Cell surface protein mRNAs show differential transcription in pyramidal and fast-spiking cells as revealed by single-cell sequencing

Lilla Ravasz[1,2], Katalin Adrienna Kékesi[1,2,3], Dániel Mittli[2], Mihail Ivilinov Todorov[2], Zsolt Borhegyi[1,2], Mária Ercsey-Ravasz[4,5], Botond Tyukodi[4,6], Jinhui Wang[7], Tamás Bártfai[8,#], James Eberwine[7,#,*], Gábor Juhász[1,2,9,#,*]

[1]ELTE NAP Neuroimmunology Research Group, Department of Biochemistry, Institute of Biology, ELTE Eötvös Loránd University, Budapest H-1117 Hungary

[2]Laboratory of Proteomics, Institute of Biology, ELTE Eötvös Loránd University, Budapest, H-1117 Hungary

[3]Department of Physiology and Neurobiology, Institute of Biology, ELTE Eötvös Loránd University, Budapest, H-1117 Hungary

[4]Faculty of Physics, Babeș-Bolyai University, Cluj-Napoca, Romania

[5]Transylvanian Institute of Neuroscience, Cluj-Napoca, Romania

[6]Martin Fisher School of Physics, Brandeis University, Waltham, MA, 02451, USA

[7]Department of Systems Pharmacology and Translational Therapeutics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, 19104, USA

[8]Department of Biochemistry and Biophysics, Stockholm University, Sweden

[9]CRU Hungary Ltd. Göd, H-2131 Hungary

[#]Authors contributed equally and are co-senior authors.

*Correspondence should be addressed: Gábor Juhász, email: gjuhasz100@gmail.com ; James Eberwine, email: eberwine@upenn.edu

Short title: Single-cell transcriptomics of mouse PFC neurons

Keywords: prefrontal cortex, interneuron, pyramidal cell, single-cell transcriptomics, drug target, schizophrenia, autism, taxonomy

Abstract

The prefrontal cortex (PFC) plays a key role in higher order cognitive functions and psychiatric disorders like autism, schizophrenia and depression. In the PFC, the two major classes of neurons are the glutamatergic pyramidal (Pyr) cells and the GABAergic interneurons like fast-spiking (FS) cells. Despite extensive electrophysiological, morphological and pharmacological studies of the PFC, the therapeutically utilized drug targets are restricted to dopaminergic, glutamatergic and GABAergic receptors. To expand the pharmacological possibilities as well as to better understand the cellular and network effects of clinically used drugs, it is important to identify cell type-selective, druggable cell surface proteins and to link developed drug candidates to Pyr or FS cell targets. To identify the mRNAs of such cell-specific/enriched proteins, we performed ultra-deep single-cell mRNA sequencing (19,685 transcripts in total) on electrophysiologically characterized intact PFC neurons harvested from acute brain slices of mice. Several selectively expressed transcripts were identified with some of the genes that have already been associated with cellular mechanisms of psychiatric diseases, which we can now assign to Pyr (e.g. *Kcnn2*, *Gria3*) or FS (e.g. *Kcnk2*, *Kcnmb1*) cells. The earlier classification of PFC neurons was also confirmed at mRNA level and additional markers have been provided.

Introduction

Cortical neurons have extreme molecular and functional heterogeneity. Pathological alterations in their cooperative function could initiate psychiatric diseases particularly when the cortical neuronal networks are mistuned by subcortical influences. The classical example is the unbalanced excitatory/inhibitory synaptic communication of pyramidal (Pyr) and fast-spiking (FS) cells in schizophrenia (Ferguson and Gao 2018) when the decreased feedback inhibition results in decreased gamma frequency coupling necessary for cognitive functions (Buzsáki and Wang 2012). Therefore, the selective fine tuning of cortical Pyr and FS cells in schizophrenia is a potential therapeutic target. Cell-selective targeting is highly developed in cancer treatment using cell surfaceome targets (surfaceome database: http://wlab.ethz.ch/surfaceome/), but it is in infancy in the field of brain diseases. However, there are evidences about the causality between the transcriptomic and electrophysiological profile of a neuron (Cadwell et al. 2016; Fuzik et al. 2016; Bomkamp et al. 2019) supporting the idea that development of cell type-selective fine tuning of cellular electrical activity could be promoted by single-cell transcriptomics.

In the past decades enormous efforts were made to distinguish cortical neuronal phenotypes according to morphological and immunohistochemical criteria (Kawaguchi and Kubota 1997; Hill 2001). The development of single-cell sequencing of large number of neurons automatically separated by cell dispersion methods established several novel neuronal subclasses on the basis of their transcriptomic profile (Zeisel et al. 2015; Tasic et al. 2016). However, in some cases the taxonomic results are questionable since many of the clusters are difficult to link to cortical functions and are mostly unrecognizable *in vivo* in the functioning brain (Zeng and Sanes 2017). On the other hand, a large body of evidence is available about the cellular mechanisms of Pyr and FS cell interactions in CNS diseases so the selective tuning of them via targeted drug therapy can be an appropriate approach. Thus, we performed single-

cell sequencing for answering the question whether the Pyr and FS cells have or have not transcript differences significant enough for seriously considering their selective pharmacological targeting.

The two electrophysiologically well-recognizable cell types in the prefrontal cortex (PFC) the Pyr and the FS cells have an immense literature describing their role in physiological and pathological processes (Zhong and Yan 2011; Hu et al. 2014; Kann 2016). In the present study we separated them using patch clamp recording before cell harvesting and used ultra-deep sequencing to establish mRNA transcript differences coding cell surface receptors and ion channels which are functionally relevant members of the neuronal surfaceome. We focused our study on the PFC since it plays a crucial role in higher order cognitive functions and psychiatric diseases (Goldman-Rakic 1995), such as schizophrenia and autism spectrum disorders (Foss-Feig et al. 2017). In addition, it is an important target for drug development (Masana et al. 2013). Functional evidences suggest cell type-selective tuning of PFC neurons for advancing pharmacotherapy, but the classical anatomical markers (VGLUT1, GAD67) of Pyr and FS cells do not allow cell-selective targeting because of subcellular localization and overall distribution in the organism (Esclapez et al. 1994; Vigneault et al. 2015). Thus, there is a need for identifying novel cell-specific proteins belonging to the neuronal surfaceome. The mechanisms determining the *in vivo* selectivity of drugs are influenced by many factors such as pharmacokinetic properties of drug-target binding or tissue distribution of target molecules (Vlot et al. 2017). Furthermore, the effective concentration and the cell type-specificity of target proteins are also relevant aspects, in the exploration of which single-cell mRNA sequencing may be a useful technique (Bartfai et al. 2012).

Here, we report on ion channel and receptor (e.g. GPCR) mRNA differences between Pyr and FS cells in the mouse PFC remarkable enough to propose the coded proteins for selective targeting after the protein level validation. It is the first study applying the patch-seq technology

combined with ultra-deep sequencing for uncovering cell surfaceome differences between the two investigated neuronal types. We are aware of the limitations of patch-seq technology and single-cell sequencing of the harvested neurons, but it seems to be a substantially robust method for *ex vivo* characterization of neurons before harvesting, and the applied ultra-deep sequencing allows to detect cytoplasmic mRNAs of very low copy number. Consequently, we should take into consideration the difficulties of validation of our transcriptomic results at protein level, but we propose our finding for further investigations aiming the selective fine tuning of Pyr and FS cells.

Materials and Methods

Animals

Cell harvesting was made in acute slices, obtained from male C57BL/6N mice (n=73) of age between 27 and 40 days (Innovo Kft., Isaszeg, Hungary). Animals were housed in an animal room maintained in a specific pathogen-free animal facility having HEPA filtered air and they were tested serologically regularly. All procedures of animal care and minimalizing suffer and pain for the animals were done under the local ethical rules of Eötvös Loránd University which is in accordance with the EU Ethical Rules of Using Animals for Research Purposes (2010/63/EU revising Directive 86/609/EEC) and the Hungarian Act of Animal Care and Experimentation (1998, XXVIII).

Brain slice preparation and patch clamp electrophysiology

All reagents were purchased from Sigma-Aldrich (St. Louis, MO, USA) and all solutions were made from autoclaved (120ºC, 20 min) PCR quality, RNase/DNase free water (Merck Millipore, Billerica, MA, USA). During the experiments three different artificial cerebrospinal fluid (ACSF) were used: (i) 85 mM NaCl, 2.5 mM KCl, 2.0 mM $MgCl_2$, 1.125 mM NaH_2PO_4,

25 mM $NaHCO_3$, 25 mM glucose, 60 mM sucrose, 1 mM $CaCl_2$, at pH 7.4 for preparing the slices, (ii) 125 mM NaCl, 2.5 mM KCl, 2.0 mM $MgCl_2$, 1.125 mM NaH_2PO_4, 25 mM $NaHCO_3$, 25 mM glucose, 1 mM $CaCl_2$, at pH 7.4 for incubating the slices, and (iii) 125 mM NaCl, 2.5 mM KCl, 1.0 mM $MgCl_2$, 1.125 mM NaH_2PO_4, 25 mM $NaHCO_3$, 25 mM glucose, 2 mM $CaCl_2$, at pH 7.4 during recording.

The patch pipettes were filled with RNase-free intracellular solution (130 mM potassium gluconate, 20 mM KCl, 10 mM HEPES, 0.16 mM EGTA, 4.0 mM ATP, 2.0 mM $MgCl_2$, 0.3 mM GTP, at pH 7.4). A silver-silver chloride electrode was inserted into the pipettes which was freshly chlorided before each session. Pipettes were pulled right before use to avoid contaminations.

We used standard brain slice preparation technology. Briefly, mice were anesthetized in 2% isoflurane, and quickly decapitated. Brains were removed and placed into ice cold ACSF. Slices were cut by vibratome (Leica VT1000 S, Leica Biosystems, Wetzlar, Germany) under ice cold cutting ACSF supplied with carbogen (5% CO_2 in 95% O_2), then were incubated in storage ACSF at room temperature in a chamber and permanently supplied by carbogen for at least 1 hour before recording. Three hundred μm coronal slices were cut from the brains so that we were able to collect cells from different layers from the PFC.

Patch pipettes were pulled to 4-10 Mohm resistance by a David Kopf 720 vertical pipette puller (David Kopf Instruments, Tujunga, CA, USA) using a patch pipette puller gear.

