1, a pyramidal neuron-specific transcript (Chan et al., 2001), and GAD65/GAD2, a GABAergic interneuron-specific transcript (Katarova et al., 2000), using Primer3 software¹, based on gene transcript information from the Ensembl genome browser². Primer sequences were: Emx2, GAGCGAGCCTTTGAGAAGAA and CTGGAACCACACCTT CACCT; GAD65/GAD2, AGATCGCCCCTGTATTTGTG and GCATGGCATAATGTTGGAG. Real-time PCR was performed in a LightCycler (Roche Diagnostics) according to DyNAmo Capillary SYBR Green qPCR protocol. Ct values were determined using the maximum second derivative function in the LightCycler software (Roche Diagnostics). Generation of PCR products was confirmed by melting curve analysis and gel electrophoresis. qPCR results were classified as absent (–) if Ct values were higher than 35 and no single peak was observed from the melting curve, detectable above background (+/–) if a single peak was observed from the melting curve and Ct values were lower than 34, and expressed significantly above background (+) if a single melting curve peak was present and Ct values were lower than 30.

ANALYSIS OF MICROARRAY DATA

All microarray data were analysed using the Acuity platform (Axon Instruments/Molecular Dynamics). Expression data from each array were normalized by locally-weighted regression (lowess) analysis. For data analysis of each group of arrays, data were filtered to remove low intensity features, features reporting gene expression in less than two thirds of the arrays in the set and features that did not show evidence for enriched expression in at least three cells in a set (empirically set at a 1.5-fold enrichment). Genes and arrays were pair-wise, hierarchically clustered using Pearson's correlation coefficients as similarity metrics. Similar results were obtained using alternative hierarchical clustering metrics (Euclidean distance, for example), with minor variations in the assignment of cells within clusters. Hierarchical cluster dendrograms were built by iterative, pair-wise similarity comparisons, using Pearson's correlation coefficients.

To identify genes with robust differential enrichment between young fast- and regular-spiking neurons, Student's *t*-testing was used to identify genes differentially expressed between the two groups ($p < 0.05$). Expression ratios of genes identified as significantly differing in expression between the groups were averaged to rank genes for enrichment in RS or FS cells, and those groups were visualized by hierarchical clustering. *In situ* hybridization data for adult expression of genes identified as enriched in each cell class were extracted from the Allen Brain Atlas' mouse brain database (mouse.brain-map.org) (Lein et al., 2007).

¹<http://frodo.wi.mit.edu/primer3/>

²<http://www.ensembl.org>

RESULTS

CLASSIFICATION OF NEURONS INTO RS OR FS CELL TYPES FOR EXPRESSION PROFILING

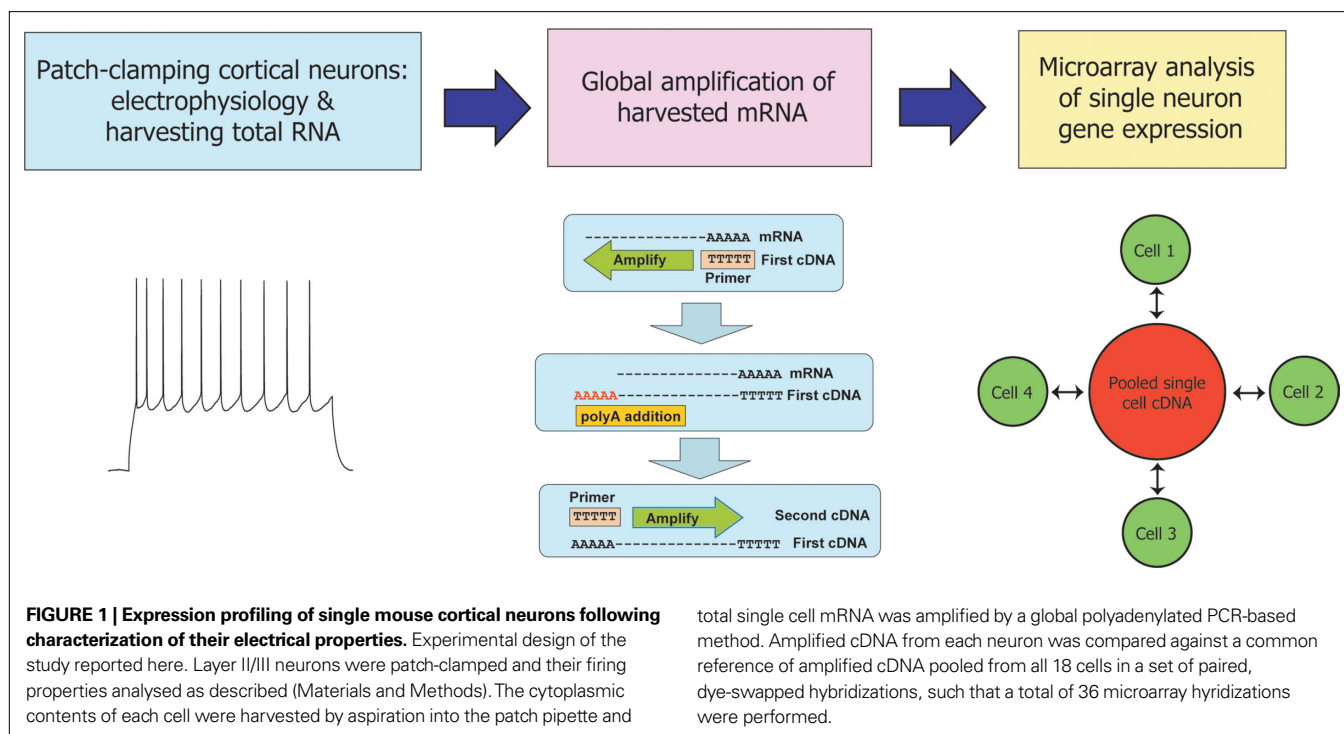
The goal of this study was to explore whether the global expression profiles of neuronal cells could be used to prospectively cluster neurons into groups that reflected their neuronal firing properties. For this purpose, we chose to expression profile neurons in two broad categories of morphologically and electrophysiologically-defined cortical neurons: regular-spiking cells of pyramidal morphology, corresponding to glutamatergic, excitatory pyramidal neurons (Connors and Gutnick, 1990); and fast-spiking, non-pyramidal cells, corresponding to GABAergic inhibitory interneurons of basket morphology that express parvalbumin (Kawaguchi and Kondo, 2002). To do so, we analysed 18 cells from layer II/III of mouse neocortex harvested between postnatal days 5 (P5) and P20: 10 young neurons (all P13 and older) and 8 immature neurons (aged between P5 and P10). The firing properties of this set of neurons were measured individually for each cell, following which the cytoplasm of each cell was harvested for global mRNA amplification (Figure 1).