Whole-cell patch clamp measurements were performed at 34°C using a standard setup built on a Leica electrophysiology microscope DM6000 FS (Leica Microsystems, Wetzlar, Germany) equipped with DIC optics and a 40x magnification water immersion objective. The microscope had an optical magnification controller from 0.35x to 1.25x for amplification of the image. A QImaging Rolera-XR high resolution camera (QImaging, Surrey, BC, Canada) was used for visualizing the cells. We used Sutter MP-285 micromanipulator (Sutter Instrument,

Novato, CA, USA) for 3D movement of microelectrodes. All equipments were placed on a Gibraltar platform and X-Y stage (Burleigh Instruments, New York, NY, USA). The patch clamping protocol started with targeting the selected cell and achieving Gohm seal resistance. Then we used a step-gradient depolarization protocol in bridge mode at sampling rate 10 kHz which started at -0.1 nA using 0.05 nA step size and 0.5 sec step duration. We continued recording up to the highest firing rate of the neurons and stored the data of all steps in CED 1401 file format. Electrophysiological signals were amplified by an AxoClamp 2B amplifier (Axon Instruments, Foster City, CA, USA). Analog data were digitized by a CED 1401 MK II (Cambridge Electronic Design, Cambridge, UK) data capture device using Signal 4.11 software for data capture and processing. All data were stored in a data bank in a standard form. We stored pictures from before patch clamping, during recording and after harvesting of the cells and also we stored a low magnification picture showing the position of the recorded cell in the PFC.

For cell identification we used anatomical and physiological markers. Pyr cells were identified based on their visible apical dendrite, the larger, triangular cell body and the regular firing of action potentials. The identification of FS cells in slices was done by their smaller, spherical cell body and the high frequency firing of narrow action potentials. It is important to note that we harvested at most 2 cells from each type from one animal to preserve the individual genetic variation. We harvested altogether 75 Pyr and 29 FS cells and after quality control we sequenced 59 Pyr and 25 FS cells.

Amplification and sequencing

Following electrophysiological recording, the cytoplasmic mRNA of the recorded cells was harvested, reverse transcribed into cDNA, amplified through two rounds of aRNA amplification (Van Gelder et al. 1990; Eberwine et al. 1992), and performed into barcoded

libraries for sequencing. Each cell was individually amplified through three rounds of aRNA amplification. A 1:4,000,000 dilution of ERCC RNA Spike-In (Life Technologies, Carlsbad, CA, USA) was added to each sample to control for technical variation between samples. For the first round of amplification, a synthesized oligo(dT)-T7 primer that contained a poly-T and phage T7 RNA polymerase promoter sequence was annealed to the poly-A tail of the mRNA. Double-stranded cDNA was synthesized using SuperScript III reverse Transcriptase (Invitrogen, Carlsbad, CA, USA), and DNA polymerase I (Invitrogen, Carlsbad, CA, USA) following the manufacture's protocols. The double-stranded cDNA served as the template for aRNA transcription using the MEGAscript T7 kit (Invitrogen, Carlsbad, CA, USA). For second and third round of amplification, aRNA was converted into cDNA using random hexanucleotide primers for first strand and oligo(dT)-T7 primer to initiate second strand. The amplified aRNA was purified with AGENCOURT RNACLEAN X beads (Beckman Coulter, Brea, CA, USA). The quality and quantity of the aRNA were assessed using Bioanalyzer RNA Picochip and Nanochip (Agilent, Santa Clara, CA, USA).

Libraries were constructed from the aRNA using the TruSeq Stranded mRNA Library Prep Kit (Illumina, San Diego, CA, USA) without the initial fragmentation incubation step. The resulting libraries were quantified using Bioanalyzer DNA1000 chip (Agilent, Santa Clara, CA, USA). Libraries were sequenced either on HiSeq2500 to produce 100-base paired-end reads, or NextSeq500 to produce 75-base paired-end reads. The GEO accession number for the data is GSE135060.

Hierarchical clustering based on electrophysiological parameters

To measure parameters of electrophysiological signals we used a script of CED written for evaluation of patch clamp data (Intracellular spike analysis, Cambridge Electronic Design,

Cambridge, UK, http://ced.co.uk/downloads/scriptsiganal). We measured 22 parameters for each cell at each depolarization step. We show these in details in Figure S1.

Anatomical evaluation of the cells was made by scoring different parameters based on low and high magnification pictures made before and after the recordings using Paxinos mouse brain atlas as reference (Paxinos and Franklin 2006). Layer position score was equal with the layer number. Dorso-ventral position score was 0 for cingulate, 1 for prelimbic and 2 for infralimbic location. Antero-posterior position score was made so that the middle range of medial prefrontal area (mainly prelimbic area) was scored as 1 and if the cell was more anterior than scored 0 and 2 if it was more posterior. A cell score was also applied based on its visible details: 0 if only a round-shaped soma could be seen, 1 if apical dendrite could be seen also, 2 if the cell was multipolar, 3 for dendrite with uncertain morphology and 4 for bipolar cells. The position of the anchoring rod (keeping the slice fixed in position) was also scored if it was so close to the recording site that it might have interfered with the physiology by pressing cells and cell processes. Finally, a score was also applied if the cytoplasm removal was full and the nucleus could be seen clearly on the tip of the pipette when it was pulled out from the tissue.

Physiological data were organized in frames meaning that neuronal response to each depolarization step was stored in a separate frame. The first frame was defined as the frame in which the cell fired at least three action potentials throughout the entire depolarizing current pulse. The last frame was the depolarization step which induced the highest firing rate of the neuron. For clustering based on the electrophysiological measurements we used data from the first frame because first frame and last frame data resulted very similar clusters. The following 22 features were taken into account: 1st, 2nd, middle spike: time, threshold, amplitude, 10-90% rise time, half width, after hyperpolarization value (AHP); frequency: 1st-2nd spike, 2nd-last spike average; last spike time; 2nd-last spike average amplitude (see Figure S1).

Each of these 22 features was normalized such that they have zero average and unit variance (Choromanska et al. 2015). Clustering being a complex optimization problem, using too many parameters could result in difficulties. For this reason a principal component (PC) analysis was applied to these features (Tipping and Bishop 1999) and the first 3 PCs were considered for the clustering.

For these three PCs a hierarchical agglomerative clustering was carried out (Müllner 2011). Between the data points an Euclidean distance was considered in the 3D space (the 3 PCs being the coordinates) and the Ward linkage criterion was applied (Bar-Joseph et al. 2001). The basic idea of agglomerative clustering is that it starts form all points being in different clusters and at each step it merges two different clusters such that an objective function is optimized. The method we used is called Ward's minimum variance criterion and it minimizes the total within-cluster variance. This agglomerative process can be illustrated through a dendrogram (Figure 1D) and one can choose to study the clusters at any level (choosing any number of clusters).

Contrast analysis of hierarchical clustering

There are many statistical methods for contrast analysis (ANOVA tests, F-tests, etc.) and, however, the number of cells in our case is so small that performing complex statistical methods could provide results that are hard to interpret and not necessarily meaningful or correct. For this reason, we decided to use a very simple approach by measuring different contrast variables calculated from the experimental values.

For each gene G and cell C we have an $x(G,C)$ measure. These show a very large variance, there are also many 0 values. Some former experiments indicate that the distribution is lognormal. For our analysis the best would be to have values close to a normal distribution. For this reason, we used the measures $y(G,C)=\ln(1+x(G,C))$. (The 1 inside the logarithm is needed

to avoid obtaining –infinity in case of x=0.) The y variables will span over a much smaller interval having a distribution closer to a Gaussian.

Next we identified the clusters we wanted to compare, we cut the dendrogram provided by the hierarchical clustering at a level that provided 4 clusters. The average of $y(G,C)$ variables inside a cluster j (j=1,…,4) will be denoted as $\mu_j(G) = \frac{1}{n_j}\sum_{C\in j} y(G,C)$, where n_j denotes the number of cells in cluster j.

A contrast variable can be defined as $V(G) = \sum_j c_j \mu_j(G)$, where $\sum_j c_j = 0$. We have defined three different contrast variables that are orthogonal to each other, each of them describing a separate comparison (see Table S5). V_1 compares the two large clusters of FS and Pyr cells: clusters (1 and 2) vs. (3 and 4). V_2 compares the two clusters inside the FS cells (cluster 1 vs. 2) and V_3 the two clusters inside the Pyr cells (clusters 3 vs. 4).

Normalization of raw single-cell sequencing data

Because of the extreme low sample volumes and difficulties to collect them from the cells of the living tissue, and also of the expectable differences in the actual state of the individual cells, the raw data of single-cell sequencing exhibits large fluctuations in the number of genes transcribed and also in their copy numbers. Therefore, the data normalization is a critical issue in differential expression analysis of a heterogeneous neuron population. First, the sequencing data were normalized using the DESeq R program package originally developed for bulk sequencing data. We expect that if a dataset is properly normalized, the sum of normalized copy numbers should correlate with the number of observed genes. Verifying this correlation is a simple test of the reliability of the normalization. Using plain DESeq data resulted in a poor correlation (see Figure S2A) indicating an insufficient normalization procedure. Transcript copy number data could be normalized on the housekeeping gene products which are

indispensable for the homeostasis of the cell and expected to be transcribed at a constant level (low dispersion). However, normalizing the raw sequencing data for 46 housekeeping genes (29 genes that are generally accepted as housekeeping reference genes and 17 genes of the citric acid cycle, Table S6) provided also poor correlation (see Figure S2B). We found that even these genes were not transcribed in all of the cells and none of the cells expressed all of them. The incomplete transcript pattern of housekeeping genes could be the result of technical uncertainties and, however, it has little if any chance because when they are transcribed, the transcript copy number is high above the technical limit. As it is known, gene transcription is performed in bursts and several of the genes could be in OFF state in the time point of harvesting. In fact, we observed that several genes are transcribed in large copy numbers but only in small number of cells, we assume that mainly gene OFF state is responsible for the zero transcription values. In turn, the reference genes for normalization should be transcribed in most of the cells and their transcription levels should be high and stable. Using these criteria, we searched for reference genes in a non-biased way (i.e., without knowing their house-keeping or any function).

(i) First, we sorted out 1,000 genes that were present most frequently in all cells.

(ii) To find the most constantly expressed genes, we carried out a "pre-normalization" on the sequencing data using these 1,000 genes as reference. For each cell, a scaling-factor was determined as the factor resulting in the lowest root-mean-square deviation (RMSD) related to the average values on all expressed genes out of the 1,000 selected genes in the cell. The cells were "pre-normalized" with these factors.

(iii) In the next step, on the "pre-normalized" data to the most frequent 1,000 genes, we sorted out the 500 most stable ones showing the lowest standard deviation.

(iv) Because the main goal of this single-cell sequencing study is to determine the transcription differences between FS and Pyr cells, the reference genes were further filtered for those

showing less than tenfold difference between FS and Pyr cells (Figure S3). The resulting 409 genes were taken as reference genes with frequent and stable expression both in FS and Pyr cells and represent an "equi-phenotype" of the sequenced neuronal cell transcriptomes. The "pre-normalized" data was not used any further.

(v) We have to note, that for most of the genes, single-cell sequencing data showed zero copy numbers in majority of the cells. The great number of zero transcription dramatically shifts (lowers) the average transcription values. The use of median for single-cell data could be misleading also. Therefore, we calculated the average copy numbers only for the non-zero values. The rationale behind our calculation is that the non-zero values average represents the expectable average transcription level when the gene is in ON state. Together with the frequency of the gene ON state, our estimation provided a more real description of single-cell transcription than the overall average or median that include zero values.