The set of all recorded cells contained 7 RS and 11 FS cells. RS cells showed a pyramidal morphology, while FS cells generally had a nonpyramidal, multipolar morphology under infrared differential interference contrast microscopy (Connors et al., 1982). As in a previous study (Tateno et al., 2004), regular-spiking neurons (Figure 2A) showed an ability to fire repetitively at very low frequencies (<10 Hz), a relatively slow spike waveform and lacked a pronounced after hyperpolarization. Fast-spiking neurons (Figure 2B), in contrast showed a very fast action potential waveform, a large-amplitude, sharp after hyperpolarization, stable periodic firing only at a relatively high frequency (>10 Hz), and usually a high maximal firing rate. Depolarization block was pronounced at higher current levels in some recordings in each type of cell, potentially as a consequence of the high levels of RNase inhibitor, which is not a normal component of whole-cell pipette solutions.

In addition to qualitatively matching the well-known RS and FS firing patterns, we distinguished the two cell types on a quantitative basis, using the smoothed first-derivative of action potentials elicited just above threshold (see Materials and Methods for details). The interval between zero-crossings of this derivative, corresponding to the interval between the peak of the action potential and the deepest point of the after-hyperpolarization was tightly clustered (Figures 2C,D), below 10 ms for FS cells (blue), while it ranged from 20 to 60 ms for RS cells (red).

Neurons studied here demonstrated some diversity in firing patterns, however all could be clearly defined as FS or RS spiking types (Figure 3; see Table S1 in Supplementary Material for details). Particularly at young ages, depolarization block is pronounced and can obscure the typical adult firing pattern, but the spike shape and the pattern of spiking within a certain window of stimulus currents leading to periodic firing, enabled a quite robust classification of such cells.

After analysis and recording of the electrophysiological properties of each neuron, the cytoplasmic contents of each cell were harvested. Total RNA was purified from each cell, reverse transcribed to cDNA and globally amplified (see Materials and Methods for details). Typically, 32–34 cycles of PCR produced 15–20 μ g of amplified cDNA. Prior to global expression profiling, we further



investigated the expression of cell-specific transcripts in the single cells by quantitative PCR (qPCR) analysis, using the amplified cDNA as a template. Quantitative PCR for the pyramidal cell-specific transcription factor *Emx1* and for the GABAergic neuron-specific enzyme glutamic acid decarboxylase-2 (*GAD2/GAD65*) was used to distinguish between the two cell types (Figure 3). For example, cells 9 and 13 were defined as RS by firing properties (Table S1 in Supplementary Material) and expressed significant levels of *Emx1* and little or no *GAD65* mRNA (Figure 3). Similarly, cell 12 was classed as FS by firing properties and qPCR found that this cell was *GAD65*-positive and *Emx1*-low.

However, there are notable exceptions in which expression of these markers was not detected in cells with clear electrophysiological phenotypes. For example, *Emx1* was not detected in the RS cell 7. This absence of PCR-based detection of markers probably reflects drop-out or sampling of transcripts during the global cDNA amplification step, as we have observed previously in neural stem cells (Subkhankulova et al., 2008). In contrast, the RS cell 1 was found to have both *Emx1* and *GAD65* mRNA by qPCR, but has pyramidal morphology. It is unlikely that this cell corresponds to a regular-spiking, non-pyramidal cell. As discussed below, cluster analysis based on gene expression indicates that this cell is a regular-spiking pyramidal neuron. Therefore, it may be that the presence of *GAD65* mRNA represents contamination or a false-positive result due to the use of globally-amplified cDNA as the template for the subsequent qPCR.

INDIVIDUAL EXPRESSION PROFILING DISTINGUISHED TWO GROUPS OF YOUNG CELLS CORRESPONDING TO FS AND RS NEURONS

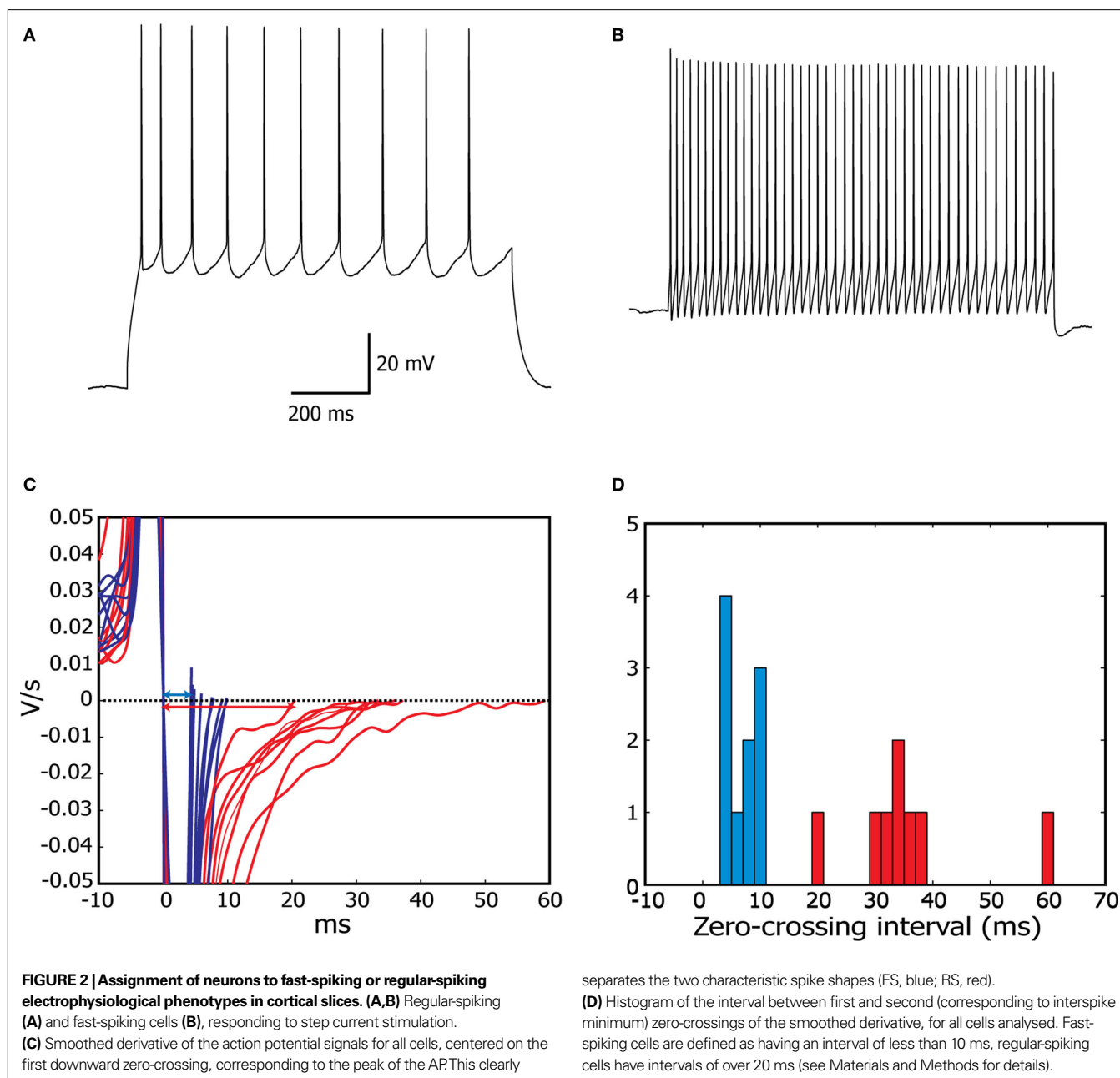
The experimental design here was one in which gene expression in each cell was compared to a reference sample of cDNA pooled from the entire set of 18 young and immature cells studied in

two dye-swap hybridisations. Statistical analysis found that the expression of the abundant transcript *GAPDH*, printed on the arrays more than 200 times, did not vary significantly among the cells (data not shown), confirming the overall quality of the combined procedures of RNA harvesting, amplification and microarray hybridisation.

To test the ability of global gene expression profiling to group and distinguish well-characterised types of neurons, we used the set of 10 young cells (P13–22) to examine if the classification of neurons by global expression analysis corresponded to the physiological classification of RS and FS cells. Cells were assigned to groups by unsupervised clustering according to the degree of similarity between pairs of cells (as calculated by Pearson correlation coefficients; Figure 4). The cells formed two distinct groups of four and six cells each, as defined by gene expression (Figure 4). Those two groups corresponded to the two electrophysiologically-defined groups, with the three RS cells forming one branch of the overall dendrogram, with one misplaced FS neuron, and a set of five FS cells plus one RS cell forming the other. Within the FS sub-tree there was a high degree of diversity of expression, although there were some cells that were highly similar in gene expression (cells 14, 15 and 19, Figure 4).

SINGLE-CELL EXPRESSION PROFILING PROSPECTIVELY CLASSIFIED RS AND FS NEURONS AT DIFFERING DEVELOPMENTAL STAGES

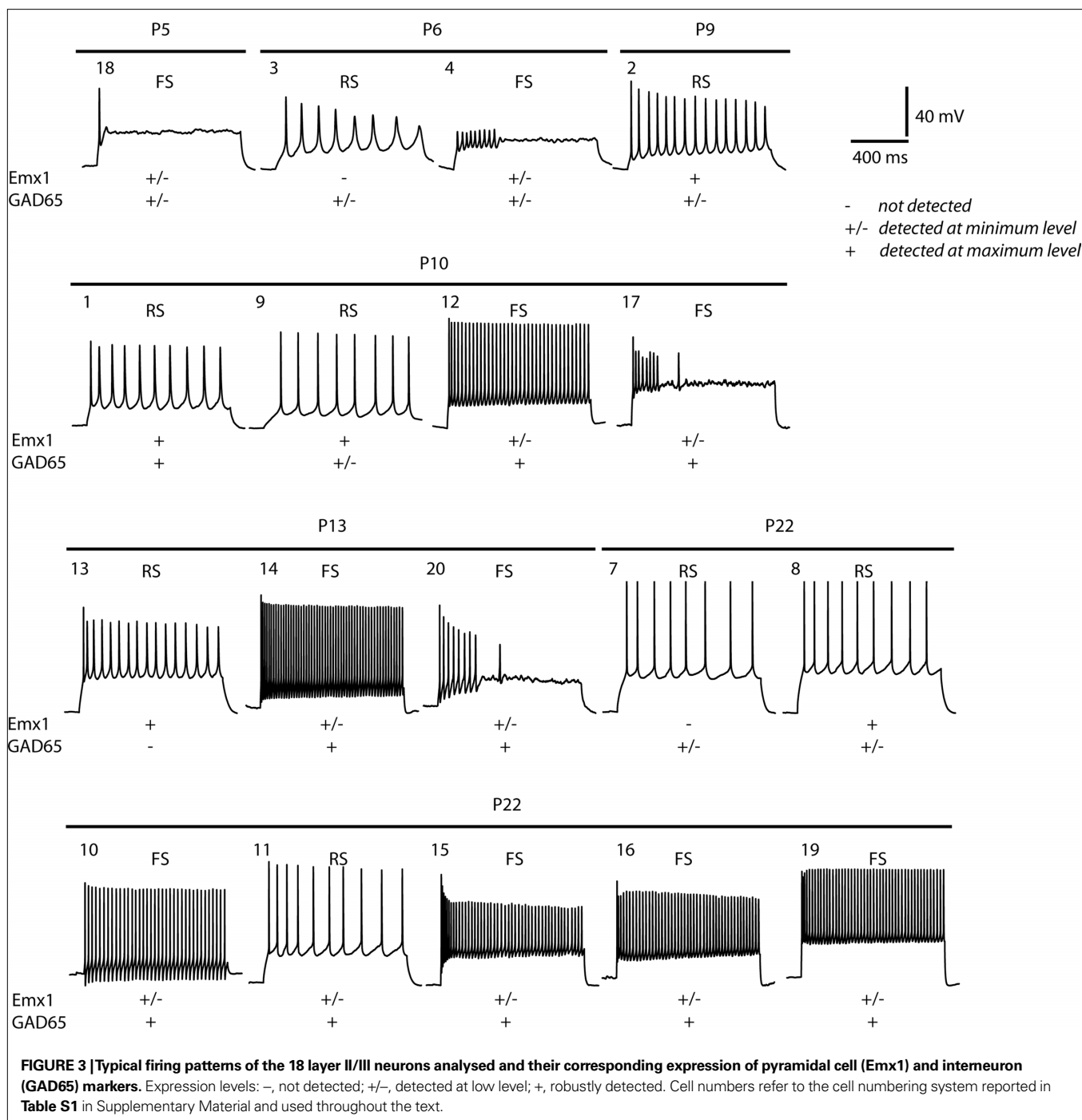
To study whether gene expression could group and classify a more complex population of neurons, we proceeded to use unsupervised cluster analysis to analyse the entire group of 18 layer II/III neurons ranging in age from P5 to P22. Based on firing properties, all neurons in this set could be designated as either RS or FS, but with some variation in their developmental stage, as reflected in their electrical properties (Figure 3; Table S1 in Supplementary Material).