(vi) The selected 409 genes of the "equi-phenotype" were used to normalize the raw data of the cells for expression difference analysis. The normalization factor for a cell was the multiplication-factor that provided the lowest RMSD on the 409 reference genes compared to the averages of the non-zero values. The correlation between the number of recovered genes and the sum of copy numbers on the normalized dataset is presented in Figure S2C and indicated that our normalization was more reliable than other normalization processes.

Investigation of the reliability of the comparison of FS and Pyr cells – the difference reliability score

Our aim was to compare the transcription levels for the sequenced genes between FS and Pyr cells. Because of the relatively low number of sequenced cells and because most of the genes were only present in a fraction of the cells resulting lot of zero values, it was difficult to find an appropriate statistic to verify the reliability of comparison and eliminate false discovery

rate. Outlier data, defined as higher than the median of the log-transformed data (for non-zero values) with more than twice of the standard deviation, were discarded. To validate the data quality for differential expression analysis, we randomly divided the FS cells into two groups of equal number of cells and added to these groups the randomly selected two halves of Pyr cells. Thus, we randomly selected cells to two mixed-type cell groups of identical FS and Pyr cell numbers in each. If the data is solid and suitable for quantitative analysis, comparison of that two randomly composed groups should give similar copy numbers even if the gene exhibits large differences between FS and Pyr cells. If the copy numbers are not reliable (e.g. largely fluctuating) or the gene could be detected only in a very few cells, such a randomized group comparison should show large fluctuations. We evaluated one million random distribution of the cells into the two groups of equivalent composition and defined a gene reliable if in >95% of the random choices resulted less than tenfold difference between the two groups in the copy numbers. This filtering verifies the reliability for differently expressing genes between FS and Pyr cells (Figure S4). The same method was applied in the comparison of layer 2, 3 and layer 5, 6 Pyr cells.

The raw sequencing data sets of FS and Pyr cells were compared on the 409 genes of "equi-phenotype" and on all the genes observed. As an average, the harvesting of the cytosol resulted in twice higher amount of mRNA molecules from Pyr cells which might be explained by their larger size and volume (Table S7).

Verification of the sequencing data by PCR

The mRNA content of the single-cell was subjected to one round of aRNA amplification, after the aRNA was converted into cDNA using random hexanucleotide primers. This cDNA was used as template for the PCR reactions. We used synthetic oligonucleotide primers derived from the *Celsr1*, *Kcnmb1*, *Gria3* and *Adora1* mRNA for amplification of the corresponding mRNA

fragments. PCR amplification was conducted in a 50 µl reaction mixture containing 1 µl of each cDNA template, 2.5 µl of each 10 µM primer, 4 µl dNTP (2.5mM), 1U of Q5 High-Fidelity DNA polymerase (NEB, #M0491L), and 10 µl of 5x reaction buffer according to the manufacture's protocol. The PCR reaction went for 40 cycles in a Bio-Rad T100 thermal cycler. Each cycle consisted of denaturation at 98°C for 10 sec, annealing at 55°C for 30 sec, and elongation at 72°C for 30 sec. The amplified PCR products were detected using D5000 Screentape (Agilent Santa Clara, CA, USA). PCR products were purified after electrophoresis through 2% agarose gels in the presence of 0.5 µg of ethidium bromide per ml and submitted for sequencing.

Results

Electrophysiological classification of PFC neurons

We used a step gradient depolarization protocol for electrophysiological characterization of FS and Pyr cells (Figure 1B). Twenty-two parameters of their electrical activity were measured as described in Figure S1. The PC analysis showed that the first 3 PCs provide 89% accuracy in clustering the PFC neurons into Pyr cell group and FS interneuron group, in agreement with previous publications (Krimer et al. 2005) (Figure 1C). Clustering on the basis of the first frame of continuous firing and the last frame of highest firing rate showed the same results. Therefore, we relied upon the first frame data in our study. Based on this cluster analysis, out of the 25 cells previously selected as FS, two cells were classified as Pyr cell and only two out of the 59 Pyr cells were found to be FS. This indicates the accuracy of electrophysiology-based cell selection for single-cell transcriptome analysis of PFC neurons. There were sub-clusters of FS and Pyr cells discernible on the basis of their electrical activity. We were unable to correlate these sub-clusters with any of the assessed brain functions or

location (layer) and, therefore, we used only the main cluster of FS and Pyr cells in further analysis.

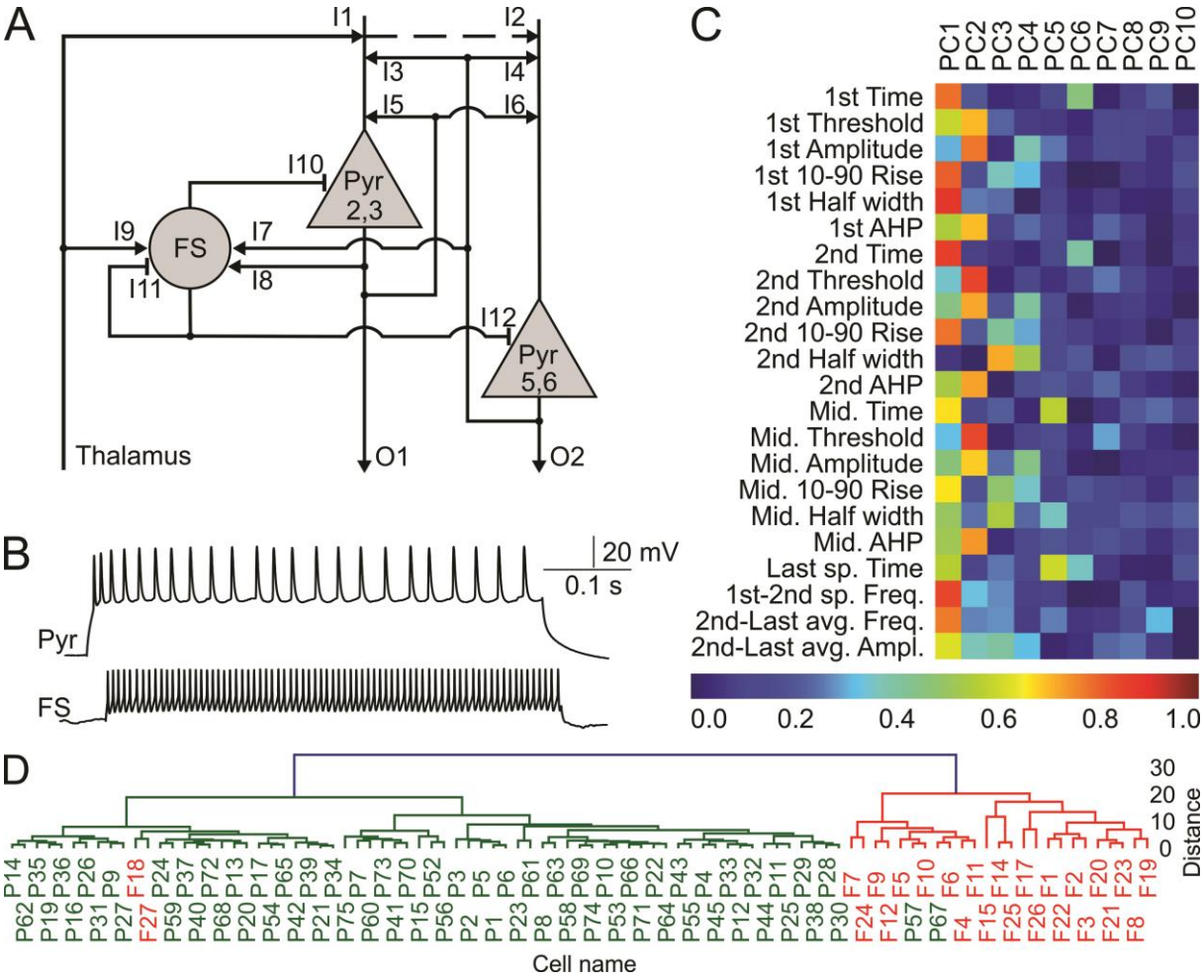

Figure 1. Physiological clustering of PFC cells. (A) The canonical network of PFC containing inputs (I1-I12) and outputs (O1-O2) of FS and Pyr cells. **(B)**Typical firing pattern of Pyr and FS cells. Representative frames from step-gradient measurements. **(C)** Principal component (PC) analysis of 22 physiological parameters and contribution of each physiological parameter to the first 10 PCs. These contributions are comparable within and between PCs, because of the proper normalization (see Materials and Methods and Figure S1). **(D)** Hierarchical clustering of FS and Pyr cells confirmed that FS and Pyr cells are clearly distinguishable and physiologically determined clusters of PFC neurons. Green represents the FS cell cluster (only 2 Pyr cells are included), red represents the Pyr cluster (only 2 FS cells are included).

Transcriptional profiling of Pyr and FS cell surfaceome

After deep single-cell sequencing (10 million reads per cell) we found 19,685 transcripts in total of PFC neurons. Normalization of single-cell transcriptomics data, especially when the

sample is harvested from the living cell by patch clamp technique, presents some special challenges because of the great number of zero expression values. The application of the normalization process described above resulted in the positive correlation between the number of recovered transcripts and the observed gene number of raw data (Figure S2C). Plotting the average normalized copy numbers of transcripts in FS and Pyr cells against each other, we identified expression differences between the two types of neurons (Figure 2). The expression levels were tested for reliability (see reliability test in Materials and Methods).