Cluster analysis by gene expression of all 18 neurons, independent of their firing properties, identified three major clusters (Figure 5). The first group contained four regular-spiking neurons of a range of developmental stages (Figure 5). The second cluster of 6 cells contained four FS neurons, of which three were immature, together with two cells, 2 (P9) and 11 (P22), whose firing properties were those of typical RS cells. The third cluster contains 6 fast-spiking neurons and two regular-spiking neurons (one young and one immature). Thus unbiased clustering based on global gene expression grouped this complex population of single cortical neurons in a manner that broadly reflects their firing properties, with a subset of neurons whose cluster assignment did not accurately reflect their firing properties.

IDENTIFICATION OF TRANSCRIPTS ENRICHED IN REGULAR-SPIKING NEURONS COMPARED WITH FAST-SPIKING CELLS

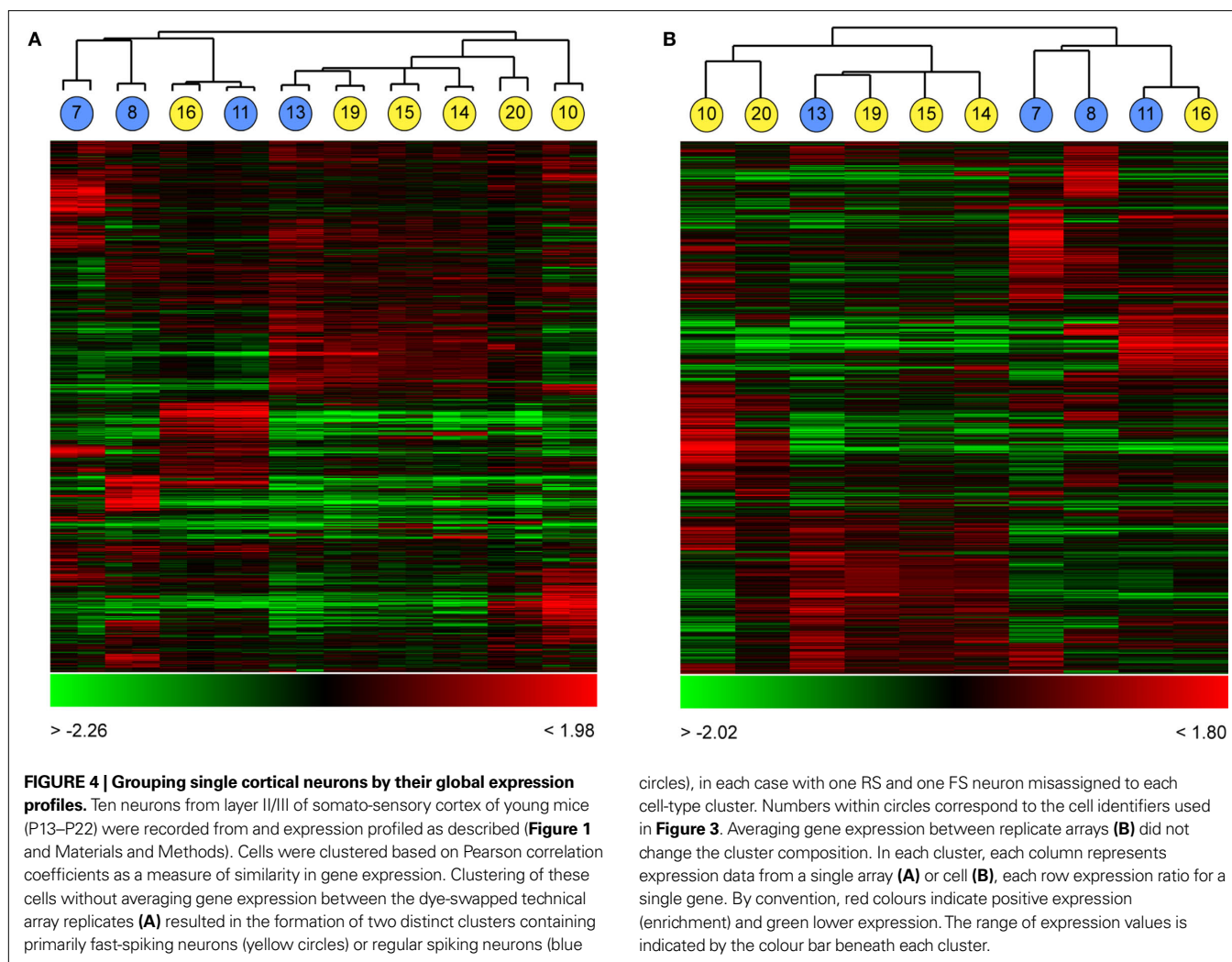
We used the firing properties and qPCR data to group the ten young (>P13) neurons into RS- and FS-type cells for the purpose of identifying genes differentially expressed between the two groups. Genes showing the most reproducible differences in expression between the two groups were selected using Student's *t*-test (see Materials and Methods for details), and the gene expression values for each gene averaged within each group. Those genes were then clustered for visualization (Figure 6A). At the statistical cut-off applied ($p < 0.05$), 482 mRNA transcripts with differential expression between the two groups were identified, 279 of which were enriched in RS cells and 203 enriched in FS cells (Table S2 in



Supplementary Material). The sets of genes enriched in each group were functionally heterogeneous, including transcription factors, intracellular enzyme and plasma membrane proteins (**Table S2** in Supplementary Material).

To explore whether the enriched expression of genes identified by the array analysis as such in RS or FS cells, we made use of the public, large-scale *in situ* hybridization dataset for the adult mouse brain, the Allen Brain Atlas (Lein et al., 2007). Many of the genes showing statistically significant differences in expression between the two groups of cells, but of low magnitude, were found to be

pan-neuronally expressed (data not shown). However, for those genes that were significantly enriched in RS cells, that is pyramidal cells of the neocortex, we found many examples of genes that were either cortex-specific or highly enriched in expression in cortical and, in many cases, hippocampal pyramidal cells, when the expression patterns were compared with that of *Emx1* (**Figure 6**). Several of those genes are not just enriched in cortex, but also demonstrate specific expression in layers 2–4 of the cortex (see **Figure 6B** for examples), demonstrating the ability of the single cell profiling approach to identify cell type-specific genes.