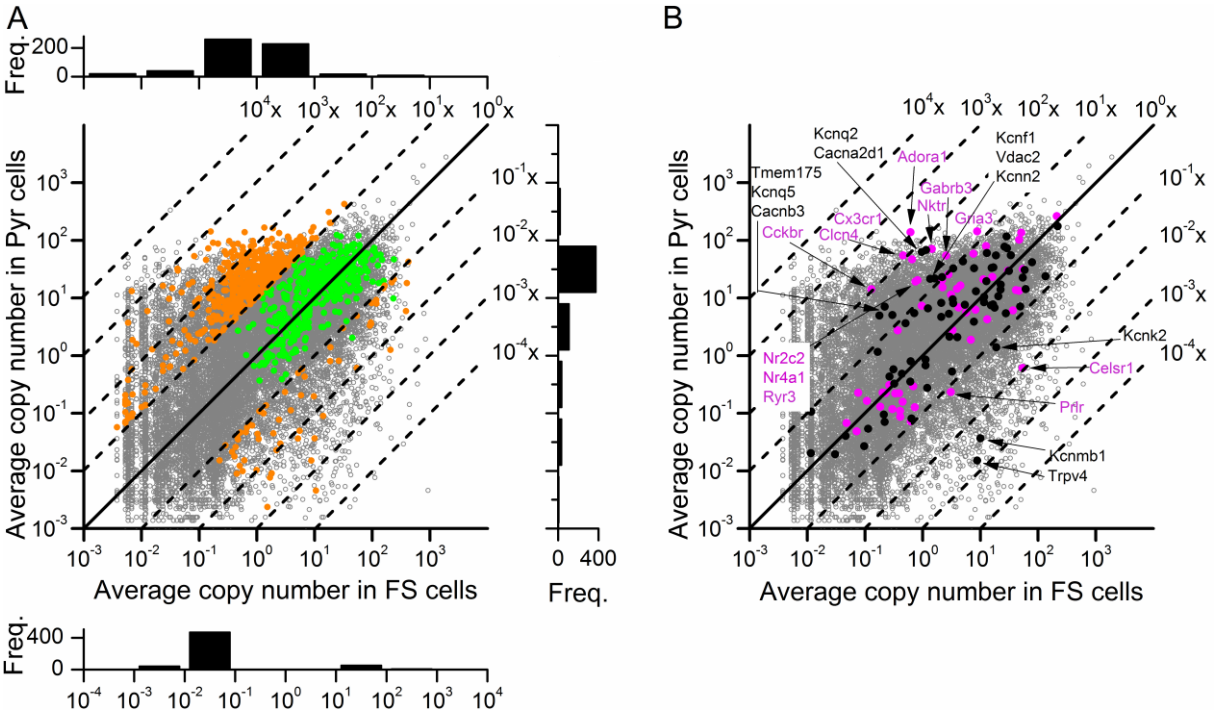

Figure 2. Transcriptomic differences between Pyr and FS cells. (A) Average copy numbers in Pyr cells vs. FS cells on logarithmic scale. Points represent individual genes. Genes representing the "equi-phenotype" and used for normalization are presented in green. Those reliable genes that show higher than tenfold difference in transcription levels between FS and Pyr cells are presented in orange. The histograms show the distribution of these genes as a function of average expression level in FS cells (top) and in Pyr cells (right), furthermore as a function of FS/Pyr average expression level ratio (bottom). Note that the indicated values belong to the diagonal dashed lines. **(B)** Comparison of receptor (purple) and ion channel (black) transcripts in FS and Pyr cells. Only genes with reliable copy numbers are highlighted.

The scatter plot of FS and Pyr cell transcript abundances (Figure 2A) revealed that at least 54% of the transcripts are evenly expressed and 22% of them were highly reliable in both major

neuron types. The evenly expressed transcripts are coding housekeeping proteins and scaffolding proteins as well as proteins related to neurogenesis, neuronal differentiation and development. The abundant expression of housekeeping genes (e.g. *Actb* and *Gapdh*) indicates the accuracy of harvesting and the reliability of dataset (van den Hurk et al. 2018). We found that *Actb* was expressed by 87% and *Gapdh* by 60% of the cells, other housekeeping genes like *Eef1a1* and *Tuba1a* were also detected in the vast majority of cells (in 85% and 92% of them, respectively). A fraction of evenly expressed transcripts were used for normalization labeled by green on Figure 2A. In addition, transcripts of several neurotransmitter and neuropeptide receptors (Figure 2B) were evenly expressed, such as some of the acetylcholine, adenosine, angiotensin, histamine, lysophospholipid, neuropeptide FF, NPY, cholecystokinin A and melanocortin receptors (Table S1). Transcripts coding receptors that facilitate lateral synaptic communication in the PFC, such as glutamate and GABA receptors transcripts were expressed abundantly in both Pyr and FS neurons in agreement with pharmacological and electrophysiological studies (Mclennan 1983; McCormick et al. 1993). Further the mRNAs of receptors for cytokines, chemokines, and purinergic receptors, normally assumed to be expressed by microglia, astroglia, dendritic cells and to lesser extent by neurons in the PFC (Hanisch 2002; Ambrosini et al. 2005; Kataoka et al. 2009) were found evenly distributed in FS and Pyr neurons. The pan-neuronal marker genes *Snap25* (Tasic et al. 2016), *Stmn2* and *Thy1* (Fuzik et al. 2016) showed stable expression in the sequenced neurons (they were detected in 86%, 76% and 65% of the cells, respectively). We also found the mRNAs of classical Pyr and FS cell immunostaining markers in the single-cell transcriptomes (Figure 3B). The glutamatergic Pyr neurons expressed the classical vesicular glutamate transporter marker gene *Vglut1* (mouse gene symbol: *Slc17a7*) at 25-fold higher level than the FS cells. *Camk2b*, an excitatory neuron marker gene (Jones et al. 1994) also showed enriched (13-fold) expression in Pyr cells. The expression of the GABAergic cell marker gene *Gad1* was 65-fold higher in FS

cells, however some of the Pyr cells also expressed it. The mRNA level of *Slc6a1* coding the GABA transporter GAT-1 was also 27-fold higher in FS interneurons. Thus, our results regarding the expression pattern of pan-neuronal, pan-excitatory and pan-inhibitory markers seem to be consistent with earlier single-cell sequencing studies based on patch-seq (Fuzik et al. 2016) and sort-seq (Zeisel et al. 2015; Tasic et al. 2016) techniques.

The transcriptome of FS and Pyr cells contained large number of differentially expressed mRNAs showing 10-20-fold differences in copy numbers. Relying on the copy numbers and expression probability, we identified receptor and ion channel mRNAs for further investigation as candidates for cell-specific pharmacological targeting of Pyr and FS cells (Figure 2B, 3, Table 1). The expression levels of the potassium channel subunit encoding genes like *Kcnf1*, *Kcnk4*, *Kcnn2*, *Kcnq2*, *Kcnq5*, *Tmem175* were more than tenfold higher in Pyr cells than in FS cells. Voltage-dependent calcium channel subunit alpha-2/delta-1 (*Cacna2d1*) and L-type calcium channel subunit beta-3 (*Cacnb3*) were also highly expressed in Pyr cells. Also, Pyr cells highly expressed the mitochondrial voltage dependent anion-selective channel (*Vdac2*) mRNA. The *Ryr3* coding ryanodine receptor 3, an endoplasmic reticulum calcium channel was also highly expressed in Pyr cells. The transcript numbers for GABA receptor subunit beta-3 (*Gabrb3)* and the ionotropic glutamate receptor AMPA type subunit-3 (*Gria3*) were several times higher in Pyr cells than in FS cells. In Pyr cells, the cholecystokinin B receptor (*Cckbr*) was also highly expressed (118-fold higher than in FS cells). There was high expression of adenosine A1 receptor (*Adora1*), proton/chloride exchange transporter 4 (*Clcn4*), C-X3-C motif chemokine receptor 1 (*Cx3cr1*) and NK-tumor recognition protein (*Nktr*). Opioid receptor Mu 1 (*Oprm1*) was found exclusively in Pyr cells. The nuclear receptors *Nr2c2*, *Nr4a1* are also remarkable because they express ligand dependent transcription factors activated by hormones (e.g. steroids) and other signaling molecules (e.g. retinoic acid) acting on neurons (Chen et al. 2008; Y. Chen et al. 2014).

In FS cells, the transcript number of transient receptor potential cation channel (*Trpv4*) permeable for calcium was ~580-fold higher than in Pyr cells (Figure 2B, 3, Table 1). Also, the average expression of calcium-activated potassium channel subunit beta-1 (*Kcnmb1*) was ~270 times higher in FS cells than in Pyr cells. The transcripts of potassium voltage-gated channel subfamily C member 2 (*Kcnc2*) and potassium channel subfamily K member 2 (*Kcnk2*) were also enriched in FS cells. Additional significantly enriched mRNAs in FS cells are coding angiotensin receptor (*Agtr2*), prolactin receptor (*Prlr*), histamine receptor (*Hrh2*) and cadherin EGF LAG seven-pass G-type receptor 1 (*Celsr1*). The highly selective expression of the gastrin releasing peptide (*Grp*) is also noteworthy in FS cells.

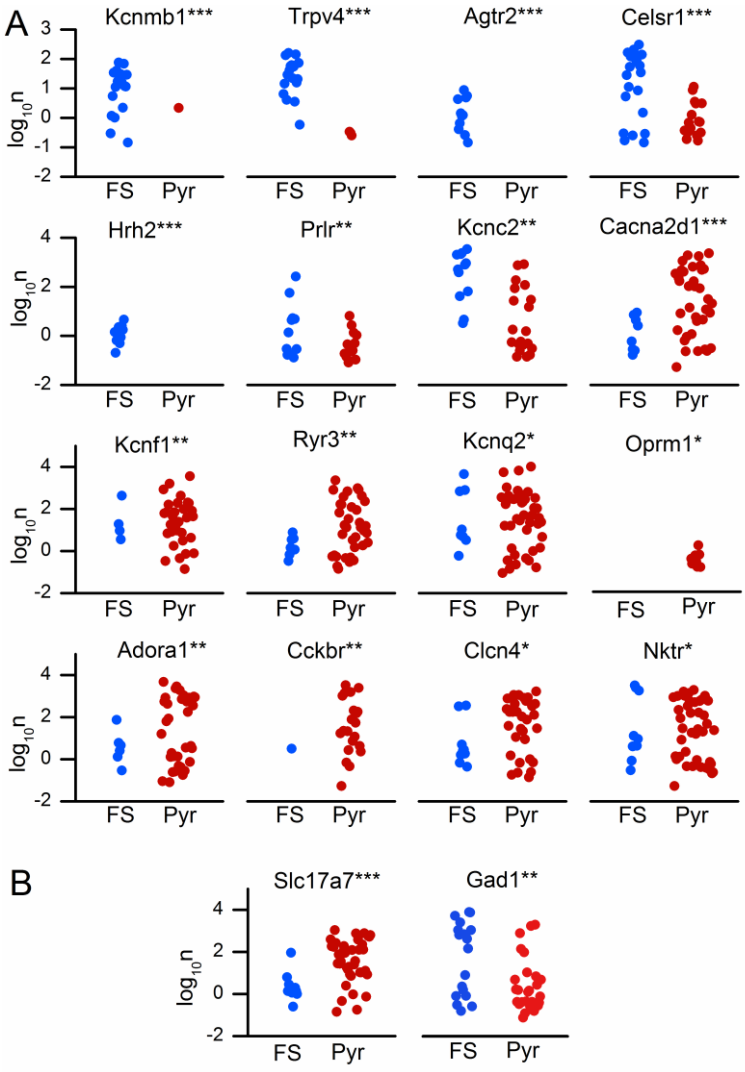

Figure 3. The distribution of differentially expressed transcripts. The figure shows the normalized copy numbers and the distribution by cell type of some of the differentially expressed receptor and ion channel transcripts **(A)** and classical marker

genes **(B)**. Points represent individual cells. Only genes having at least 0.95 reliability score are presented. Please note the difference between the number of the sequenced FS (n=25) and Pyr (n=59) cells (*: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$ according to the two-sample *t*-test).

For verification of the sequencing results we performed PCR experiments on 12 Pyr and 5 FS cells and, however, we note here that single-cell sequencing is more appropriate for identifying endogenous cell sequences as it is more specific and sensitive. We examined the presence of 4 differentially expressed transcripts and similarly to the sequencing results found varying densities in the investigated cells (for a representative gel image see Figure S5). PCR products from *Gria3* analysis were confirmed by sequencing.

We compared our differentially expressed mRNA results to the cell surfaceome database (http://wlab.ethz.ch/surfaceome/) (Figure 4). The surfaceome database was constructed on the basis of the human genome and only about 6000 genes could be matched to the mouse genome, thus we limited our analysis on them. The surfaceome scores in the database represent the probability of a protein to be located on the outer surface of the cell membrane. Our data showed a marked overlap with the human surfaceome database and the transcripts of GPCRs and other receptors having tenfold or higher difference between Pyr and FS cells (Figure 4B) have high surface score as well (above 0.9). We note here that beyond receptors and ion channels of more than tenfold transcript number differences we found also other differentially expressed transcripts of high surface location score. The focus of this study was on receptors and ion channels and, however, some of the other proteins having high surface location score could be used for cell-selective targeting in the future.

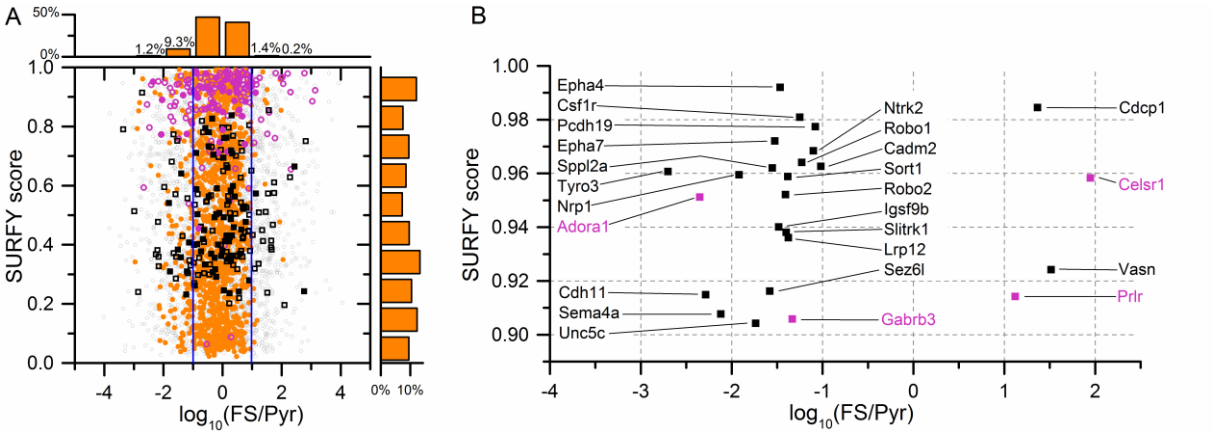

Figure 4. Comparison of our data and the human surfaceome database. The figure shows only those ~6000 mouse genes that have human orthologs in the database. (A) The surfaceom (SURFY) score as a function of average copy numbers in FS/Pyr cells. Highlighted genes: genes with reliable copy numbers (filled orange dots), genes coding receptors (empty purple dots), reliable genes coding receptors (filled purple dots), genes coding ion channels (empty black squares) and reliable genes coding ion channels (filled black squares). (B) The reliable differentially expressed genes with high surfaceome score (>0.9) suggesting cell surface localization of the coded protein. In addition to the above identified cell type-enriched receptor transcripts (purple), the in silico analysis revealed several other differentially expressed genes coding potential cell surface proteins (black).

Transcriptional differences between layer 2, 3 and layer 5, 6 Pyr cells

Layer 2, 3 and layer 5, 6 Pyr cells in the PFC are directly connected by axon collaterals projecting to the dendritic layer (lateral connections) while Pyr cells of different layers have distinct efferent projections (Figure 1A). The anatomical connection differences between Pyr cells of layer 2, 3 and layer 5, 6 suggest functionally different roles in healthy and diseased PFC; however, the parameters derived from the step gradient depolarization paradigm did not cluster them as electrophysiologically different groups (Figure 1D).

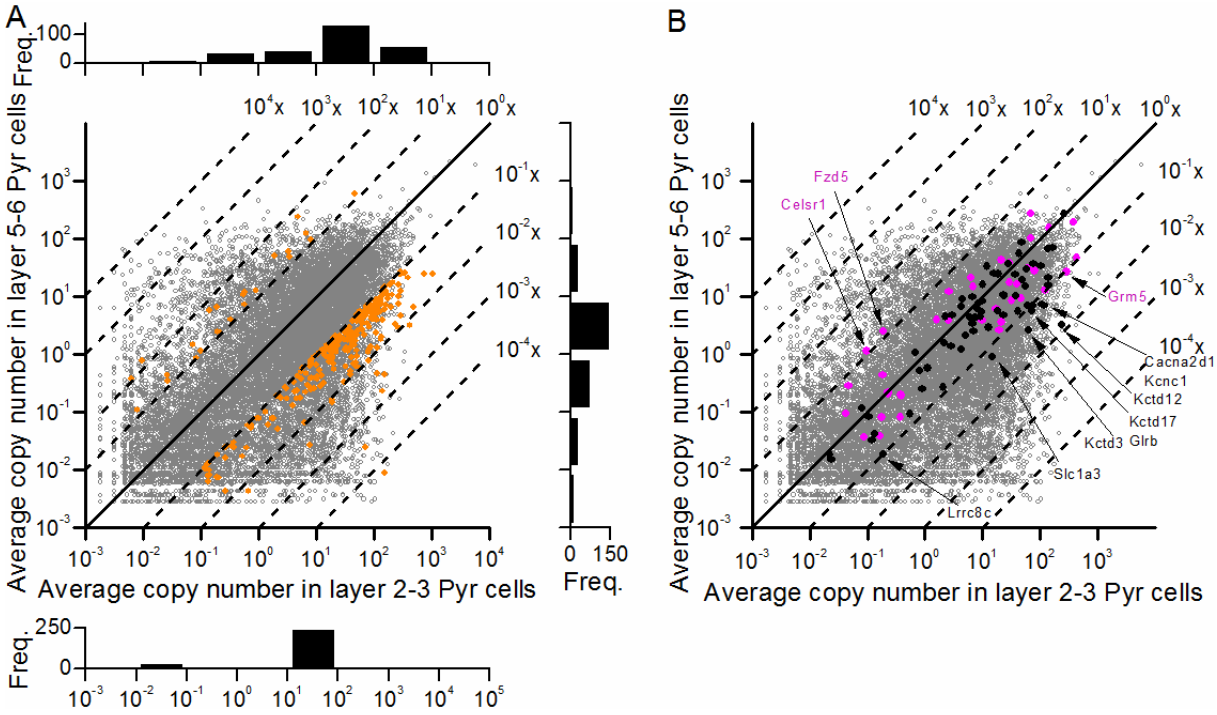

Figure 5. Gene expression differences between Pyr cells of layer 2, 3 and layer 5, 6 in PFC. (A) Average copy numbers in layer 2, 3 cells vs. layer 5, 6 cells on logarithmic scale. Points represent individual genes. Genes that show higher than tenfold difference in transcription levels are presented in orange. The histograms show the distribution of these genes as a function of average expression level in layer 2, 3 Pyr cells (top) and in layer 5, 6 Pyr cells (right), furthermore as a function of layer 2, 3/layer 5, 6 average expression level ratio (bottom). Note that the indicated values belong to the diagonal dashed lines. **(B)** Comparison of receptor (purple) and ion channel (black) genes in layer 2, 3 and layer 5, 6 Pyr cells. Only genes with reliable copy numbers are highlighted.

We found several differences between the transcripts of layer 2, 3 and layer 5, 6 Pyr cells supporting layer specific functional differences (Figure 5, Table 1, Table S4). Layer 2, 3 Pyr cells expressed *Kcnc1* inward rectifying rapid potassium channel ~17 times higher than layer 5, 6 Pyr cells. Also, layer 2, 3 Pyr cells transcribed *Cacna2d1* at a ~15 times higher level than layer 5, 6 Pyr cells. The transcript number of sodium and potassium dependent excitatory amino acid (glutamate and aspartate) transporter 1 (*Slc1a3*) was ~16 times higher in layer 2, 3 Pyr cells suggesting more efficient glutamate reuptake after excitation in layer 2, 3 than in layer 5, 6. The non-essential component of volume regulated anion channel (*Lrrc8c*) expression was ~tenfold higher in layer 2, 3 Pyr cells. Also, the glycine-gated hyperpolarizing ion channel

(*Glrb*) expression was ~13 times higher in layer 2, 3 Pyr cells suggesting enhanced glycine sensitivity of upper layer Pyr neurons. There were other potassium channels highly expressed in layer 2, 3 Pyr cells, such as *Kctd3*, *Kctd12* and *Kctd17*.

The only neuronal GPCR transcript difference is the ~11-fold enrichment of *Grm5* in the layer 2, 3 Pyr cells. The ~14-fold higher presence of *Nr4a1* mRNAs is currently difficult to interpret. We found 243 miscellaneous protein transcripts (mostly metabolic and scaffolding proteins) that are highly expressed in layer 2, 3. They are not cell-surface molecules and, therefore, not easily druggable or useful for selective cell targeting.

Layer 5, 6 Pyr cells showed fewer highly transcribed mRNAs compared to layer 2, 3 cells (Figure 5B). Two genes coding GPCRs, the *Celsr1* and frizzled-5 (*Fzd5*) were highly expressed in layer 5, 6 Pyr cells. We have to note that *Celsr1* showed the most abundant expression in FS cells (~89-fold enrichment), we merely mention it here because of the layer-specific expression in Pyr cells. We did not find any ion channel mRNA which had at least tenfold higher copy number in layer 5, 6, than in layer 2, 3 Pyr cells. Also, 21 miscellaneous protein transcripts were selectively transcribed (with more than tenfold difference) that do not code cell surface molecules.

To demonstrate the usability of our single-cell transcriptomic results we chose *Prlr* (highly expressed in FS cells), which has not been previously investigated in PFC FS or Pyr cells and, however, its electrophysiological effect was described in hypothalamic neurons (Brown et al. 2012). Thus, we tested the PRL-R functionality with 500 ng/ml prolactin (R&D Systems, Minneapolis, MN, USA) bath-application on acute brain slices during whole-cell patch clamp recordings. We have seen an increase in 4 and a decrease in 1 out of 5 FS cells in the first-second spike frequency during prolactin application (106.33±7.56% compared to control). In turn, out of 5 Pyr cells 2 showed an increase and 3 showed a decrease in frequency of the two first spikes (101.04±15.45% compared to control). Since we applied prolactin on

brain slices containing a complex neuronal network only the presence of the functional PRL-R was confirmed but we cannot tell anything about the detailed mechanisms of the response.