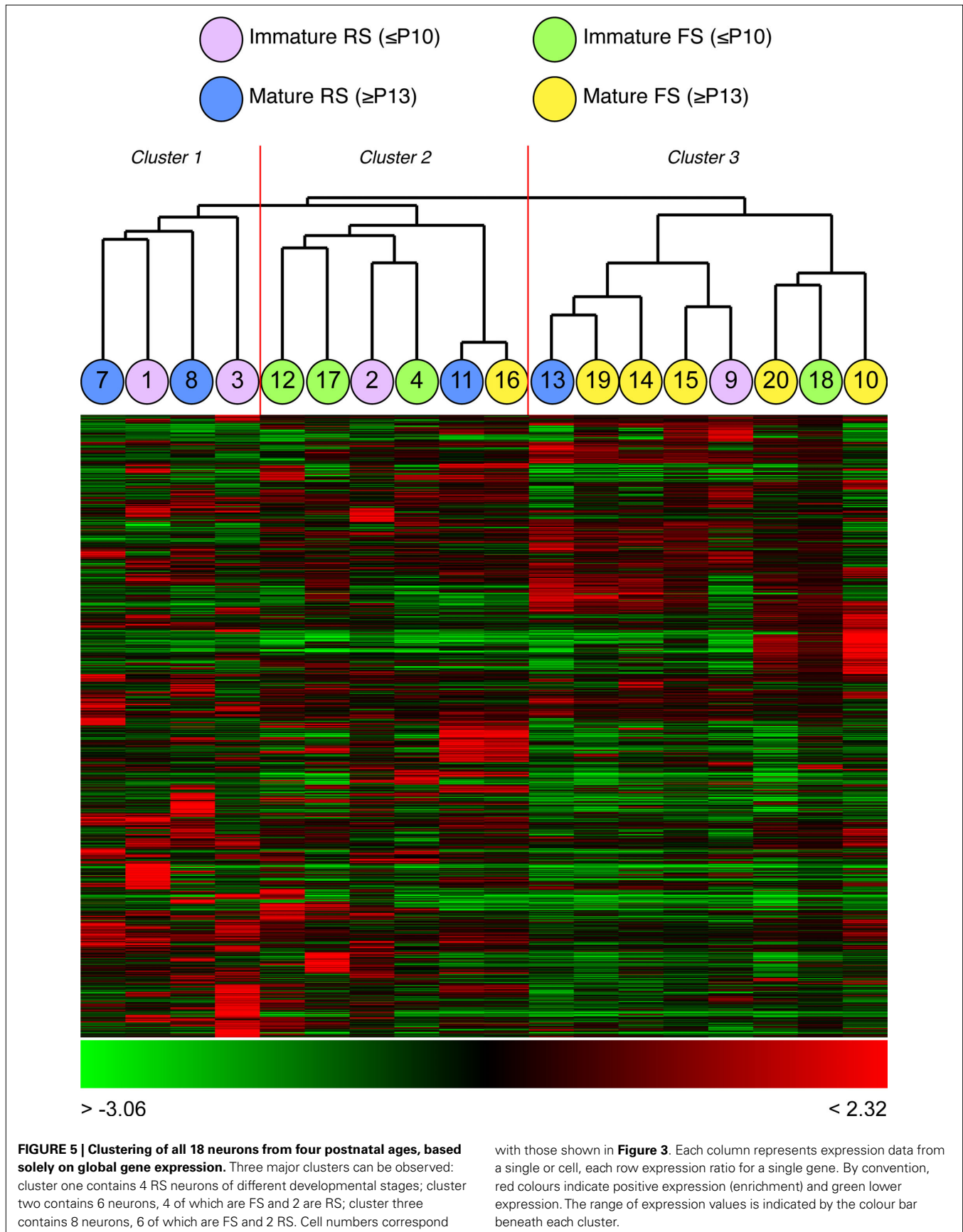


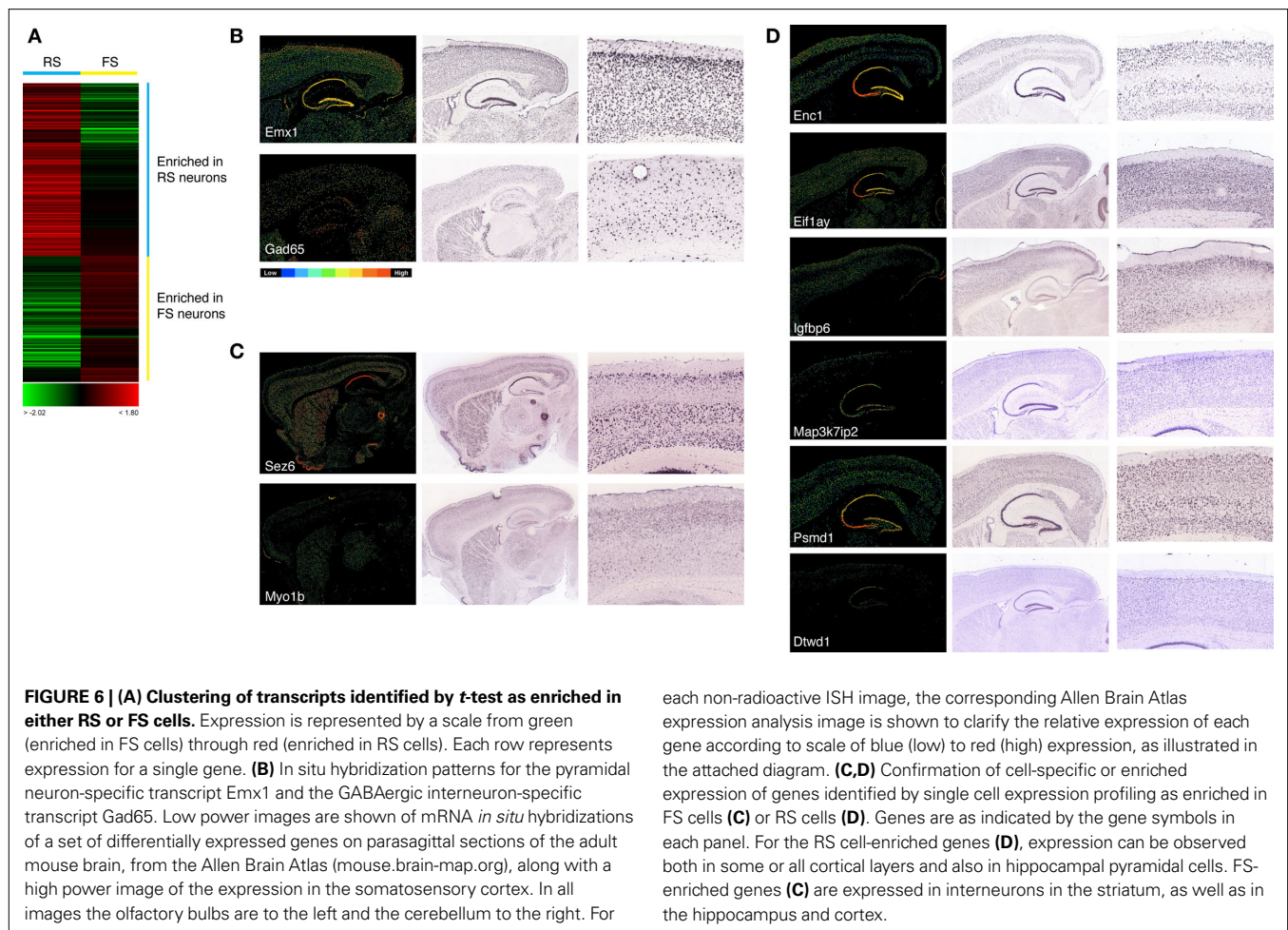
DISCUSSION

In this paper we have addressed the question of whether single neuron gene expression profiling can prospectively resolve neuronal subtypes into groups, independent of any phenotypic information, and whether those groups reflect meaningful biological properties of those neurons. To do so, we applied methods we have developed to compare gene expression among single neural stem cells to study global gene expression in 18 randomly picked neurons from layer II/III of mouse neocortex. The morphology, firing characteristics and electrical properties of the neurons enabled the classification of each cell as either fast- or regular-spiking, corresponding to inhibitory interneurons or excitatory pyramidal cells (Connors and Gutnick, 1990). Unsupervised clustering of neurons by global gene expression resolved the cells into three groups and those corresponded broadly with the groups of fast- and regular-spiking neurons. Genes specifically enriched in RS neurons were identified from the expression dataset and confirmed as such in the public mouse brain *in situ* hybridisation databases. These results provide an initial proof of principle that single-cell gene expression profiling has the potential to be used to group and classify neurons in a manner reflecting their known biological properties and can be used to identify novel, cell type-specific transcripts.

GLOBAL EXPRESSION PROFILING CAN GROUP SINGLE NEURONS INTO CLASSES THAT REFLECT MEANINGFUL BIOLOGICAL PROPERTIES

Previous expression profiling studies at the resolution of defined neuronal cell types have either focused on a small number of functionally important transcripts, such as ion channels or neurotransmitter systems (Cauli et al., 1997, 2000; van Hooft et al., 2000), in single neurons or studied groups of neurons defined by a shared property: synaptic target (Arlotta et al., 2005) or common expression of a transgene (Sugino et al., 2006). While very powerful, those studies did not address the question of the ability of global expression profiling at the single cell level to prospectively group and classify neuronal cell types. In addition, the use of a single property to define group identity, by its very nature, will give an averaged expression profile of that population with no indication of underlying cellular heterogeneity. Given the difficulties encountered in classifying cortical interneurons using combinations of morphology, electrical properties and marker protein or gene expression (Ascoli et al., 2008), we wished to investigate whether an unbiased, global assay of gene expression could identify groups of neurons independent of any other cellular property.





We generated global expression data from a relatively small sample of 18 morphologically and electrophysiologically characterized neurons and used those data to group the neurons by hierarchical cluster analysis. It is noteworthy that the data were not filtered to first derive a subset of transcripts that displayed variation in expression among the cells that would then be used to group the cells. In this case, the clustering based on gene expression was carried out independent of any phenotypic information. We found that this approach can group and classify cells that would be considered quite different in their cellular properties, including neurotransmitter type, morphology and developmental origins. However, we did find errors in assigning cells to clusters within this small and complex population of neurons. Secondly, we found that this approach also enabled the identification of novel cell type-enriched transcripts. It is not clear currently if single cell profiling is capable of distinguishing related, but distinct, classes of neurons, such as pyramidal cells from different cortical layers.