Discussion

Using patch-seq analysis in cell-selective target discovery

Single-cell transcriptomic studies in the CNS aimed unbiased discovery of neuronal subtypes (Zeisel et al. 2015; Tasic 2018), thus in the past decade large number of cells were sequenced in clustering studies (Valihrach et al. 2018). Beside of electrophysiological clustering the aim of this study was to discover transcriptomic differences between the two classical cell types of the PFC because the knowledge about Pyr/FS cell interplay in pathological functions raise the need for selective fine tuning of them in brain diseases. Druggability of cell surface molecules suggests uncovering differences in mRNAs coding cell surface receptors and ion channels beyond the already known markers of the two cell types (Bartfai et al. 2012).

To avoid modifications of the transcriptome induced in transgenic animals, as well as the gene expression alterations due to cell dissociation techniques, we used acute brain slices, where the neurons had their local synaptic system organizing the excitatory/inhibitory connections in the PFC mainly preserved. Performing single-cell sequencing based on cell dispersion methods can investigate the transcriptome of thousands of cells and, however, our goals required the electrophysiological identification of neurons and ultra-deep sequencing. In addition, as Pyr cells are larger than FS cells, during the dispersion procedure the cell isolating wells may contain more than one FS cell. For these reasons we preferred to use the patch-seq technique.

In general, former patch-seq studies were performed on a limited number of neurons (Tripathy et al. 2018) allowing ultra-deep sequencing up to the complete transcriptome analysis

of cells. Ultra-deep sequencing requires different data processing and quality controlling strategies than the low-depth sequencing studies (Sims et al. 2014). The detection of extremely low copy number of transcripts generates the problem of high number of zero copy numbers. As for the interpretation of zero expression values, Levine et al. (2013) reviewed the pulsatile nature of cellular processes suggesting pulsatile dynamics for the gene expression as well. Thus, some of the genes might are transcribed in bursts at certain frequency determined by the function of the coded proteins in pulsatile cellular functions (Suter et al. 2011; Liu et al. 2016). Since mRNA transcription is oscillating, at a certain time point several genes can be in OFF state while others in ON state (Munsky et al. 2012). Therefore, the zero expression value of a gene in a cell could be of biological origin at a certain time point as well (Lin et al. 2017). It cannot be excluded that zero expression values might derive from methodological uncertainties. We emphasize that zero expression values of technical origin may influence the data quality obtained from a cell which cannot be restricted to certain genes but can result low copy number for several genes. In turn, we got rid of the cells in which all copy numbers were extraordinarily low.

A technical and theoretical problem in patch-seq data processing is the normalization of the raw data. Using different normalization methods, we revealed that in some cases the total mRNA copy number in a cell does not correlate positively with the total number of genes in the raw sequencing dataset as it should be. In some studies the copy numbers of housekeeping protein mRNAs were used for normalization (Lin et al. 2019), but it is known that the expression of them can show pulsatile dynamics at single-cell level (Liu et al. 2016). However, when we applied the lowest dispersion and highest copy number transcripts for normalization without any respect to the function of the coded proteins, positive correlation was established between the total normalized copy number and the total original gene number (Figure S2C).

Receptor and ion channel differences are linked to brain diseases and drug effects

We were able to identify many differentially expressed mRNAs which are coding proteins of pathological importance. The proteins coded by some of the uncovered cell-selective transcripts (Table 1) have functions already described in pathophysiology pathways of brain diseases (Table S2, S3). Differential transcription patterns are important in Fragile-X syndrome (Martin et al. 2016), circadian rhythm regulation (Gannon and Millan 2011), mental retardation (Lehman et al. 2017) and autism (C.-H. Chen et al. 2014). In particular, PRL-R (transcribed highly in FS cells) can enhance neuronal firing (Brown et al. 2012) and is associated with schizophrenia (Albayrak et al. 2014), *Hrh2* (transcribed highly in FS cells) plays a role in Fragile-X syndrome (Wright et al. 2017), and *Gabrb3* (transcribed highly in Pyr cells) is associated with epilepsy and autism (C.-H. Chen et al. 2014). Many of the receptors and ion channels that showed cell type-specific mRNA expression are involved in the regulation neuronal excitability (e.g. *Adora1*, *Oprm1*, *Trpv4*, *Kcnf1* and *Kcnq2*) (Shibasaki et al. 2007; Cioli et al. 2014; Soh et al. 2014; Rombo et al. 2016; Wang et al. 2018). The proteins coded by *Kctd3*, *Kctd12* and *Kctd17* are tetramerization domain members of potassium channels and are involved in Tic-disorder and myoclonic diseases (Cao-Ehlker et al. 2013; Correale et al. 2013; Mencacci et al. 2015). *Agtr2* has a role in X-linked intellectual disability (Vervoort et al. 2002) and Alzheimer's disease (Ge and Barnes 1996). *Kcnmb1*, highly transcribed in FS cells (Table 1), codes the regulatory subunit of the high conductance voltage and calcium dependent potassium channel. It controls the calcium sensitivity and gating mechanisms of the channel indicating an efficient dampening regulatory role of fast depolarization (Jiang et al. 1999). The most important finding of the present study is that we revealed many unevenly distributed mRNAs in Pyr and FS cells of the murine PFC, coding proteins involved in psychiatric diseases in human. Moreover, the mechanism of action of some drug candidates can be linked to certain cell types of the PFC.

Table 1. Transcripts encoding putative drug targets that are differently expressed at high levels in FS vs. Pyr cells. Only genes with reliable copy numbers are listed*

Transcript expression	Ion channel subunit	GPCR and other receptor subunit	Classical markers
High in FS cells	Kcnmb1#, Kcnk2, Trpv4#**,** Kcnc2	Agtr2#, Celsr1#, Hrh2#, Prlr#	Calb1, Calb2, Gad1†, Gad2, Pvalb, Sst, Vip, Slc6a1†
High in Pyr cells	Cacnb3, Cacna2d1, Kcnf1, Kcnk4#, Kcnn2, Kcnq2, Kcnq5, Tmem175, Vdac2	Adora1#, Cckbr, Cx3cr1#, Gabrb3, Gria3, Nktr, Nr2c2, Nr4a1, Oprm1#, Ryr3#	Cck†, Slc17a7†, Camk2b†
High in layer 2, 3 vs. layer 5, 6 Pyr cells	Cacna2d1, Kcnc1, Kctd3, Kctd12, Kctd17, Lrrc8c#	Glrb, Grm5, Nktr, Nr4a1	
High in layer 5, 6 vs. layer 2, 3 Pyr cells	Not above the average of Pyr cells	Celsr1#, Fzd5	

* For comparison, we list the transcripts encoding the classical markers of these neuron types and note that these are also differently i.e. more than tenfold higher expressed in one than in the other cell type. †classical markers expressed differently based on transcriptomics, #labels transcripts of subliminal expression in Allan Brain Atlas

The proteins coded by the differentially expressed mRNAs are in conjunction with pharmacological experiments done on drug candidates commercially available. It was shown on mice that Spadin (Borsotto et al. 2015), a novel antidepressant is the inhibitor of TREK1 channel (coded by *Kcnk2*) which was highly expressed in PFC FS cells suggesting its FS cell-specific action. Riluzole, a TRAAK channel (coded by *Kcnk4*) antagonist decreases the glutamate release and is clinically used for the treatment of amyotrophic lateral sclerosis (ALS) (Duprat et al. 2000). *Kcnk4* was exclusively expressed by Pyr cells. Linopirdine and XE991 are

cognitive enhancers acting on ACh release and are antagonists of voltage-gated potassium channel subunit Kv7.5 (coded by *Kcnq5*) (Schnee and Brown 1998; Wang et al. 2000) which was highly transcribed in PFC Pyr cells. Finally, Erastine (Yagoda et al, 2007) is a VDAC-2 and glutathione synthesis inhibitor (Yagoda et al. 2007) and the *Vdac2* was also highly expressed in Pyr cells. Since changes in excitatory/inhibitory balance of the PFC require selective modulation of FS and/or Pyr cells, the pharmacological data indirectly support the differential expression of the target proteins suggested by our data.

The *in silico* comparison of our data to the human cell surfaceome database confirmed the surface localization of the selected receptors and ion channels but also showed several other differentially expressed mRNAs coding miscellaneous cell surface proteins (Figure 4B). We note here that these membrane proteins having no known receptor or ion channel functions could attract interest as anchor proteins for cell-selective drug targeting; however, the currently available technologies are developed for investigating receptors (e.g. GPCRs) and ion channels. We suggest further studies on these proteins in the future.

The layered architecture of the mammalian neocortex and the functional differences between Pyr cells of different cortical layers suggest that the mRNA and protein expression patterns could be layer-specific (Shepherd 2004). Indeed, we report here layer-specific mRNA differences, but the number of remarkable expression differences between Pyr cells of superficial and deep layers was limited in comparison with the Pyr/FS cell differences. Thus, we assume that layer-specific targeting of Pyr cells despite of its functional relevance probably cannot be performed.

Uncertainties in the relationship between mRNA and protein levels

Because of the limitations of single-cell proteomics, the most useful way for getting reliable information about the molecular composition of individual cells is the single-cell

mRNA sequencing. The correlation between the transcript copy numbers and the concentration of the coded protein might be weak in the case of steady state level mRNAs but it can be high in the case of variable level mRNAs (Maier et al. 2009). In addition, the correlation between mRNA and protein levels is gene and tissue specific, and the correlation factors might be determined by the current condition of the cells (Liu et al. 2016).

Assuming that the synthesis intensity of a certain protein in different neurons may be similar, the mRNA differences could reflect protein differences in some extent. Therefore, we assume that a tenfold or higher difference in mRNA copy numbers may indicates a difference of similar magnitude in the protein levels as well, which could be high enough for selective targeting of cells. It should also be noted that the presence of a receptor transcript in a cell do not allow a direct conclusion on the functionality of the receptor protein, but it may indicate the potential of a cellular response (Bartfai et al. 2012). We also should take into account the probable correlation between mRNA copy numbers and cell size, considering Pyr cells are usually two or three times larger than FS cells. The same density of a receptor or ion channel in the cell membrane probably needs twice as much mRNAs in a Pyr cell then in an FS cell. It can explain the generally higher mRNA copy numbers in Pyr cells than in FS cells. Thus, based on transcriptomic differences we offer here a cell surface receptor and ion channel set probably differentially expressed in Pyr and FS cells which might enable the selective pharmacological targeting of the two cell types. We should emphasize that it is only an assumption because the exact validation of the results by single-cell proteomics is unrealistic at the current state of the art and the *in vivo* target selectivity of drugs depends on several complex mechanisms (Vlot et al. 2017). Probably the understanding of mRNA copy number and protein level correlation is one of the challenges in the future single-cell OMICS. In conclusion we propose classical experimentation to demonstrate how the proteins coded by evenly or differentially expressed transcripts could selectively tune the cellular activity in the PFC.

Conclusion

As a consequence of the extreme sensitivity of ultra-deep single-cell sequencing we uncovered several transcriptomic differences between Pyr and FS cells that have not been detected in the mouse PFC before (Table 1. # labeled differences). Also, we confirmed the mRNA expression differences of the already known Pyr and FS cell marker proteins. The results support our principal idea that application of single-cell sequencing to reveal novel druggable surfaceome differences can be achieved efficiently on electrophysiologically characterized neuronal cell types. We suggest the application of this workflow for discovery of molecular differences between functionally and pathologically important cell types of the neocortex to improve drug development for connectome diseases like schizophrenia, autism or depression. The present data suggest that the cell type-selective drug targeting of Pyr and FS cells is a feasible possibility. The layer-selective targeting of Pyr cells may prove to be more challenging as there are less selective transcript differences between the Pyr cells of layer 2, 3 and layer 5, 6. Beyond identifying several cell-specific surfaceome differences in the PFC, we demonstrated that single-cell sequencing of electrophysiologically selected, genetically unaltered and morphologically intact neurons is useful in novel and cell type-specific drug target discovery. Moreover, it can provide important insight into the molecular mechanism of action of psychoactive compounds that is envisioned to support drug discovery efforts for treatment of mental illnesses.

Acknowledgements. The present studies were supported by the National Research, Development and Innovation Office of Hungary (KTIA NAP_B_13-2-2014-0004 to LR and MIT, 2017-1.2.1-NKP-2017-00002 to KAK and GJ, FIEK_16-1-2016-0005 to DM, KAK, ZB, LR and GJ). ME-R is supported by grants of EU Horizon 2020 No 668863 (SyBil-AA) - and

CNCS-UEFISCDI COFUND FLAGERA II-CORTICITY and PN-III-P1-1.1-TE-2016-1457. BT is supported by the Brandeis Center for Bioinspired Soft Materials, an NSF MRSEC, DMR-1420382. JE and JW were supported by NIH grant MH10180. We thank András Micsonai and József Kardos for working out the methodology of normalization and the reliability test for expression differences between cell types and Éva Bulyáki for control experiments.

Author contributions. LR, ZB, TB, JE and GJ designed the experiments, LR, MIT, ZB, JW, JE and GJ performed the experiments, LR, DM, KAK, ZB, ME-R, BT, JW, TB, JE, GJ analyzed the data, LR, DM, KAK, TB, JE and GJ wrote the paper and all co-authors reviewed the manuscript and provided comments.

References

Albayrak Y, Beyazyüz M, Beyazyüz E, Kuloğlu M. 2014. Increased serum prolactin levels in drug-naive first-episode male patients with schizophrenia. Nord J Psychiatry. 68(5):341–346. doi:10.3109/08039488.2013.839739.

Ambrosini E, Remoli ME, Giacomini E, Rosicarelli B, Serafini B, Lande R, Aloisi F, Coccia EM. 2005. Astrocytes produce dendritic cell-attracting chemokines in vitro and in multiple sclerosis lesions. J Neuropathol Exp Neurol. 64(8):706–15. doi:10.1097/01.jnen.0000173893.01929.fc.

Bar-Joseph Z, Gifford DK, Jaakkola TS. 2001. Fast optimal leaf ordering for hierarchical clustering. Bioinformatics. 17. doi:10.1093/bioinformatics/17.suppl_1.S22.

Bartfai T, Buckley PT, Eberwine J. 2012. Drug targets: Single-cell transcriptomics hastens unbiased discovery. Trends Pharmacol Sci. 33(1):9–16. doi:10.1016/j.tips.2011.09.006.

Bomkamp C, Tripathy SJ, Gonzales CB, Hjerling-Leffler J, Craig AM, Pavlidis P. 2019.

Transcriptomic correlates of electrophysiological and morphological diversity within and across excitatory and inhibitory neuron classes. PLoS Comput Biol. 15(6):e1007113. doi:10.1371/journal.pcbi.1007113.

Borsotto M, Veyssiere J, Moha Ou Maati H, Devader C, Mazella J, Heurteaux C. 2015. Targeting two-pore domain K+ channels TREK-1 and TASK-3 for the treatment of depression: a new therapeutic concept. Br J Pharmacol. 172(3):771–784. doi:10.1111/bph.12953.

Brown RSE, Piet R, Herbison AE, Grattan DR. 2012. Differential actions of prolactin on electrical activity and intracellular signal transduction in hypothalamic neurons. Endocrinology. 153(5):2375–2384. doi:10.1210/en.2011-2005.

Buzsáki G, Wang X-J. 2012. Mechanisms of Gamma Oscillations. Annu Rev Neurosci. 35:203–225. doi:10.1146/annurev-neuro-062111-150444.

Cadwell CR, Palasantza A, Jiang X, Berens P, Deng Q, Yilmaz M, Reimer J, Shen S, Bethge M, Tolias KF, et al. 2016. Electrophysiological, transcriptomic and morphologic profiling of single neurons using Patch-seq. Nat Biotechnol. 34(2):199–203. doi:10.1038/nbt.3445.

Cao-Ehlker X, Zong X, Hammelmann V, Gruner C, Fenske S, Michalakis S, Wahl-Schott C, Biel M. 2013. Up-regulation of hyperpolarization-activated cyclic nucleotide-gated channel 3 (HCN3) by specific interaction with K+channel tetramerization domain-containing protein 3 (KCTD3). J Biol Chem. 288(11):7580–7589. doi:10.1074/jbc.M112.434803.

Chen C-H, Huang C-C, Cheng M-C, Chiu Y-N, Tsai W-C, Wu Y-Y, Liu S-K, Gau SS-F. 2014. Genetic analysis of GABRB3 as a candidate gene of autism spectrum disorders. Mol Autism. 5(1):36. doi:10.1186/2040-2392-5-36.

Chen Y-T, Collins LL, Chang S-S, Chang C. 2008. The roles of testicular orphan nuclear

receptor 4 (TR4) in cerebellar development. The Cerebellum. 7(1):9–17. doi:10.1007/s12311-008-0006-3.

Chen Y, Wang Y, Ertürk A, Kallop D, Jiang Z, Weimer RM, Kaminker J, Sheng M. 2014. Activity-induced Nr4a1 regulates spine density and distribution pattern of excitatory synapses in pyramidal neurons. Neuron. 83(2):431–443. doi:10.1016/j.neuron.2014.05.027.

Choromanska A, Henaff M, Mathieu M, Arous G Ben, LeCun Y, Ioffe S, Szegedy C, Wu S, Zhong S, Liu Y, et al. 2015. Batch normalization: accelerating deep network training by reducing internal covariate shift. Data Min with Decis Trees. doi:10.1007/s13398-014-0173-7.2.

Cioli C, Abdi H, Beaton D, Burnod Y, Mesmoudi S. 2014. Differences in human cortical gene expression match the temporal properties of large-scale functional networks. PLoS One. 9(12):1–28. doi:10.1371/journal.pone.0115913.

Correale S, Esposito C, Pirone L, Vitagliano L, Gaetano S Di, Pedone E. 2013. A biophysical characterization of the folded domains of KCTD12: Insights into interaction with the GABAB2receptor. J Mol Recognit. 26(10):488–495. doi:10.1002/jmr.2291.

Duprat F, Lesage F, Patel a J, Fink M, Romey G, Lazdunski M. 2000. The neuroprotective agent riluzole activates the two P domain K(+) channels TREK-1 and TRAAK. Mol Pharmacol. 57(5):906–12.

Eberwine J, Yeh H, Miyashiro K, Cao Y, Nair S, Finnell R, Zettel M, Coleman P. 1992. Analysis of gene expression in single live neurons. Proc Natl Acad Sci U S A. 89(7):3010–3014. doi:10.1073/pnas.89.7.3010.

Esclapez M, Tillakaratne NJK, Kaufman DL, Tobin AJ, Houser CR. 1994. Comparative localization of two forms of glutamic acid decarboxylase and their mRNAs in rat brain

supports the concept of functional differences between the forms. J Neurosci. 14(3):1834–55. doi:10.1523/jneurosci.14-03-01834.1994.

Ferguson BR, Gao W-J. 2018. PV interneurons: critical regulators of E/I balance for prefrontal cortex-dependent behavior and psychiatric disorders. Front Neural Circuits. 12:37. doi:10.3389/fncir.2018.00037.

Foss-Feig JH, Adkinson BD, Ji JL, Yang G, Srihari VH, McPartland JC, Krystal JH, Murray JD, Anticevic A. 2017. Searching for cross-diagnostic convergence: neural mechanisms governing excitation and inhibition balance in schizophrenia and autism spectrum disorders. Biol Psychiatry. 81(10):848–861. doi:10.1016/j.biopsych.2017.03.005.

Fuzik J, Zeisel A, Mate Z, Calvigioni D, Yanagawa Y, Szabo G, Linnarsson S, Harkany T. 2016. Integration of electrophysiological recordings with single-cell RNA-seq data identifies novel neuronal subtypes. Nat Biotechnol. 34(2):175–183. doi:10.1038/nbt.3443.

Gannon RL, Millan MJ. 2011. Positive and negative modulation of circadian activity rhythms by mGluR5 and mGluR2/3 metabotropic glutamate receptors. Neuropharmacology. 60(2–3):209–215. doi:10.1016/j.neuropharm.2010.08.022.

Ge J, Barnes NM. 1996. Alterations in angiotensin AT1 and AT2 receptor subtype levels in brain regions from patients with neurodegenerative disorders. Eur J Pharmacol. 297(3):299–306. doi:10.1016/0014-2999(95)00762-8.

Van Gelder RN, von Zastrow ME, Yool a, Dement WC, Barchas JD, Eberwine JH. 1990. Amplified RNA synthesized from limited quantities of heterogeneous cDNA. Proc Natl Acad Sci U S A. 87(5):1663–1667. doi:10.1073/pnas.87.5.1663.

Goldman-Rakic PS. 1995. Cellular basis of working memory. Neuron. 14(3):477–485. doi:10.1016/0896-6273(95)90304-6.

Hanisch UK. 2002. Microglia as a source and target of cytokines. Glia. 40(2):140–55. doi:10.1002/glia.10161.

Hill E. 2001. Cellular Diversity in Mouse Neocortex Revealed by Multispectral Analysis of Amino Acid Immunoreactivity. Cereb Cortex. 11(8):679–90. doi:10.1093/cercor/11.8.679.

Hu H, Gan J, Jonas P. 2014. Fast-spiking, parvalbumin+ GABAergic interneurons: From cellular design to microcircuit function. Science (80-). 345(6196):1255263. doi:10.1126/science.1255263.

van den Hurk M, Erwin JA, Yeo GW, Gage FH, Bardy C. 2018. Patch-Seq Protocol to Analyze the Electrophysiology, Morphology and Transcriptome of Whole Single Neurons Derived From Human Pluripotent Stem Cells. Front Mol Neurosci. 11:261. doi:10.3389/fnmol.2018.00261.