However, there are inherent limitations to this approach at present. Single cell expression profiling presents specific technical and biological challenges. Technically, current expression profiling technologies require amplification of the population of single cell mRNA, a process that can and does introduce noise, resulting in an altered transcript distribution. Importantly, this can lead to low abundance transcripts being called as absent in a random manner, due to inefficiencies in the early steps of reverse transcription and

amplification (Subkhankulova et al., 2008). In the data reported here, for example, we did not identify some genes that are known to be pyramidal or interneuron-specific, such as *Emx1* or *GAD65* (Katarova et al., 2000; Chan et al., 2001). From our previous work in neural stem cells, it is likely that this is due to a combination of technical issues, including inefficiencies in the global amplification method, the sensitivity of the array technique and the use of the pooled cells as a common reference. This last aspect of the experimental design is useful as it ensures that all genes expressed in the population are represented in the reference sample. However, the drawback of this approach is that, depending on the abundance of cell-specific transcripts in the pooled reference, some cell-enriched transcripts may not appear enriched when compared directly to the reference, rather they will appear depleted in the cells that do not express those genes.

A second potentially confounding biological problem is the noise inherent in gene expression in any cell and the consequent variation in absolute transcript numbers among phenotypically identical cells. We have previously found that there are significant differences in transcript abundance between phenotypically identical neural stem cells and that technical noise did not account for those differences (Subkhankulova et al., 2008). Therefore, it is likely that it is necessary to sample the population deeply to ensure sufficient representation of each cell type and thereby compensate for the

intrinsic variability of single cell transcript levels. Alternatively, reducing expression levels from continuous measurements to discrete bins or, in the extreme, to simple binary presence or absence calls, would potentially simplify the classification of cells based on expression of thousands of transcripts.

CAN GLOBAL EXPRESSION PROFILES BE USED TO PREDICT BIOLOGICAL PROPERTIES OF NEURONS?

One long-term goal of single cell profiling is to use the combinatorial expression data to predict the biological properties of the neuron, essentially by constructing cellular models from single cell transcriptome data. Previous studies, based on single cell PCR data of expression of sets of ion channels, indicated that it is possible to correlate electrical properties with the expression of particular combinations of channels (Toledo-Rodriguez et al., 2004). However, scaling this approach from tens to thousands of transcripts presents a considerable challenge, particularly if the relative levels of those transcripts, and not simply presence or absence of a given transcript, are included in the cellular modeling. Analyses of gene expression in single neurons have found that the levels of mRNAs for specific genes can vary considerably among cells of the same type (Schulz et al., 2006, 2007; Tobin et al., 2009). It is likely that models of neuronal phenotype and electrical properties will require data on the levels of the corresponding proteins and also the possible protein-protein interactions and their functional significance. Current technologies for protein analysis have not achieved single cell resolution, but large-scale proteomic analyses of neurons have begun to provide data that will be useful in this context (Emes et al., 2008).