Jiang Z, Wallner M, Meera P, Toro L, Al JET. 1999. Human and rodent MaxiK channel [illegible] -

subunit genes : cloning and characterization. Genomics. 55(1):57–67. doi:10.1006/geno.1998.5627.

Jones EG, Huntley GW, Benson DL. 1994. Alpha calcium/calmodulin-dependent protein kinase II selectively expressed in a subpopulation of excitatory neurons in monkey sensory-motor cortex: Comparison with GAD-67 expression. J Neurosci. 14(2):611–29. doi:10.1523/jneurosci.14-02-00611.1994.

Kann O. 2016. The interneuron energy hypothesis: Implications for brain disease. Neurobiol Dis. 90:75–85. doi:10.1016/j.nbd.2015.08.005.

Kataoka A, Tozaki-Saitoh H, Koga Y, Tsuda M, Inoue K. 2009. Activation of P2X7 receptors induces CCL3 production in microglial cells through transcription factor NFAT. J Neurochem. 108(1):115–25. doi:10.1111/j.1471-4159.2008.05744.x.

Kawaguchi Y, Kubota Y. 1997. GABAergic cell subtypes and their synaptic connections in rat frontal cortex. Cereb Cortex. 7(6):476–86. doi:10.1093/cercor/7.6.476.

Krimer LS, Zaitsev A V, Czanner G, Kröner S, González-burgos G, Povysheva N V, Iyengar S, Lewis DA, Kro S, Gonza G, et al. 2005. Cluster Analysis – Based Physiological Classification and Morphological Properties of Inhibitory Neurons in Layers 2 – 3 of Monkey Dorsolateral Prefrontal Cortex. 94(5):3009–3022. doi:10.1152/jn.00156.2005.

Lehman A, Thouta S, Mancini GMS, Naidu S, van Slegtenhorst M, McWalter K, Person R, Mwenifumbo J, Salvarinova R, Adam S, et al. 2017. Loss-of-function and gain-of-function mutations in KCNQ5 cause intellectual disability or epileptic encephalopathy. Am J Hum Genet. 101(1):65–74. doi:10.1016/j.ajhg.2017.05.016.

Levine JH, Lin Y, Elowitz MB. 2013. Functional roles of pulsing in genetic circuits. Science (80-). 342(6163):1193–1200. doi:10.1126/science.1239999.

Lin Y, Ghazanfar S, Strbenac D, Wang A, Patrick E, Lin DM, Speed T, Yang JYH, Yang P. 2019. Evaluating stably expressed genes in single cells. Gigascience. 8(9). doi:10.1093/gigascience/giz106.

Lin Y, Ghazanfar S, Strbenac D, Wang A, Patrick E, Speed T, Yang J, Yang P. 2017. Housekeeping genes, revisited at the single-cell level. bioRxiv. bioRxiv 22. doi:10.1101/229815.

Liu Y, Beyer A, Aebersold R. 2016. On the Dependency of Cellular Protein Levels on mRNA Abundance. Cell. 165(3):535–50. doi:10.1016/j.cell.2016.03.014.

Maier T, Güell M, Serrano L. 2009. Correlation of mRNA and protein in complex biological samples. FEBS Lett. 583((24)):3966–73. doi:10.1016/j.febslet.2009.10.036.

Martin HGS, Lassalle O, Manzoni OJ. 2016. Differential adulthood onset mGlu 5 signaling

saves prefrontal function in the fragile X mouse. Cereb Cortex. 27(12):5592–5602. doi:10.1093/cercor/bhw328.

Masana M, Santana N, Artigas F, Bortolozzi A. 2013. Dopamine neurotransmission and atypical antipsychotics in prefrontal cortex: a critical review. Curr Top Med Chem. 999(999):1–7. doi:10.2174/15680266112129990068.

McCormick DA, Wang Z, Huguenard J. 1993. Neurotransmitter control of neocortical neuronal activity and excitability. Cereb Cortex. 3(5):387–398. doi:10.1093/cercor/3.5.387.

Mclennan H. 1983. Receptors for the excitatory amino acids in the mammalian central nervous system. Prog Neurobiol. 20(3–4):251–271. doi:10.1016/0301-0082(83)90004-7.

Mencacci NE, Rubio-Agusti I, Zdebik A, Asmus F, Ludtmann MHR, Ryten M, Plagnol V, Hauser AK, Bandres-Ciga S, Bettencourt C, et al. 2015. A missense mutation in KCTD17 causes autosomal dominant myoclonus-dystonia. Am J Hum Genet. 96(6):938–947. doi:10.1016/j.ajhg.2015.04.008.

Müllner D. 2011. Modern hierarchical, agglomerative clustering algorithms. (1973):1–29. doi:10.1109/LSP.2012.2188026.

Munsky B, Neuert G, Van Oudenaarden A. 2012. Using gene expression noise to understand gene regulation. Science (80-). 336(6078):183–7. doi:10.1126/science.1216379.

Paxinos G, Franklin KBJ. 2006. The rat brain in stereotaxic coordinates: hard cover edition. Amsterdam (NL): Elsevier.

Rombo DM, Dias RB, Duarte ST, Ribeiro JA, Lamsa KP, Sebastião AM. 2016. Adenosine A 1 receptor suppresses tonic GABA A receptor currents in hippocampal pyramidal cells and in a defined subpopulation of interneurons. Cereb Cortex. 26(3):1081–1095. doi:10.1093/cercor/bhu288.

Schnee ME, Brown BS. 1998. Selectivity of linopirdine (DuP 996), a neurotransmitter release enhancer, in blocking voltage-dependent and calcium-activated potassium currents in hippocampal neurons. J Pharmacol Exp Ther. 286(2):709–17.

Shepherd GM. 2004. The Synaptic Organization of the Brian. Oxford (UK): Oxford University Press.

Shibasaki K, Suzuki M, Mizuno A, Tominaga M. 2007. Effects of Body Temperature on Neural Activity in the Hippocampus: Regulation of Resting Membrane Potentials by Transient Receptor Potential Vanilloid 4. J Neurosci. 27(7):1566–1575. doi:10.1523/JNEUROSCI.4284-06.2007.

Sims D, Sudbery I, Ilott NE, Heger A, Ponting CP. 2014. Sequencing depth and coverage: Key considerations in genomic analyses. Nat Rev Genet. 15(2):121–132. doi:10.1038/nrg3642.

Soh H, Pant R, LoTurco JJ, Tzingounis A V. 2014. Conditional Deletions of Epilepsy-Associated KCNQ2 and KCNQ3 Channels from Cerebral Cortex Cause Differential Effects on Neuronal Excitability. J Neurosci. 34(15):5311–5321. doi:10.1523/JNEUROSCI.3919-13.2014.

Suter DM, Molina N, Gatfield D, Schneider K, Schibler U, Naef F. 2011. Mammalian genes are transcribed with widely different bursting kinetics. Science (80-). 332(6028):472–4. doi:10.1126/science.1198817.

Tasic B. 2018. Single cell transcriptomics in neuroscience: cell classification and beyond. Curr Opin Neurobiol. 50:242–249. doi:10.1016/j.conb.2018.04.021.

Tasic B, Menon V, Nguyen TNT, Kim TTK, Jarsky T, Yao Z, Levi BB, Gray LT, Sorensen SA, Dolbeare T, et al. 2016. Adult mouse cortical cell taxonomy revealed by single cell

transcriptomics. Nat Neurosci. 19(2):335–346. doi:10.1038/nn.4216.

Tipping ME, Bishop CM. 1999. Mixtures of probabilistic principal component analyzers. Neural Comput. 11(2):443–482. doi:10.1162/089976699300016728.

Tripathy SJ, Toker L, Bomkamp C, Mancarci BO, Belmadani M, Pavlidis P. 2018. Assessing Transcriptome Quality in Patch-Seq Datasets. Front Mol Neurosci. 11. doi:10.3389/fnmol.2018.00363.

Valihrach L, Androvic P, Kubista M. 2018. Platforms for single-cell collection and analysis. Int J Mol Sci. 19(3):22–24. doi:10.3390/ijms19030807.

Vervoort VS, Beachem M a, Edwards PS, Ladd S, Miller KE, de Mollerat X, Clarkson K, DuPont B, Schwartz CE, Stevenson RE, et al. 2002. AGTR2 mutations in X-linked mental retardation. Science (80-). 296(5577):2401–2403. doi:10.1126/science.1072191.

Vigneault É, Poirel O, Riad M, Prud'homme J, Dumas S, Turecki G, Fasano C, Mechawar N, El Mestikawy S. 2015. Distribution of vesicular glutamate transporters in the human brain. Front Neuroanat. 9:23. doi:10.3389/fnana.2015.00023.

Vlot AHC, de Witte WEA, Danhof M, van der Graaf PH, van Westen GJP, de Lange ECM. 2017. Target and tissue selectivity prediction by integrated mechanistic pharmacokinetic-target binding and quantitative structure activity modeling. AAPS J. 20(1):11. doi:10.1208/s12248-017-0172-7.

Wang D, Tawfik VL, Corder G, Low SA, François A, Basbaum AI, Scherrer G. 2018. Functional divergence of delta and mu opioid receptor organization in CNS pain circuits. Neuron. 98(1):90-108.e5. doi:10.1016/j.neuron.2018.03.002.

Wang HS, Brown BS, Mckinnon D, Cohen IS. 2000. Molecular basis for differential sensitivity of KCNQ and I(Ks) channels to the cognitive enhancer XE991. Mol Pharmacol.

57(6):1218–23.

Wright C, Shin JH, Rajpurohit A, Deep-Soboslay A, Collado-Torres L, Brandon NJ, Hyde TM, Kleinman JE, Jaffe AE, Cross AJ, et al. 2017. Altered expression of histamine signaling genes in autism spectrum disorder. Transl Psychiatry. 7(5):e1126. doi:10.1038/tp.2017.87.

Yagoda N, Von Rechenberg M, Zaganjor E, Bauer AJ, Yang WS, Fridman DJ, Wolpaw AJ, Smukste I, Peltier JM, Boniface JJ, et al. 2007. RAS-RAF-MEK-dependent oxidative cell death involving voltage-dependent anion channels. Nature. 447(7146):864–8. doi:10.1038/nature05859.

Zeisel a., Manchado a. BM, Codeluppi S, Lonnerberg P, La Manno G, Jureus A, Marques S, Munguba H, He L, Betsholtz C, et al. 2015. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. Science (80-). 347(6226):1138–1142. doi:10.1126/science.aaa1934.

Zeng H, Sanes JR. 2017. Neuronal cell-type classification: Challenges, opportunities and the path forward. Nat Rev Neurosci. 18(9):530–546. doi:10.1038/nrn.2017.85.

Zhong P, Yan Z. 2011. Differential regulation of the excitability of prefrontal cortical fast-spiking interneurons and pyramidal neurons by serotonin and fluoxetine. PLoS One. 6(2). doi:10.1371/journal.pone.0016970.