REFERENCES

- Arlotta, P., Molyneaux, B. J., Chen, J., Inoue, J., Kominami, R., and Macklis, J. D. (2005). Neuronal subtype-specific genes that control corticospinal motor neuron development *in vivo*. *Neuron* 45, 207–221.
- Ascoli, G. A., Alonso-Nanclares, L., Anderson, S. A., Barrionuevo, G., Benavides-Piccione, R., Burkhalter, A., Buzsáki, G., Cauli, B., Defelipe, J., Fairen, A., Feldmeyer, D., Fishell, G., Fregnac, Y., Freund, T. F., Gardner, D., Gardner, E. P., Goldberg, J. H., Helmstaedter, M., Hestrin, S., Karube, F., Kisvárdy, Z. F., Lambolez, B., Lewis, D. A., Marin, O., Markram, H., Muñoz, A., Packer, A., Petersen, C. C., Rockland, K. S., Rossier, J., Rudy, B., Somogyi, P., Staiger, J. F., Tamas, G., Thomson, A. M., Toledo-Rodriguez, M., Wang, Y., West, D. C., and Yuste, R. (2008). Petilla terminology: nomenclature of features of GABAergic interneurons of the cerebral cortex. *Nat. Rev. Neurosci.* 9, 557–568.
- Bota, M., and Swanson, L. W. (2007). The neuron classification problem. *Brain Res. Rev.* 56, 79–88.
- Brady, G., Billia, F., Knox, J., Hoang, T., Kirsch, I. R., Voura, E. B., Hawley, R. G., Cumming, R., Buchwald, M., and Siminovitch, K. (1995). Analysis of gene expression in a complex differentiation hierarchy by global amplification of cDNA from single cells. *Curr. Biol.* 5, 909–922.
- Brady, G., and Iscove, N. N. (1993). Construction of cDNA libraries from single cells. *Meth. Enzymol.* 225, 611–623.
- Cauli, B., Audinat, E., Lambolez, B., Angulo, M. C., Ropert, N., Tsuzuki, K., Hestrin, S., and Rossier, J. (1997). Molecular and physiological diversity of cortical nonpyramidal cells. *J. Neurosci.* 17, 3894–3906.
- Cauli, B., Porter, J. T., Tsuzuki, K., Lambolez, B., Rossier, J., Quenet, B., and Audinat, E. (2000). Classification of fusiform neocortical interneurons based on unsupervised clustering. *Proc. Natl. Acad. Sci. U.S.A.* 97, 6144–6149.
- Chan, C. H., Godinho, L. N., Thomaidou, D., Tan, S. S., Gulisano, M., and Parnavelas, J. G. (2001). Emx1 is a marker for pyramidal neurons of the cerebral cortex. *Cereb. Cortex* 11, 1191–1198.
- Chomczynski, P., and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium triocyanate-phenol-chloroform extraction. *Anal. Biochem.* 162, 156–159.
- Connors, B. W., and Gutnick, M. J. (1990). Intrinsic firing patterns of diverse neocortical neurons. *Trends Neurosci.* 13, 99–104.
- Connors, B. W., Gutnick, M. J., and Prince, D. A. (1982). Electrophysiological properties of neocortical neurons *in vitro*. *J. Neurophysiol.* 48, 1302–1320.
- Emes, R. D., Pocklington, A. J., Anderson, C. N., Bayes, A., Collins, M. O., Vickers, C. A., Croning, M. D., Malik, B. R., Choudhary, J. S., Armstrong, J. D., and Grant, S. G. (2008). Evolutionary expansion and anatomical specialization of synapse proteome complexity. *Nat. Neurosci.* 11, 799–806.
- Helmstaedter, M., Sakmann, B., and Feldmeyer, D. (2009a). L2/3 interneuron groups defined by multiparameter analysis of axonal projection, dendritic geometry, and electrical excitability. *Cereb. Cortex* 19, 951–962.
- Helmstaedter, M., Sakmann, B., and Feldmeyer, D. (2009b). The relation between dendritic geometry, electrical excitability, and axonal projections of L2/3 interneurons in rat barrel cortex. *Cereb. Cortex* 19, 938–950.
- 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-picogram quantities of mRNA. *Nat. Biotechnol.* 20, 940–943.
- Karagiannis, A., Gallopin, T., David, C., Battaglia, D., Geoffroy, H., Rossier, J., Hillman, E. M., Staiger, J. F., and Cauli, B. (2009). Classification of NPY-expressing neocortical interneurons. *J. Neurosci.* 29, 3642–3659.
- Katarova, Z., Sekerkova, G., Prodan, S., Mugnaini, E., and Szabo, G. (2000). Domain-restricted expression of two glutamic acid decarboxylase genes in midgestation mouse embryos. *J. Comp. Neurol.* 424, 607–627.
- Kawaguchi, Y., and Kondo, S. (2002). Parvalbumin, somatostatin and cholecystokinin as chemical markers for specific GABAergic interneuron types in the rat frontal cortex. *J. Neurocytol.* 31, 277–287.
- Lein, E. S., Hawrylycz, M. J., Ao, N., Ayres, M., Bensinger, A., Bernard, A., Boe, A. F., Boguski, M. S., Brockway, K. S., Byrnes, E. J., Chen, L., Chen, L., Chen, T. M., Chin, M. C., Chong, J., Crook, B. E., Czaplinska, A., Dang, C. N., Datta, S., Dee, N. R., Desaki, A. L., Desta, T., Diep, E., Dolbeare, T. A., Donelan, M. J., Dong, H. W., Dougherty, J. G., Duncan, B. J., Ebbert, A. J., Eichele, G., Estlin, L. K., Faber, C., Facer, B. A., Fields, R., Fischer, S. R., Floss, T. P., Frensley, C., Gates, S. N., Glattfelder, K. J., Halverson, K. R., Hart, M. R., Hohmann, J. G., Howell,

TOWARDS A CATALOGUE OF CORTICAL CELL TYPES

Previous global expression studies of cortical neurons have produced a taxonomy of forebrain cell types (Sugino et al., 2006) and identified many novel cell-specific transcripts in pyramidal cells (Arlotta et al., 2005). Those studies have demonstrated that there are many examples of transcripts that can be used singly or in combination as discriminators of cortical cell types. The data presented here indicate that the global expression profiling approach has the potential to be extended to the single cell level, independent of any cellular property of cortical neurons, to sample the population of neurons in a cortical region in order to assess cellular diversity at the genetic level, group those cells and, ultimately, classify them by gene expression. As such, this approach provides a bottom-up approach that complements the top-down approach of identifying cell-specific expression patterns based on some group-defining property of a set of neurons. Together, these complementary approaches may enable the generation of a catalogue of cortical cell types, based on gene expression, which can be one element of a holistic classification of cortical neurons.

ACKNOWLEDGMENTS

Tatiana Subkhankulova and Kojiro Yano were supported by the EU FP6 STREP, Neuro-Functional Genomics.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at <http://www.frontiersin.org/molecularneuroscience/paper/10.3389/fmol.2010.00010/>

- M. P., Jeung, D. P., Johnson, R. A., Karr, P. T., Kaval, R., Kidney, J. M., Knapik, R. H., Kuan, C. L., Lake, J. H., Laramée, A. R., Larsen, K. D., Lau, C., Lemon, T. A., Liang, A. J., Liu, Y., Luong, L. T., Michaels, J., Morgan, J. J., Morgan, R. J., Mortrud, M. T., Mosqueda, N. F., Ng, L. L., Ng, R., Orta, G. J., Overly, C. C., Pak, T. H., Parry, S. E., Pathak, S. D., Pearson, O. C., Puchalski, R. B., Riley, Z. L., Rockett, H. R., Rowland, S. A., Royall, J. J., Ruiz, M. J., Sarno, N. R., Schaffnit, K., Shapovalova, N. V., Sivisay, T., Slaughterbeck, C. R., Smith, S. C., Smith, K. A., Smith, B. I., Sodt, A. J., Stewart, N. N., Stumpf, K. R., Sunkin, S. M., Sutram, M., Tam, A., Teemer, C. D., Thaller, C., Thompson, C. L., Varnam, L. R., Visel, A., Whitlock, R. M., Wohnoutka, P. E., Wolkey, C. K., Wong, V. Y., Wood, M., Yaylaoglu, M. B., Young, R. C., Youngstrom, B. L., Yuan, X. F., Zhang, B., Zwingman, T. A., and Jones, A. R. (2007). Genome-wide atlas of gene expression in the adult mouse brain. *Nature* 445, 168–176.
- Markram, H., Toledo-Rodriguez, M., Wang, Y., Gupta, A., Silberberg, G., and Wu, C. (2004). Interneurons of the neocortical inhibitory system. *Nat. Rev. Neurosci.* 5, 793–807.
- Masland, R. H., and Raviola, E. (2000). Confronting complexity: strategies for understanding the microcircuitry of the retina. *Annu. Rev. Neurosci.* 23, 249–284.
- Monyer, H., and Jonas, P. (1995). “Polymerase chain reaction analysis of ion channel expression in single neurons of brain slices” in *Single Channel Recording*, eds B. Sakmann and E. Neher (New York: Plenum Press), 357–373.
- Nelson, S. B., Sugino, K., and Hempel, C. M. (2006). The problem of neuronal cell types: a physiological genomics approach. *Trends Neurosci.* 29, 339–345.
- Rossner, M. J., Hirrlinger, J., Wichert, S. P., Boehm, C., Newrzella, D., Hiemisch, H., Eisenhardt, G., Stuenkel, C., von Ahsen, O., and Nave, K. A. (2006). Global transcriptome analysis of genetically identified neurons in the adult cortex. *J. Neurosci.* 26, 9956–9966.
- Schulz, D. J., Goaillard, J. M., and Marder, E. (2006). Variable channel expression in identified single and electrically coupled neurons in different animals. *Nat. Neurosci.* 9, 356–362.
- Schulz, D. J., Goaillard, J. M., and Marder, E. E. (2007). Quantitative expression profiling of identified neurons reveals cell-specific constraints on highly variable levels of gene expression. *Proc. Natl. Acad. Sci. U.S.A.* 104, 13187–13191.
- Subkhankulova, T., Gilchrist, M. J., and Livesey, F. J. (2008). Modelling and measuring single cell RNA expression levels find considerable transcriptional differences among phenotypically identical cells. *BMC Genomics* 9, 268.
- Subkhankulova, T., and Livesey, F. J. (2006). Comparative evaluation of linear and exponential amplification techniques for expression profiling at the single cell level. *Genome Biol.* 7, R18.
- Sugino, K., Hempel, C. M., Miller, M. N., Hattox, A. M., Shapiro, P., Wu, C., Huang, Z. J., and Nelson, S. B. (2006). Molecular taxonomy of major neuronal classes in the adult mouse forebrain. *Nat. Neurosci.* 9, 99–107.
- Tateno, T., Harsch, A., and Robinson, H. P. (2004). Threshold firing frequency-current relationships of neurons in rat somatosensory cortex: type 1 and type 2 dynamics. *J. Neurophysiol.* 92, 2283–2294.
- Tateno, T., Jimbo, Y., and Robinson, H. P. (2005). Spatio-temporal cholinergic modulation in cultured networks of rat cortical neurons: spontaneous activity. *Neuroscience* 134, 425–437.
- Thomson, A. M., and Lamy, C. (2007). Functional maps of neocortical local circuitry. *Front. Neurosci.* 1, 19–42.
- Tobin, A. E., Cruz-Bermudez, N. D., Marder, E., and Schulz, D. J. (2009). Correlations in ion channel mRNA in rhythmically active neurons. *PLoS ONE* 4, e6742. doi:10.1371/journal.pone.0006742.
- Toledo-Rodriguez, M., Blumenfeld, B., Wu, C., Luo, J., Attali, B., Goodman, P., and Markram, H. (2004). Correlation maps allow neuronal electrical properties to be predicted from single-cell gene expression profiles in rat neocortex. *Cereb. Cortex* 14, 1310–1327.
- van Hooft, J. A., Giuffrida, R., Blatow, M., and Monyer, H. (2000). Differential expression of group I metabotropic glutamate receptors in functionally distinct hippocampal interneurons. *J. Neurosci.* 20, 3544–3551.
- Yano, K., Subkhankulova, T., Livesey, F. J., and Robinson, H. P. (2006). Electrophysiological and gene expression profiling of neuronal cell types in mammalian neocortex. *J. Physiol.* 575, 361–365.

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

Received: 11 November 2009; paper pending published: 04 December 2009; accepted: 20 March 2010; published online: 13 April 2010.

Citation: Subkhankulova T, Yano K, Robinson HPC and Livesey FJ (2010) Grouping and classifying electrophysiologically-defined classes of neocortical neurons by single cell, whole-genome expression profiling. *Front. Mol. Neurosci.* 3:10. doi: 10.3389/fnmol.2010.00010
Copyright © 2010 Subkhankulova, Yano, Robinson and Livesey. This is an open-access article subject to an exclusive license agreement between the authors and the Frontiers Research Foundation, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited.