Single-nuclei transcriptomics of schizophrenia prefrontal cortex primarily implicates neuronal subtypes

Benjamin C. Reiner[1,3,*], Richard C. Crist[1,3], Lauren M. Stein[1], Andrew E. Weller[1], Glenn A. Doyle[1], Gabriella Arauco-Shapiro[1], Thomas N. Ferraro[2], Matthew R. Hayes[1], and Wade H. Berrettini[1]

[1] Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania. [2] Department of Biomedical Sciences, Cooper Medical School of Rowan University, Camden, NJ. [3] These authors contributed equally.

***Corresponding author**

Benjamin C. Reiner, PhD

125 S. 31st St., room 2208-1

Philadelphia, PA 19104

Office: (215)-573-0278

Email: bcreiner@pennmedicine.upenn.edu

Abstract

Transcriptomic studies of bulk neural tissue homogenates from persons with schizophrenia and controls have identified differentially expressed genes in multiple brain regions. However, the heterogeneous nature prevents identification of relevant cell types. This study analyzed single-nuclei transcriptomics of ~311,000 nuclei from frozen human postmortem dorsolateral prefrontal cortex samples from individuals with schizophrenia (n = 14) and controls (n = 16). 2,846 differential expression events were identified in 2,195 unique genes in 19 of 24 transcriptomically-distinct cell populations. ~97% of differentially expressed genes occurred in five neuronal cell types, with ~63% occurring in a subtype of *PVALB*+ inhibitory neurons and *HTR2C*+ layer V excitatory neurons. Differentially expressed genes were enriched for genes localized to schizophrenia GWAS loci. Cluster-specific changes in canonical pathways, upstream regulators and causal networks were identified. These results expand our knowledge of disrupted gene expression in specific cell types and permit new insight into the pathophysiology of schizophrenia.

Introduction

Schizophrenia is a chronic psychotic illness affecting ~1% of the population worldwide. With typical onset in late adolescence or early adulthood and a chronic unremitting course, schizophrenia is a devastating disease (reviewed[1]). The heritability of schizophrenia has been estimated at 64%[2] and substantial efforts have been dedicated to identifying genetic variants that are associated with vulnerability to developing schizophrenia. The largest genome wide association study (GWAS) identified 108 genomic loci associated with schizophrenia[3], and when combined with an additional dataset, subsequent meta-analysis identified 145 genomic risk loci[4]. Additionally, rare structural variants, such as microdeletions and copy number variants contribute to the heritability of schizophrenia[5-7]. However, these and other studies have only been able to explain a portion of the heritability of schizophrenia, suggesting that gene by

environment interactions may play an important role[8]. Supporting this interpretation is accumulating evidence that demonstrates increased risk for developing schizophrenia is associated with maternal infection and obstetrics complications, early life trauma and social adversity, urbanicity, and substance abuse (reviewed[9]). While these studies have helped elucidate the genetic and environmental components of risk for developing schizophrenia, determining the underlying mechanisms has proven difficult. One approach to identifying the mechanisms that underlie schizophrenia is transcriptomics.

Transcriptomic studies utilizing bulk homogenates of frozen human postmortem brain tissue from persons with schizophrenia and controls have identified differentially expressed genes (DEGs) in the amygdala[10], hippocampus[11, 12], superior temporal gyrus[13], anterior cingulate cortex[14, 15], and dorsolateral prefrontal cortex (dlPFC)[16, 17], with the largest of these studies identifying ~700 DEGs in the dlPFC[17]. However, the heterogenous cellular composition of the bulk homogenates prevents identification of the specific cell types in which relevant genes are differentially regulated and expressed. To address these limitations, prior studies used laser capture microdissection followed with transcriptomic analysis by microarray to elucidate the effect of schizophrenia on the transcriptome of individual neural cell types. Examination of layer III and V pyramidal neurons in the dlPFC identified ~1,400 DEGs[18] in the context of schizophrenia, whereas a study of parvalbumin positive (PVALB+) interneurons in the dlPFC identified ~900 DEGs[19]. A substantial portion of the DEGs identified in these studies were not detected in previous examinations of bulk homogenates of the same brain regions, supporting the need for more targeted approaches[18, 19]. These findings suggest that examination of transcriptome changes associated with schizophrenia at the level of neural cellular subpopulations is necessary to fully appreciate the differences in cell type-specific transcript levels. Laser capture microdissection studies of human postmortem brain tissue are limited by: their ability to examine a small number of cell types in a targeted fashion, rather than the

75 broader cellular landscape of the brain; the relative low throughput of the technique; and the
76 pooling of cells, which loses the variability of the transcriptome between cells and may collapse
77 transcriptomically-distinct cellular subpopulations.
78
79 Recent advances in single nuclei RNA sequencing (snRNAseq) allow for simultaneous
80 transcriptomic profiling of thousands of nuclei, across all neural cell types in frozen human
81 postmortem brain homogenate, with simultaneous indexing of transcripts at the sample,
82 nucleus, and individual transcript level (unique molecular identifier, UMI). The utility of this
83 approach for human postmortem study is supported by evidence suggesting that single cells
84 and their nuclei have similar transcriptomes, with ~98% of transcripts having the same relative
85 levels[20]. snRNAseq has identified cell type-specific transcriptomic changes in human
86 postmortem brain samples from Alzheimer's disease[21], autism[22], multiple sclerosis[23], and major
87 depressive disorder[24].
88
89 In this study, we performed snRNAseq of ~311,000 nuclei from dlPFC of individuals with
90 schizophrenia (n = 14) and controls (n = 16). We chose to examine the dlPFC due to the
91 evidence of dlPFC dysfunction in schizophrenia[25]. We identified 2,195 DEGs in 19 of 24
92 transcriptomically-distinct cell populations. ~97% of the DEGs occurred in five neuronal cell
93 types, with a subtype of PVALB+ inhibitory neurons and 5-Hydroxytryptamine Receptor 2C
94 positive (HTR2C+) layer V excitatory neurons accounting for ~63% of DEGs. The DEGs were
95 enriched for genes associated with schizophrenia GWAS loci and overrepresented in gene
96 ontologies previously associated with the pathophysiology of schizophrenia. Canonical pathway
97 analysis identified cluster-specific alterations in metabolic and cell signaling pathways, and
98 upstream regulator and causal network analysis identified cluster-specific effects of shared DEG
99 regulators. Taken together, the results of this study help elucidate the cell type specific
100 transcriptomic and neurobiological changes that underlie schizophrenia.

101

102 **Results**

103 Single nuclei RNA sequencing and identification of cell types

104 To better understand the effects of schizophrenia on the neural transcriptome, we
105 utilized snRNAseq to profile dlPFC nuclei from frozen human postmortem brain samples from
106 14 individuals with schizophrenia and 16 controls (Supplementary Table 1). Schizophrenia and
107 control groups did not differ significantly in mean age, postmortem interval, or pH
108 (Supplementary Fig. 1). Tissue samples from the participants were processed to generate
109 nuclei suspensions, which were utilized for snRNAseq using version 3.0 of 10x Genomics 3'
110 gene expression assay. After sequencing to an average depth of ~497 million reads per sample,
111 we initially identified a total of 411,171 nuclei with average medians of ~2,895 genes and ~6,588
112 UMI per nucleus (Supplementary Table 2). The number of sequencing reads per sample,
113 number of nuclei per sample, average sequencing reads per nuclei per sample, median genes
114 per nuclei per sample, and median UMI per nuclei per sample was consistent between the
115 schizophrenia and control groups (Supplementary Fig. 2). After quality control measures were
116 applied (Methods), we identified 311,297 high confidence nuclei, 149,943 from control samples
117 and 161,354 from schizophrenia samples. These nuclei were used for all further analyses.

118

119 To identify nuclei from transcriptomically-distinct cellular populations we used the
120 snRNAseq counts for nuclei from all individuals to perform unbiased clustering using Seurat[26]
121 (Methods). 24 distinct cellular clusters were identified (Fig. 1a). Nuclei from the schizophrenia
122 and control groups were found in all clusters (Fig. 1a), with nuclei from individual samples
123 contributing across the range of most clusters (Supplementary Table 3). Clusters were
124 annotated by expression of known markers for specific cellular subtypes and all major neural
125 cell types were identified, including inhibitory neurons (InNeuro), excitatory neurons (ExNeuro),
126 oligodendrocytes (Oligo), oligodendrocyte precursor cells (OPC), astrocytes (Astro), microglia

127 (Micro), and endothelial cells (Endo; Fig. 1a-b and Supplementary Fig. 3a). Whereas the
128 number of nuclei from each major cell type varied by sample, the total number of nuclei from
129 each major cell type was consistent between schizophrenia and control groups (Supplementary
130 Fig. 3b-c and Supplementary Table 3). ExNeuro clusters were further annotated using known
131 markers specific for different cortical layers. Populations of layer II/III *CUX2*+ (ExNeuro4 and
132 ExNeuro5), layer V *HTR2C*+ (ExNeuro1), layer V *SULF2*+ (ExNeuro7), layer V *KCNK2*+
133 (ExNeuro2 and ExNeuro8), layer V/VI *RXFP1*+ (ExNeuro3 and ExNeuro6), and layer VI
134 *NR4A2*+ (ExNeuro9) excitatory neurons were identified (Fig. 1c). Similarly, InNeuro clusters
135 were annotated using genes associated with known subpopulations. Somatostatin positive
136 (*SST*+, InNeuro1 and InNeuro7), vasoactive intestinal peptide positive (*VIP*+, InNeuro6), and
137 *PVALB*+ positive (InNeuro2 and InNeuro3) inhibitory neurons were identified (Fig. 1d).
138
139 Differential expression
140 Cell type specific changes in nuclear transcript levels were compared between the
141 schizophrenia and control groups using the hurdle model, a linear mixed model (LMM), in
142 MAST[27]. A total of 2,846 differential expression events were identified in 2,195 unique DEGs (q-
143 value < 0.1; expression change ≥10%; Supplementary Table 4). DEGs were detected in 19 of
144 the 24 cellular clusters (Fig. 2a and Supplementary Fig. 4a), with 48% and 51% of the
145 differential expression events occurring in the InNeuro and ExNeuro major cell types,
146 respectively (Fig. 2b. and Supplementary Table 5). No DEGs were detected in the OPC and
147 Micro cell types and only a single unique DEG was detected in the Endo cell type (Fig. 2a-b and
148 Supplementary Tables 4-5). The number of differential expression events per cellular cluster
149 was not related to the number of nuclei per cluster (Supplementary Fig. 4b). Of the unique
150 DEGs, ~97% (2,135/2,195) were detected in five neuronal cell clusters (referred to hereafter
151 collectively as the primary neuronal clusters), with ~63% (1,382/2,195) occurring in just two
152 neuronal cell types, the *PVALB*+ InNeuro2 and *HTR2C*+ ExNeuro1 cellular clusters (Fig. 2c).

153 No DEG was differentially expressed in more than five cellular clusters (Supplementary Fig. 4c
154 and Supplementary Table 4). The majority of the differential expression events (~66%) were
155 down regulatory in schizophrenia compared to controls (Fig. 2a-b, Supplementary Fig. 4a and
156 Supplementary Tables 4-5).
157
158 DEGs associated with GWAS loci
159 The most recent schizophrenia GWAS meta-analysis identified 145 genomic loci
160 associated with schizophrenia[4]. The LMM analyzed a total of 12,239 genes, of which 330 were
161 associated with these GWAS loci. 67 of the 2,195 DEGs identified in this analysis overlapped
162 GWAS loci, representing a total of 44 unique loci (Supplementary Table 6). This number of
163 genes represents an ~13% enrichment over what would be expected by chance
164 (hypergeometric test, $p = 0.029$). DEGs associated with schizophrenia GWAS loci were only
165 present in nine clusters. Cluster-specific hypergeometric tests identified significant
166 overrepresentation of genes related to schizophrenia GWAS loci in DEGs from InNeuro5
167 (~90%), ExNeuro1 (~32%), and InNeuro2 (~23%; Fig. 3a and Supplementary Table 7). These
168 results suggest that some of the DEGs identified in this study may represent the transcriptomic
169 changes that underlie specific schizophrenia GWAS risk loci.
170
171 Gene Ontologies and KEGG pathways
172 To understand the neurobiological consequences of the DEGs in the primary neuronal
173 clusters, DAVID was used to identify gene ontologies (GO; Supplementary Tables 8-22) and
174 KEGG pathways (Supplementary Tables 23-27) in which the DEGs from each cell cluster were
175 over or under represented. The gene ontology analysis identified cluster-specific representation
176 changes of DEGs in ontologies currently believed to be related to the pathophysiology of
177 schizophrenia, such as the biological processes “RNA splicing” and “aerobic respiration” and the
178 cellular compartments “lysosomal membrane” and “mitochondrion”. Similarly, the KEGG

179 pathway analysis identified changes in the representation of DEGs in the KEGG pathways
180 “Spliceosome”, “Oxidative phosphorylation” and “Synaptic vesicle cycle”. Taken together, these
181 results suggest that the cluster-specific DEGs may represent some of the neurobiological
182 changes that underlie schizophrenia-associated neuropathophysiology.
183
184 Canonical Pathways
185 To better identify the metabolic and cell signaling pathways that are likely to be altered in
186 the primary neuronal clusters in the context of schizophrenia, canonical pathway analysis was
187 performed with Ingenuity Pathway Analysis (IPA) using the DEGs occurring in each of the
188 primary neuronal clusters. Shared and unique affected canonical pathways were identified for
189 the primary neuronal clusters. The top six shared canonical pathways for the primary neuronal
190 clusters, as determined by IPA activation z-score, a prediction of the direction and magnitude of
191 dysregulation, are presented in Fig. 3b and the top six canonical pathways from each of the
192 primary neuronal clusters, as determined by -log(p-value), are presented in Supplementary
193 Table 28. Given that the majority of DEGs only occurred in the ExNeuro1 and InNeuro2 cellular
194 clusters, shared canonical pathways for these clusters were also determined (Fig. 3c).
195 Interestingly, whereas some clusters shared predicted directions of effect for some pathways,
196 the DEGs underlying the effect for each cellular cluster were a combination of both shared and
197 unique DEGs (Supplementary Table 28). For example, function of the Oxidative
198 Phosphorylation canonical pathway is predicted to be substantially downregulated in
199 schizophrenia in the ExNeuro1 (z-score = -4.146, p-value = 3.98e-11) and InNeuro2 (z-score = -
200 3.742, p-value = 3.52e-4) cellular clusters (Fig. 3c). Examination of the DEG heat map (Fig. 3d)
201 and signaling pathway (Supplementary Figure 5) for the Oxidative Phosphorylation canonical
202 pathway show a combination of both multi-cluster and cluster-specific DEGs contributing to the
203 overall dysregulation. While the overall results of the canonical pathway analysis suggest that
204 the primary neuronal clusters share alterations in neurobiologically relevant metabolic and cell

205 signaling pathways in the context of schizophrenia, with respect to both direction and magnitude
206 of effect, examination of the underlying changes suggests that it is actually a combination of
207 shared and cluster-specific alterations of the transcriptome that drive overall pathway
208 dysfunction.
209
210 Upstream Regulators of DEGs
211 The majority of DEGs were only detected in the ExNeuro1 or InNeuro2 cellular clusters,
212 suggesting the possibility that there may be cell type-specific upstream regulators of DEGs
213 unique to these cellular clusters. IPA was utilized to identify putative upstream regulators of
214 DEGs. 52 and 87 predicted upstream regulators of DEGs, for which a z-score could be
215 calculated, were identified for the ExNeuro1 cluster and InNeuro2 cluster, respectively
216 (Supplementary Tables 29-30). The top six shared upstream regulators of DEGs, as determined
217 by IPA activation z-score, are presented in Fig. 3e. Predicted upstream regulators with the three
218 highest and lowest activation z-scores for the ExNeuro1 and InNeuro2 cellular clusters are
219 displayed in Fig. 3f-g. Similar to the canonical pathways, upstream regulators predicted to be
220 involved in the function of both the ExNeuro1 and InNeuro2 cellular clusters demonstrate both
221 shared and cell type-specific DEG targets. For example, the function of the transcription factor
222 *NFE2L2*, was predicted to be substantially downregulated in both the ExNeuro1 (z-score = -
223 3.747, p-value = 2.54e-2) and InNeuro2 (z-score = -2.659, p-value = 7.64e-3) cell clusters (Fig.
224 3e-g and Supplementary Tables 29-30), with 9 shared DEGs out of the 23 and 32 downstream
225 DEG targets, respectively (Supplementary Fig. 6a-b). These results suggest that the shared
226 predicted upstream regulators identified in this analysis may act in a cell cluster-specific fashion
227 to regulate the DEGs. Supporting the validity of the upstream regulator analysis, 8 DEGs in the
228 ExNeuro1 cluster and 13 DEGs in the InNeuro2 cluster were predicted to be upstream
229 regulators of other DEGs in their respective clusters. All 21 upstream DEGs had altered
230 transcription in the predicted direction of effect (Supplementary Table 31). In total, these results

231 suggest that the predicted upstream regulators identified in this analysis may play a role in
232 regulating cluster-specific DEGs in the context of schizophrenia.
233
234 Causal Networks of DEG regulation
235 Knowing that regulation of DEGs in the ExNeuro1 and InNeuro2 cellular clusters would
236 not be limited to only direct upstream regulators, we utilized IPA to predict upstream causal
237 networks that putatively regulate the DEGs. The analysis identified 788 and 622 predicted
238 causal networks that regulate the DEGs in the ExNeuro1 and InNeuro2 cellular clusters,
239 respectively (Supplementary Tables 32-33). The top six shared causal network master
240 regulators, as determined by IPA activation z-score, are presented in Fig. 4a. Predicted causal
241 network master regulators with the three highest and lowest activation z-scores for the
242 ExNeuro1 and InNeuro2 cellular clusters are displayed in Fig. 4b-c. Similar to the upstream
243 regulators, there are both shared and cluster-specific DEG targets associated with causal
244 network master regulators predicted to be involved in the ExNeuro1 and InNeuro2 cellular
245 clusters (Supplementary Tables 32-33). The validity of causal network analysis is supported by
246 the prediction of 32 DEGs in the ExNeuro1 cluster and 30 DEGs in the InNeuro2 cluster as
247 causal network master regulators (Supplementary Tables 32-33). For example, in the ExNeuro1
248 cluster, *ST3GAL5* was predicted to be an upregulated causal network master regulator that
249 acted through four known regulatory targets and caused the observed downregulation of 25
250 DEGs (Fig. 4d). In agreement with this prediction, *ST3GAL5* was upregulated ~19% in the
251 ExNeuro1 cluster. Taken together, these results suggest that at least a portion of the predicted
252 master regulators identified in the causal network analysis may be important regulatory
253 elements that control cell cluster-specific transcriptomic changes in the context of
254 schizophrenia.
255
256 **Discussion**

257 This study examined the transcriptome of ~311,000 single nuclei from the dlPFC of
258 persons with schizophrenia and controls. The median UMI and gene counts per nucleus were
259 approximately double those of three recent human postmortem cortical snRNAseq studies,
260 although the ratio of median UMI count to median gene count was similar[22-24]. Differences
261 between versions 2.0 and 3.0 of the 10x Genomics gene expression assay explain at least a
262 portion of our increased gene and UMI yield. The identification of 24 transcriptomically distinct
263 cellular clusters is consistent with other snRNAseq studies of human postmortem cortex[21, 22, 24]
264 and the number of clusters per major cell type closely approximates those of a recent
265 snRNAseq study of the dlPFC, including the identification of a single *HTR2C*+ cluster of
266 pyramidal neurons and two distinct *PVALB*+ clusters of inhibitory neurons[24]. 2,195 DEGs were
267 detected in 19 of 24 cellular clusters, with ~97% of DEGs occurring in five neuronal cell types
268 and ~63% occurring in either a cluster of *PVALB*+ inhibitory neurons (InNeuro2), a cluster of
269 *HTR2C*+ pyramidal neurons (ExNeuro1), or both. Prior evidence suggests that GWAS loci and
270 gene sets associated with schizophrenia are related to a limited subset of neurons, including
271 PVALB+ inhibitory neurons and glutamatergic pyramidal neurons[28]. Similarly, GWAS loci-related
272 genes were overrepresented in the DEGs of the *PVALB*+ InNeuro2 and *HTR2C*+ ExNeuro1
273 clusters, as well as the *CALB1*+, *NOS1*+ InNeuro5 cluster. Taken together, these results
274 support the hypothesis that the common genetic variants associated with schizophrenia are
275 relevant in limited sets of neuronal cell types and schizophrenia-related transcriptome
276 alterations are primarily limited to these cells. The observation that no DEG occurred in more
277 than five cellular clusters and that no DEG was present in all the clusters of any multi-clustered
278 major cell type underscores the importance of utilizing single nuclei/single cell approaches to
279 neural transcriptomics.
280
281 Analysis of the DEGs in the primary neuronal clusters identified enrichment of genes
282 associated with a substantial number of GO terms and KEGG pathways, with many aligning to

283 known pathophysiology of schizophrenia. For example, multiple clusters were enriched for
284 DEGs related to RNA splicing. Broad differences in alternative splicing have been identified in
285 brain tissue from individuals with psychiatric disorders, including schizophrenia[13, 29]. Many of
286 these changes have been shown to occur in candidate genes associated with schizophrenia
287 pathology (reviewed[30]) or genes associated with schizophrenia GWAS loci[29]. Additionally, our
288 analyses identified multiple GO terms and KEGG pathways across the primary neuronal
289 clusters related to energy metabolism and oxidative stress (e.g. mitochondria, mitochondrial
290 inner membrane, aerobic respiration, and oxidative phosphorylation). Oxidative stress, an
291 increase in intracellular reactive oxygen species (ROS), is associated with increased risk for
292 developing schizophrenia and may be integral to the pathology of the disease[31] (reviewed[32]).
293 ROS cause both single- and double-strand DNA breaks (reviewed[33]) and oxidative stress is
294 associated with reduced telomere length *in vitro* and *in vivo* (reviewed[34]). Similarly, several
295 studies have found that individuals with schizophrenia have increased levels of DNA damage
296 (reviewed[35]) and shortened telomeres[36-38]. As an example, the InNeuro2 cluster exhibited a
297 downregulation of genes encoding antioxidant proteins, including *TXNRD2*, *PRDX1*, and
298 *PRDX5*, consistent with an increase in oxidative stress. Furthermore, the InNeuro2 cluster was
299 also predicted to have alteration of the mitochondrial dysfunction canonical pathway, which may
300 be potentially caused by ROS and a producer of additional oxidative stress (reviewed[39]). The
301 overlapping cluster-specific nature of the GO term, KEGG pathway and canonical pathway
302 results suggests that the identified mechanisms may represent some of the cell type-specific
303 neural pathophysiology of schizophrenia.
304
305 It is somewhat surprising that the results of the present study primarily highlight the
306 neuronal transcriptome, given that a substantial body of prior research suggests a role for glia in
307 the pathophysiology of schizophrenia (reviewed[40]). Thus, it is important to consider why so few
308 DEGs were identified in the glial populations. Counts of different cell types in human cortical

309 grey matter have indicated that glia are more prevalent than neurons, with a glia to neuron ratio
310 of 1.48:1[41]. In contrast, our results found neuronal nuclei to outnumber glial nuclei (glia-to-
311 neuron ratio 0.79:1). Similar results have been observed in other studies using snRNAseq in
312 human cortex[21, 22, 24], despite the use of varying nuclear isolation techniques, suggesting that the
313 apparent discrepancy may be bioinformatic in nature. Glia have consistently shown lower
314 numbers of UMIs than neurons in human snRNAseq experiments, including the present study[21,
315 22, 24]. Two reports even detail their need to alter initial UMI quality control cutoffs to prevent the
316 exclusion of microglia[21, 24]. While this method was successful in these studies with low median
317 gene and UMI counts, lowering the UMI cutoff ultimately results in the unwanted inclusion of
318 low-quality background nuclei in the dataset. The effect could be either a failure to detect
319 distinct cell populations that would be lost in the "noise" below the cutoff or broadly removing
320 valid nuclei from the low end of all glial populations during application of quality control methods.
321 More complex quality control methods and reduction of background noise in the snRNAseq data
322 may be necessary to address this problem in the future. Furthermore, the reduced detection of
323 glial nuclei in snRNAseq does not discount the glial DEGs that are observed in the present
324 study, as these transcriptome changes may be highly relevant to glial pathophysiology in the
325 context of schizophrenia. For example, *PTK2* was a unique DEG to the Astro2 cellular cluster
326 and literature evidence demonstrates that *PTK2* is uniquely spliced in astrocytes[42] and that
327 chromosomal deletions and duplications in *PTK2* are associated with increased risk for
328 schizophrenia[43]. Taken together, this suggest that the glial DEGs identified in this study may
329 represent interesting targets for future investigation.
330
331 Several limitations of the present study must be noted. First, the relatively small number
332 of postmortem samples analyzed increases the possibility that the subjects are not
333 representative of the broader schizophrenia population or that variability between individuals is
334 masking consistent smaller differences in expression between cases and controls. Replication

335 in a larger sample will be essential for ensuring the biological significance of these results. As
336 larger numbers of cases and controls are added, it is also likely that additional clusters of nuclei
337 will be defined. This improved granularity may permit the identification of additional cellular
338 subpopulations that are disrupted in schizophrenia. Second, schizophrenia patients frequently
339 have comorbidities (e.g. nicotine use disorder and obesity) that are less common in control
340 individuals, presenting analytical confounds. Similarly, schizophrenia patients usually have a
341 history of chronic antipsychotic treatment, whereas controls do not. Thus, it is impossible to
342 know at present whether any of the identified DEGs reflect causality or response to chronic
343 pharmacotherapy. This difficulty is compounded by the use of different antipsychotics with
344 different pharmacologic profiles. It may be possible to partially address this difficult confound
345 through parallel experiments in animal models treated chronically with antipsychotics. It may
346 also be possible to address this issue by studying postmortem brains of persons with
347 schizophrenia who never had antipsychotic treatment. However, at least in the United States,
348 these patients are uncommon. Third, transcriptome-based methods such as snRNAseq have
349 the potential to miss relevant genes that are regulated primarily at the level of translation, since
350 these genes may not display notable mRNA differences between cases and controls. Our use of
351 a 3' gene expression assay also limits the ability to interrogate splicing changes, which may also
352 help to shape transcriptomic architecture and be relevant to schizophrenia pathology. Finally,
353 this project was limited to a single brain region from individuals over 18 years of age. Therefore,
354 spatial and temporal changes in gene expression occurring over the course of the disease
355 would not be identified in our analysis. Different cortical and sub-cortical regions are likely to
356 have both shared and unique gene expression changes in the context of schizophrenia, all of
357 which may be relevant to the development of new therapies. The findings from this study of
358 dlPFC cannot be extrapolated to other areas of the brain, justifying the need for more
359 comprehensive studies. There is also substantial evidence for neurodevelopmental origins in
360 schizophrenia (reviewed[44, 45]), suggesting that there are likely to be relevant transcriptomic

361 differences between cases and controls before adulthood. In summary, we have begun to
362 characterize transcriptome alterations in schizophrenia at the level of single neural cells and
363 extension of this work may provide a new basis for the development of effective treatment
364 strategies.
365
366 **Online Methods**
367 Brain Samples
368 This study was approved by the University of Pennsylvania Institutional Review Board.
369 Fresh frozen postmortem dorsolateral prefrontal cortex (dlPFC) tissue from individuals with
370 schizophrenia (n = 16) and controls (n =16) were obtained from the Douglas-Bell Canada Brain
371 Bank at McGill University, the Human Brain and Spinal Fluid Resource Center at UCLA and the
372 New South Wales Brain Tissue Resource Center. Schizophrenia cases were individuals who
373 were clinically diagnosed with schizophrenia using DSM-IV criteria and controls were individuals
374 without history of psychiatric disease who died of acute non-central nervous system-related
375 reasons. All reported age, sex, ethnicity, postmortem interval, and prefrontal cortex pH data are
376 based on associated medical records. Gray matter samples from the dlPFC were dissected by
377 trained neuroanatomists at their respective brain banks. 14 schizophrenia cases and 16 controls
378 were used in differential gene expression analyses (see below).
379
380 Isolation of Nuclei
381 Nuclei were isolated from frozen postmortem dlPFC (~30mg) using a modified version of
382 a previously described protocol[24]. Briefly, tissue was dounce homogenized in 1 mL lysis buffer
383 (Nuclease-free water with 10 mM Tris-HCL, 10 mM NaCl, 3 mM $MgCl_2$, 0.5% NP-40) using 20
384 strokes of the loose pestle and 15 strokes of the tight pestle. The homogenate was diluted to 5

385 mL with lysis buffer and left to lyse on ice for an additional 5 minutes, followed by addition of an
386 equal volume of wash buffer (1X PBS with 2% BSA, 1:1000 RNase inhibitor, 0.25% Glycerol).
387 Samples were then passed through a 30 µm cell strainer and centrifuged at 500 x g for 5
388 minutes at 4°C. Supernatants were decanted and pellets were resuspended in 10 mL wash
389 buffer. Samples were passed through a 30 µm cell strainer for a second time and centrifuged at
390 500 x g for 7 minutes at 4°C. Supernatants were decanted and pellets were resuspended in 5
391 mL wash buffer and centrifuged at 500 x g for 10 minutes at 4°C. Supernatants were decanted,
392 pellets were resuspended in 500 µL wash buffer, and mixed with 500 µL of 50% iodixanol
393 solution (Optiprep), to make a 25% iodixanol solution containing the washed nuclei. The 1 mL
394 25% iodixanol nuclei solution was layered on 1 mL of a 29% iodixanol cushion and centrifuged
395 at 10,000 x g for 30 minutes at 4°C to pellet nuclei. Finally, nuclei were counted by the Trypan
396 Blue method using a hemocytometer and resuspended in wash buffer at a concentration of
397 1,000 nuclei/µL.

398

399 <u>10x Library Preparation, Sequencing, and Initial Quality Control</u>

400 Microfluidics capture and sequencing library preparation was performed with the 10x
401 Genomics Chromium Single Cell 3' GEM, Library and Gel Bead Kit v3.0 at the Children's
402 Hospital of Philadelphia Center for Applied Genomics per manufacturer's instructions. Target
403 capture was ~10,000 nuclei per sample. Libraries were sequenced in pools of 8 on Illumina
404 NovaSeq 6000 S2 flow cells. Pools contained 4 schizophrenia samples and 4 control samples
405 to minimize any batch effects. CellRanger version 3.1 was used to align reads to the hg38 pre-
406 mRNA transcriptome and make initial cell calls. Filtered read count matrices for all subjects and
407 nuclei were merged into a single Seurat object for subsequent quality control and clustering
408 using Seurat version 3.1 (Supplementary Table 34; Initial Seurat File). Mitochondrial transcripts
409 were removed from the dataset. For initial quality control assessment, the distributions of the

410 numbers of genes and UMIs were determined. Nuclei with the lowest 1% of genes (< 462
411 genes) were removed, as they were unlikely to be informative in downstream analyses.
412 Likewise, nuclei in the top 1% of UMI count (UMI > 63,213) were removed to reduce the
413 presence of multiplets in downstream analyses (Supplementary Table 34; nFeature & nUMI
414 Cut).

415

416 Calculation of PCs, Clustering

417 Transcript counts were normalized to 10,000 counts per subject and scaled. Variably
418 expressed genes were identified with the FindVariableFeatures function in Seurat using the
419 mean.var.plot selection method and analyzing only genes with mean scaled expression
420 between 0.003 and 2. These parameters identified 2,215 highly variable genes, which were
421 used to generate principal components (PCs). Clustering was performed in Seurat using the first
422 50 PCs. Initial clustering was performed at a resolution of 0.25 and generated 31 clusters. Five
423 cell populations with low mean UMIs were removed, along with two schizophrenia cases who
424 did not cluster with the other samples (Supplementary Table 29; Low UMI Clusters & Outlier
425 Subjects Removed). The remaining nuclei were reclustered at a higher resolution (0.40) and
426 cell populations with >85% of nuclei coming from ≤2 subjects were removed, resulting in 27
427 clusters (Supplementary Table 29; Clusters From Few Subjects Removed).

428

429 Annotation of Cell Clusters

430 Major cell types were identified using known cell type markers and methods described in
431 Nagy et al.[24]: Microglia: - *SP1, MRC1, CX3CR1*; Endothelial – *CLDN5*; Astrocytes – *GLUL,*
432 *SOX9, AQP4, GJA1, NDRG2, GFAP, ALDH1A1, ALDH1L1*; Oligodendrocyte Precursor Cells –
433 *PDGFRA, PCDH15, OLIG1, OLIG2*; Oligodendrocytes – *PLP1, MAG, MOG, MOBP, MBP*;

434 Excitatory Neurons – *SATB2, SLC17A7*; Neurons - *SNAP25, STMN2, RBFOX3*; Inhibitory
435 Neurons – *GAD1, GAD2, SLC32A1*. Three clusters with mixed cell type markers were removed
436 for a final total of 24 clusters (Supplementary Table 29; Mixed Marker Clusters Removed). The
437 cortical layer of excitatory neuron clusters was identified using previously described layer
438 markers[24]: Layer II: *GLRA3*; Layer II/III: *CUX2, CARTPT, LAMP5*; Layer V: *KCNK2, SULF2,*
439 *PCP4, HTR2C, FEZF2*; Layer V/VI: *RXFP1, TOX, ETV1, RPRM, FOXP2*; Layer VI: *NR4A2,*
440 *SYNPR, NTNG2*. Similarly, inhibitory neuron cell types were identified using previously
441 described markers[24]: *GAD1, GAD2, CCK, CALB1, CALB2, PVALB, SST, RELN, VIP, NOS1*.

442

443 Differential Gene Expression Analysis

444 Count data were normalized to one million reads, extracted from Seurat, and converted
445 to log2 counts per million (cpm). Metadata and cpm were merged to form a SingleCellAssay
446 object for each cluster. Genes expressed in <20% of the nuclei in a cluster were excluded from
447 downstream analyses. Differential expression analysis between cases and controls was
448 performed using the *MAST* R package[27] by fitting the following linear mixed model:

449

```
450 m <- zlm(~casestatus + gdr + age + sex + batch + (1|subject), sca, parallel = TRUE, method =
451 "glmer", ebayes = FALSE, silent=TRUE)
```

452

453 Case status, age, sex, the capture and sequencing batch, and the number of genes detected in
454 each nucleus (gdr) were included as fixed effects. Subject was included as a random effect to
455 account for correlations between the nuclei coming from a single person. To optimize the
456 random and fixed effects coefficients in the penalized iteratively reweighted least squares step,

457 the integer scalar in the *LME4* R package was set equal to zero, as previously described[22]
458 (https://github.com/DmitryVel/KriegsteinLab/blob/master/snRNAseq_DGE.R). Likelihood Ratio
459 Test was performed in *MAST* to test for differences between the model with and without
460 schizophrenia case status, identifying gene expression differences associated with
461 schizophrenia. Differentially expressed genes were defined as those that were a) statistically
462 significant after multiple testing correction with FDR = 0.1 and b) had at least a 10% difference
463 in expression between case and controls (log2 fold change ≥0.14).

464

465 Gene Ontology and KEGG Pathways

466 Overrepresentation of DEGs in gene ontologies and KEGG pathways was analyzed
467 using DAVID v6.8 (https://david.ncifcrf.gov/home.jsp). A list of the Ensembl IDs for the DEGs in
468 each of the primary neuronal clusters was used as input. Statistical significance was determined
469 as a DAVID EASE Score < 0.1.

470

471 Ingenuity Pathway Analysis

472 Canonical pathway, upstream regulator and causal network analysis were conducted
473 using Ingenuity Pathway Analysis software[46] (www.ingenuity.com). Each of the five primary
474 neuronal clusters was analyzed independently using the Core Analysis, Analyze Dataset
475 function and a list of DEG Ensembl IDs, *MAST* raw p-values, FDR corrected p-values (i.e. q-
476 values), and log2 fold change was used as input (Supplementary Table 4). The analyses from
477 the five primary neuronal clusters or the ExNeuro1 and InNeuro2 clusters were compared using
478 the Core Analysis, Compare Analyses function. P-values were calculated using a right-tailed
479 Fisher's exact test.

480

481 Availability of Data

482 Raw sequencing data and sample annotations are available at dbGaP accession
483 phsXXXXXX.

484

485 **Acknowledgements**

486 The authors wish to thank those providing support for this work. BCR is supported, in
487 part, by a 2017 NARSAD Young Investigator Grant (#26634) from the Brain and Behavior
488 Research Foundation as the Patrick A. Coffer Investigator, funding for which was generously
489 provided by Ronald and Kathy Chandonais. BCR was supported during part of this work by
490 T32MH014654 (PI is WHB). This work was funded by a National Institutes of Health grant to
491 WHB (R01MH109260). LMS was supported by NIDDK F32 DK118818. MRH was supported by
492 NIDDK R01 DK115762. The authors would also like to thank Rachel Kember for her assistance
493 with R. The authors would also like to thank Jane Gaisinsky for her help with tissue preparation.
494 Additionally, the authors wish to thank all the participants, and their families, who donated their
495 brain tissue to research, for making these studies possible.

496

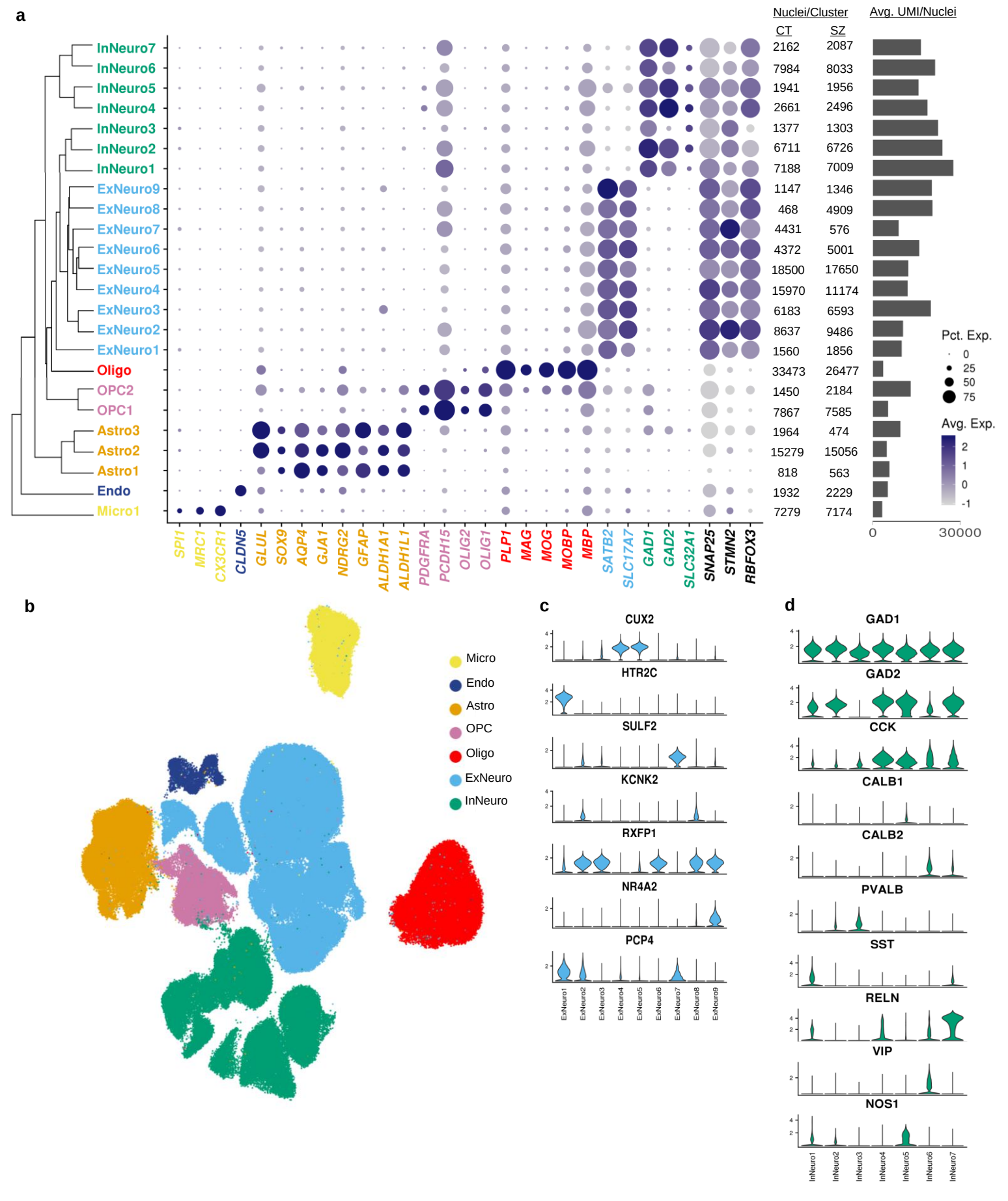
497 **Author Contributions**

498 BCR conceptualized this study and performed experiments. BCR and RCC performed
499 bioinformatics and wrote the manuscript. GA contributed to the performance of experiments and
500 manuscript preparation. LMS, AEW, GAD, TNF and MRH contributed to data interpretation and
501 manuscript preparation. WHB provided oversight, including experimental design, data
502 interpretation and manuscript preparation.

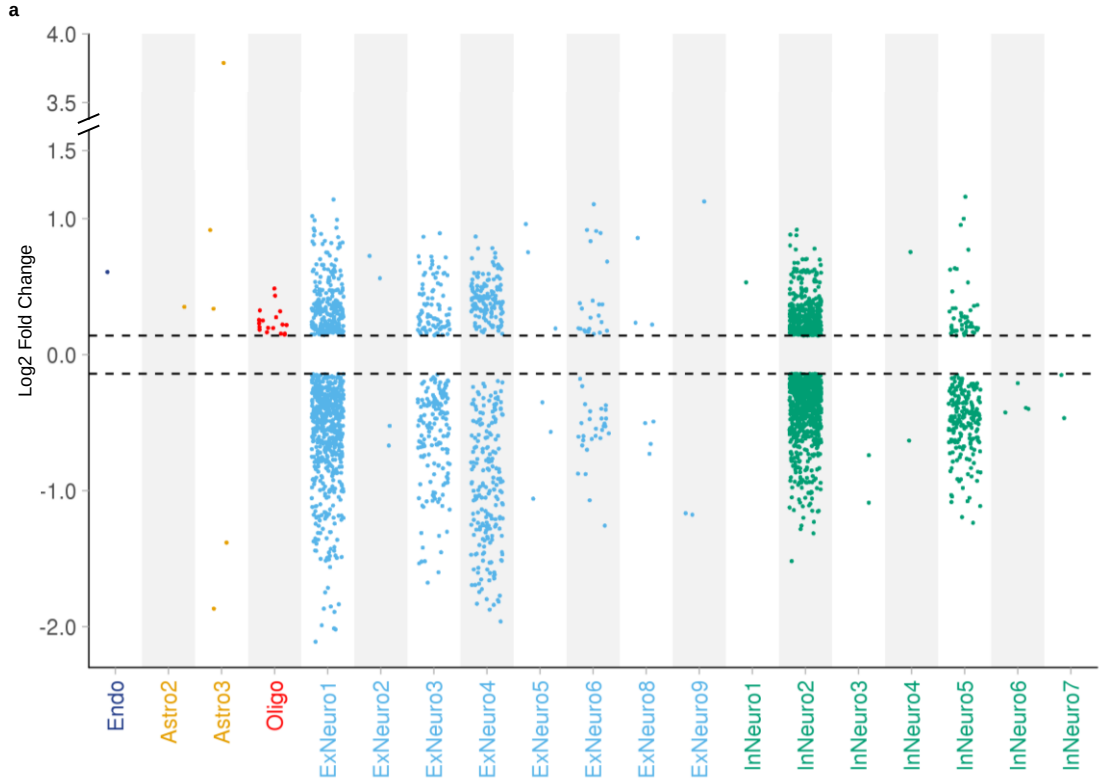
503

504 **Competing Interests Statement**

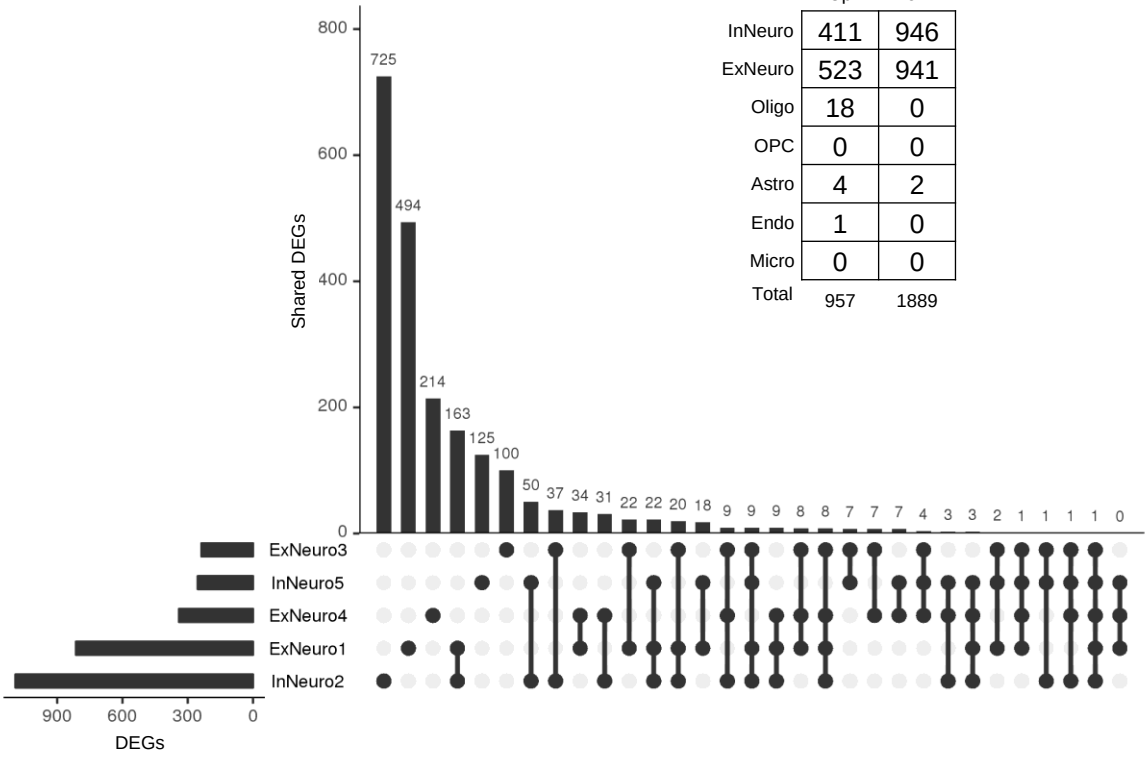
MRH receives separate financial support from Boehringer Ingelheim and Eli Lily & Co that was not used for or related to these studies.


a
Nuclei/Cluster
CT
SZ
Avg. UMI/Nuclei
InNeuro7 2162 2087
InNeuro6 7984 8033
InNeuro5 1941 1956
InNeuro4 2661 2496
InNeuro3 1377 1303
InNeuro2 6711 6726
InNeuro1 7188 7009
ExNeuro9 1147 1346
ExNeuro8 468 4909
ExNeuro7 4431 576
ExNeuro6 4372 5001
ExNeuro5 18500 17650
ExNeuro4 15970 11174
ExNeuro3 6183 6593
ExNeuro2 8637 9486
ExNeuro1 1560 1856
Oligo 33473 26477
OPC2 1450 2184
OPC1 7867 7585
Astro3 1964 474
Astro2 15279 15056
Astro1 818 563
Endo 1932 2229
Micro1 7279 7174
SPI1
MRC1
CX3CR1
CLDN5
GLUL
SOX9
AQP4
GJA1
NDRG2
GFAP
ALDH1A1
ALDH1L1
PDGFRA
PCDH15
OLIG2
OLIG1
PLP1
MAG
MOG
MOBP
MBP
SATB2
SLC17A7
GAD1
GAD2
SLC32A1
SNAP25
STMN2
RBFOX3
Pct. Exp.
0
25
50
75
Avg. Exp.
2
1
0
-1
0
30000
b
Micro
Endo
Astro
OPC
Oligo
ExNeuro
InNeuro
c
CUX2
HTR2C
SULF2
KCNK2
RXFP1
NR4A2
PCP4
ExNeuro1
ExNeuro2
ExNeuro3
ExNeuro4
ExNeuro5
ExNeuro6
ExNeuro7
ExNeuro8
ExNeuro9
d
GAD1
GAD2
CCK
CALB1
CALB2
PVALB
SST
RELN
VIP
NOS1
InNeuro1
InNeuro2
InNeuro3
InNeuro4
InNeuro5
InNeuro6
InNeuro7

651 **Fig. 1 | snRNAseq and clustering. a**, The transcriptome profile of ~311,000 nuclei were
652 utilized for unbiased clustering (dendrogram) and the resulting clusters were annotated with
653 genes known to be markers for major neural cell types (dot plot, cell types and their respective
654 markers are color coded). The size and color of dots is proportional to the percentage of cells
655 expressing the gene (Pct. Exp.) and the average expression level of the gene (Avg. Exp.),
656 respectively. The number columns list number of nuclei from the control (CT) and schizophrenia
657 (SZ) groups that are present in each cluster. The bar graph depicts the average number of
658 unique molecular identifiers (UMI) per nucleus for each cluster. InNeuro, inhibitory neuron;
659 ExNeuro, excitatory neuron; Oligo, oligodendrocyte; OPC, oligodendrocyte precursor cell; Astro,
660 astrocyte; Endo, endothelial cell; Micro, microglia/macrophages. **b,** A uniform manifold
661 approximation and projection (UMAP) dimension reduction plot of all nuclei color coded by
662 major neural cell type. **c,** Excitatory neuron clusters were annotated with known markers of
663 cortical layer II/III (CUX2), layer V (HTR2C, SULF2, KCNK2 and PCP4), layer V/VI (RXFP1),
664 and layer VI (NR4A2) to determine the cortical layer represented by each excitatory neuron
665 cluster. **d,** Inhibitory neuron clusters were annotated using known markers of inhibitory neuron
666 subtypes, including calcium binding proteins (PVALB, CALB1, and CALB2), neuropeptides
667 (SST, VIP, and CCK), and other known markers of cortical inhibitory neurons (GAD1, GAD2,
668 RELN, and NOS1).
669
670
671
672
673
674
675



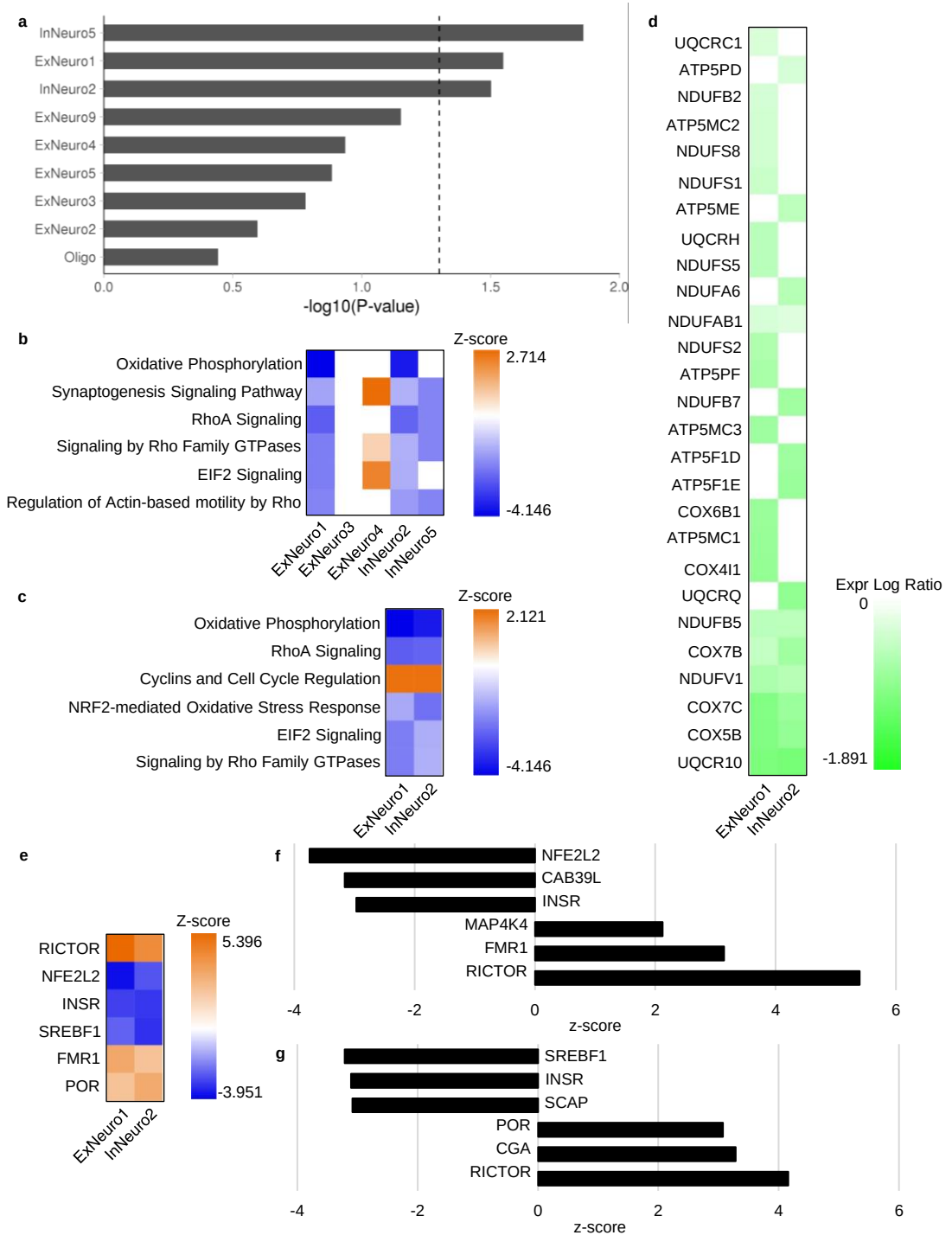
c



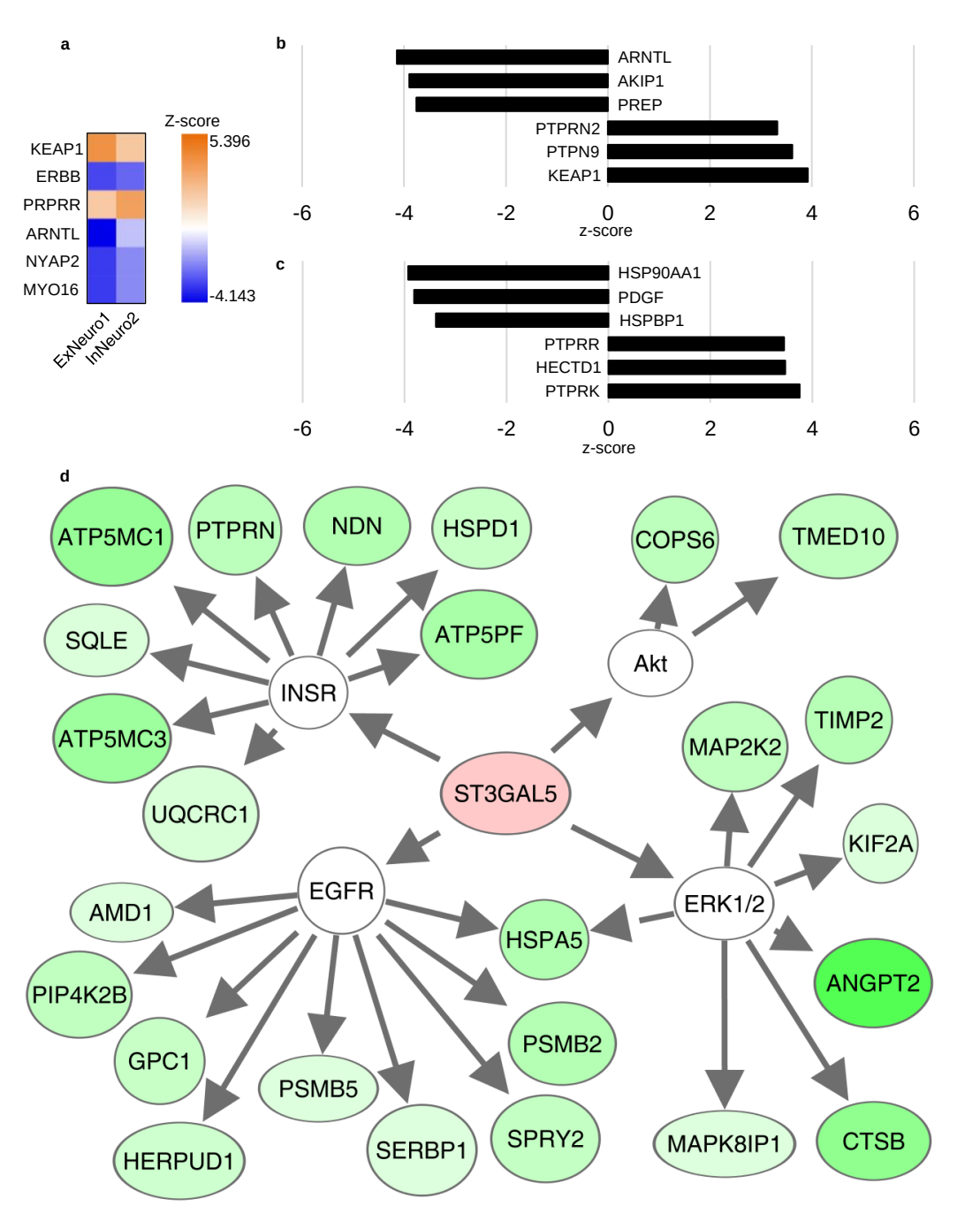
b

	Up	Down
InNeuro	411	946
ExNeuro	523	941
Oligo	18	0
OPC	0	0
Astro	4	2
Endo	1	0
Micro	0	0
Total	957	1889

676 **Fig. 2 | Differential expression. a,** A categorical scatter plot depicting the $\log_2$ fold change of all
677 differential expression events. Dashed lines indicate the minimum bioinformatic cutoff of 10%
678 change in gene expression. This figure is depicted with a continuous y-axis in Supplementary
679 Fig. 4b. **b,** The number of differential expression events per major cell type. **c,** UpSet plot of the
680 number of unique DEGs that are unique or shared between the five neuronal cell cluster that
681 encompass the majority (~97%) of all DEGs.
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700


a
InNeuro5
ExNeuro1
InNeuro2
ExNeuro9
ExNeuro4
ExNeuro5
ExNeuro3
ExNeuro2
Oligo
0.0
0.5
1.0
1.5
2.0
-log10(P-value)
b
Oxidative Phosphorylation
Synaptogenesis Signaling Pathway
RhoA Signaling
Signaling by Rho Family GTPases
EIF2 Signaling
Regulation of Actin-based motility by Rho
ExNeuro1
ExNeuro3
ExNeuro4
InNeuro2
InNeuro5
Z-score
2.714
-4.146
c
Oxidative Phosphorylation
RhoA Signaling
Cyclins and Cell Cycle Regulation
NRF2-mediated Oxidative Stress Response
EIF2 Signaling
Signaling by Rho Family GTPases
ExNeuro1
InNeuro2
Z-score
2.121
-4.146
d
UQCRC1
ATP5PD
NDUFB2
ATP5MC2
NDUFS8
NDUFS1
ATP5ME
UQCRH
NDUFS5
NDUFA6
NDUFAB1
NDUFS2
ATP5PF
NDUFB7
ATP5MC3
ATP5F1D
ATP5F1E
COX6B1
ATP5MC1
COX4I1
UQCRQ
NDUFB5
COX7B
NDUFV1
COX7C
COX5B
UQCR10
ExNeuro1
InNeuro2
Expr Log Ratio
0
-1.891
e
RICTOR
NFE2L2
INSR
SREBF1
FMR1
POR
ExNeuro1
InNeuro2
Z-score
5.396
-3.951
f
NFE2L2
CAB39L
INSR
MAP4K4
FMR1
RICTOR
-4
-2
0
2
4
6
z-score
g
SREBF1
INSR
SCAP
POR
CGA
RICTOR
-4
-2
0
2
4
6
z-score

701 **Fig. 3 | Functional correlations of DEGs. a,** Hypergeometric test p-values for cluster specific
702 overlap between DEGs associated with individual cellular clusters and genes associated with
703 the schizophrenia GWAS loci. **b & c,** Canonical pathways predicted to be disrupted in the (b)
704 primary neuronal clusters and in the (c) ExNeuro1 and InNeuro2 cellular clusters, ranked by
705 absolute magnitude of IPA activation z-score, showing multi-cluster and cluster-specific
706 directions of effect. **d,** DEGs associated with disruption of the oxidative phosphorylation
707 canonical pathway. **e,** Shared predicted upstream regulators of DEGs, ranked as in (c). **f & g,**
708 Predicted upstream regulators in the (f) ExNeuro1 and (g) InNeuro2 cellular clusters with the
709 three highest and lowest predicted activation z-scores.
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725


a
Z-score
5.396
-4.143
KEAP1
ERBB
PRPRR
ARNTL
NYAP2
MYO16
ExNeuro1
InNeuro2
b
ARNTL
AKIP1
PREP
PTPRN2
PTPN9
KEAP1
-6
-4
-2
0
2
4
6
z-score
c
HSP90AA1
PDGF
HSPBP1
PTPRR
HECTD1
PTPRK
-6
-4
-2
0
2
4
6
z-score
d
ATP5MC1
PTPRN
NDN
HSPD1
COPS6
TMED10
SQLE
ATP5PF
Akt
INSR
ATP5MC3
TIMP2
MAP2K2
ST3GAL5
UQCRC1
KIF2A
EGFR
ERK1/2
AMD1
HSPA5
PIP4K2B
ANGPT2
GPC1
PSMB2
PSMB5
SPRY2
MAPK8IP1
CTSB
HERPUD1
SERBP1

726 **Fig. 4 | Causal networks. a,** Shared predicted causal network master regulators of DEGs, in
727 the ExNeuro1 and InNeuro2 cellular clusters, ranked by absolute magnitude of IPA activation z-
728 score. **b & c,** Predicted causal network master regulators of DEGs in the (b) ExNeuro1 and (c)
729 InNeuro2 cellular clusters with the three highest and lowest predicted activation z-scores. **d,** The
730 predicted causal network of ST3GAL5 in ExNeuro1. Based on the 25 downregulated DEGs
731 (green), it was predicted that the function of ST3GAL5 was upregulated (red) and acted through
732 the predicted intermediates (white). ST3GAL5 was upregulated in the ExNeuro1 cluster.
733
734

531 **References**

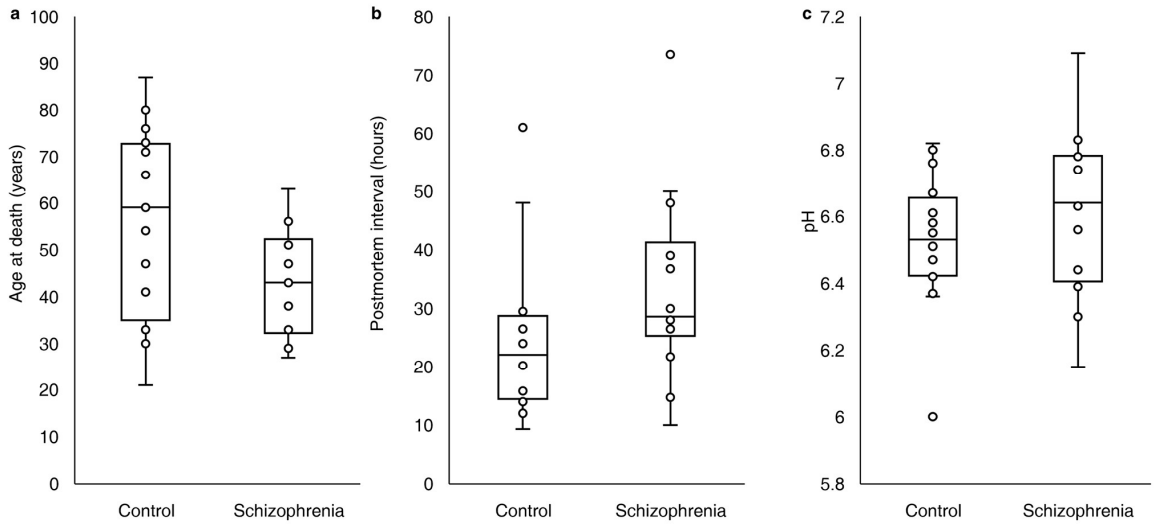
532

533 1. Owen, M.J., Sawa, A. & Mortensen, P.B. Schizophrenia. *Lancet* **388**, 86-97 (2016).
534 2. Lichtenstein, P.*, et al.* Common genetic determinants of schizophrenia and bipolar disorder in
535 Swedish families: a population-based study. *Lancet* **373**, 234-239 (2009).
536 3. Schizophrenia Working Group of the Psychiatric Genomics, C. Biological insights from 108
537 schizophrenia-associated genetic loci. *Nature* **511**, 421-427 (2014).
538 4. Pardinas, A.F.*, et al.* Common schizophrenia alleles are enriched in mutation-intolerant genes
539 and in regions under strong background selection. *Nature genetics* **50**, 381-389 (2018).
540 5. Stefansson, H.*, et al.* Large recurrent microdeletions associated with schizophrenia. *Nature* **455**,
541 232-236 (2008).
542 6. Walsh, T.*, et al.* Rare structural variants disrupt multiple genes in neurodevelopmental pathways
543 in schizophrenia. *Science* **320**, 539-543 (2008).
544 7. Xu, B.*, et al.* Strong association of de novo copy number mutations with sporadic schizophrenia.
545 *Nature genetics* **40**, 880-885 (2008).
546 8. van Os, J., Rutten, B.P. & Poulton, R. Gene-environment interactions in schizophrenia: review of
547 epidemiological findings and future directions. *Schizophr Bull* **34**, 1066-1082 (2008).
548 9. Stilo, S.A. & Murray, R.M. Non-Genetic Factors in Schizophrenia. *Curr Psychiatry Rep* **21**, 100
549 (2019).
550 10. Chang, X.*, et al.* RNA-seq analysis of amygdala tissue reveals characteristic expression profiles in
551 schizophrenia. *Transl Psychiatry* **7**, e1203 (2017).
552 11. Hwang, Y.*, et al.* Gene expression profiling by mRNA sequencing reveals increased expression of
553 immune/inflammation-related genes in the hippocampus of individuals with schizophrenia. *Transl*
554 *Psychiatry* **3**, e321 (2013).
555 12. Kohen, R., Dobra, A., Tracy, J.H. & Haugen, E. Transcriptome profiling of human hippocampus
556 dentate gyrus granule cells in mental illness. *Transl Psychiatry* **4**, e366 (2014).
557 13. Wu, J.Q.*, et al.* Transcriptome sequencing revealed significant alteration of cortical promoter
558 usage and splicing in schizophrenia. *PloS one* **7**, e36351 (2012).
559 14. Zhao, Z.*, et al.* Transcriptome sequencing and genome-wide association analyses reveal
560 lysosomal function and actin cytoskeleton remodeling in schizophrenia and bipolar disorder. *Molecular*
561 *psychiatry* **20**, 563-572 (2015).
562 15. Hong, S., Chen, X., Jin, L. & Xiong, M. Canonical correlation analysis for RNA-seq co-expression
563 networks. *Nucleic acids research* **41**, e95 (2013).
564 16. Fillman, S.G.*, et al.* Increased inflammatory markers identified in the dorsolateral prefrontal
565 cortex of individuals with schizophrenia. *Molecular psychiatry* **18**, 206-214 (2013).
566 17. Fromer, M.*, et al.* Gene expression elucidates functional impact of polygenic risk for
567 schizophrenia. *Nat Neurosci* **19**, 1442-1453 (2016).
568 18. Arion, D.*, et al.* Transcriptome Alterations in Prefrontal Pyramidal Cells Distinguish Schizophrenia
569 From Bipolar and Major Depressive Disorders. *Biol Psychiatry* **82**, 594-600 (2017).
570 19. Enwright Iii, J.F.*, et al.* Transcriptome alterations of prefrontal cortical parvalbumin neurons in
571 schizophrenia. *Molecular psychiatry* **23**, 1606-1613 (2018).
572 20. Grindberg, R.V.*, et al.* RNA-sequencing from single nuclei. *Proceedings of the National Academy*
573 *of Sciences of the United States of America* **110**, 19802-19807 (2013).
574 21. Mathys, H.*, et al.* Single-cell transcriptomic analysis of Alzheimer's disease. *Nature* **570**, 332-337
575 (2019).

576 22. Velmeshev, D.*, et al.* Single-cell genomics identifies cell type-specific molecular changes in
577 autism. *Science* **364**, 685-689 (2019).
578 23. Schirmer, L.*, et al.* Neuronal vulnerability and multilineage diversity in multiple sclerosis. *Nature*
579 **573**, 75-82 (2019).
580 24. Nagy, C.*, et al.* Single-nucleus transcriptomics of the prefrontal cortex in major depressive
581 disorder implicates oligodendrocyte precursor cells and excitatory neurons. *Nat Neurosci* **23**, 771-781
582 (2020).
583 25. Hoftman, G.D., Datta, D. & Lewis, D.A. Layer 3 Excitatory and Inhibitory Circuitry in the
584 Prefrontal Cortex: Developmental Trajectories and Alterations in Schizophrenia. *Biol Psychiatry* **81**, 862-
585 873 (2017).
586 26. Stuart, T.*, et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902 e1821
587 (2019).
588 27. Finak, G.*, et al.* MAST: a flexible statistical framework for assessing transcriptional changes and
589 characterizing heterogeneity in single-cell RNA sequencing data. *Genome biology* **16**, 278 (2015).
590 28. Skene, N.G.*, et al.* Genetic identification of brain cell types underlying schizophrenia. *Nature*
591 *genetics* **50**, 825-833 (2018).
592 29. Gandal, M.J.*, et al.* Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and
593 bipolar disorder. *Science* **362** (2018).
594 30. Glatt, S.J., Cohen, O.S., Faraone, S.V. & Tsuang, M.T. Dysfunctional gene splicing as a potential
595 contributor to neuropsychiatric disorders. *American journal of medical genetics. Part B, Neuropsychiatric*
596 *genetics : the official publication of the International Society of Psychiatric Genetics* **156B**, 382-392
597 (2011).
598 31. Flatow, J., Buckley, P. & Miller, B.J. Meta-analysis of oxidative stress in schizophrenia. *Biol*
599 *Psychiatry* **74**, 400-409 (2013).
600 32. Boskovic, M., Vovk, T., Kores Plesnicar, B. & Grabnar, I. Oxidative stress in schizophrenia. *Curr*
601 *Neuropharmacol* **9**, 301-312 (2011).
602 33. Liou, G.Y. & Storz, P. Reactive oxygen species in cancer. *Free Radic Res* **44**, 479-496 (2010).
603 34. Barnes, R.P., Fouquerel, E. & Opresko, P.L. The impact of oxidative DNA damage and stress on
604 telomere homeostasis. *Mech Ageing Dev* **177**, 37-45 (2019).
605 35. Markkanen, E., Meyer, U. & Dianov, G.L. DNA Damage and Repair in Schizophrenia and Autism:
606 Implications for Cancer Comorbidity and Beyond. *Int J Mol Sci* **17** (2016).
607 36. Russo, P.*, et al.* Shorter telomere length in schizophrenia: Evidence from a real-world population
608 and meta-analysis of most recent literature. *Schizophr Res* **202**, 37-45 (2018).
609 37. Galletly, C.*, et al.* Shorter telomere length in people with schizophrenia: A preliminary study
610 from Australia. *Schizophr Res* **190**, 46-51 (2017).
611 38. Polho, G.B., De-Paula, V.J., Cardillo, G., dos Santos, B. & Kerr, D.S. Leukocyte telomere length in
612 patients with schizophrenia: A meta-analysis. *Schizophr Res* **165**, 195-200 (2015).
613 39. Andreyev, A.Y., Kushnareva, Y.E. & Starkov, A.A. Mitochondrial metabolism of reactive oxygen
614 species. *Biochemistry (Mosc)* **70**, 200-214 (2005).
615 40. Bernstein, H.G., Steiner, J., Guest, P.C., Dobrowolny, H. & Bogerts, B. Glial cells as key players in
616 schizophrenia pathology: recent insights and concepts of therapy. *Schizophr Res* **161**, 4-18 (2015).
617 41. Azevedo, F.A.*, et al.* Equal numbers of neuronal and nonneuronal cells make the human brain an
618 isometrically scaled-up primate brain. *J Comp Neurol* **513**, 532-541 (2009).
619 42. Zhang, Y.*, et al.* An RNA-sequencing transcriptome and splicing database of glia, neurons, and
620 vascular cells of the cerebral cortex. *J Neurosci* **34**, 11929-11947 (2014).
621 43. International Schizophrenia, C. Rare chromosomal deletions and duplications increase risk of
622 schizophrenia. *Nature* **455**, 237-241 (2008).

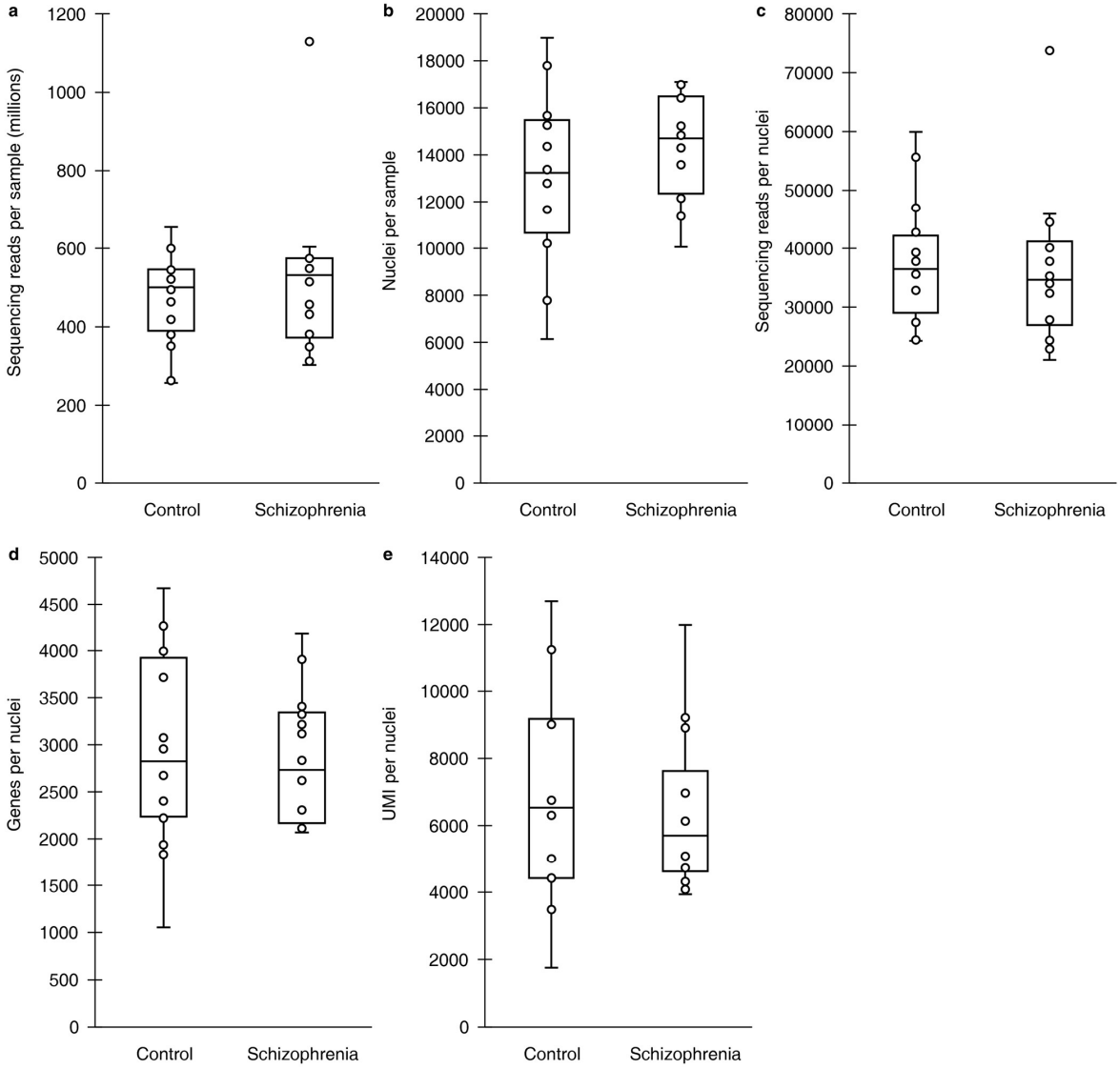
44. Birnbaum, R. & Weinberger, D.R. Genetic insights into the neurodevelopmental origins of schizophrenia. *Nature reviews. Neuroscience* **18**, 727-740 (2017).
45. Rapoport, J.L., Giedd, J.N. & Gogtay, N. Neurodevelopmental model of schizophrenia: update 2012. *Molecular psychiatry* **17**, 1228-1238 (2012).
46. Kramer, A., Green, J., Pollard, J., Jr. & Tugendreich, S. Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics* **30**, 523-530 (2014).

752 **Supplementary Figures**



753
754
755 **Supplementary Figure 1: Postmortem sample characteristics**. This study utilized human
756 postmortem brain samples from 14 individuals with schizophrenia and 16 controls. Brain
757 samples for the schizophrenia and control groups did not differ in (a) age at time of death
758 (control 56.3 ± 19.7; schizophrenia 45.0 ± 14.1; Mann-Whitney U = 65.5, z = 1.91, p = 0.05614)
759 , (b) postmortem interval (PMI; control 24.0 ± 13.3; schizophrenia 33.0 ± 15.5; Mann-Whitney U
760 = 65.0, z = -1.93, p = 0.0536), or (c) prefrontal cortex pH (control 6.5 ± 0.2; schizophrenia 6.6 ±
761 0.2; Mann-Whitney U = 91.5, z = -0.83, p = 0.41). Data presented as Mean ± Standard
762 Deviation. Box and Whisker plots display the median (center line), first and third quartiles (box),
763 and the maximum and minimum values within 1.5x the interquartile distance (whiskers). Circles
764 denote data points in addition to maximum and minimum whiskers.
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782

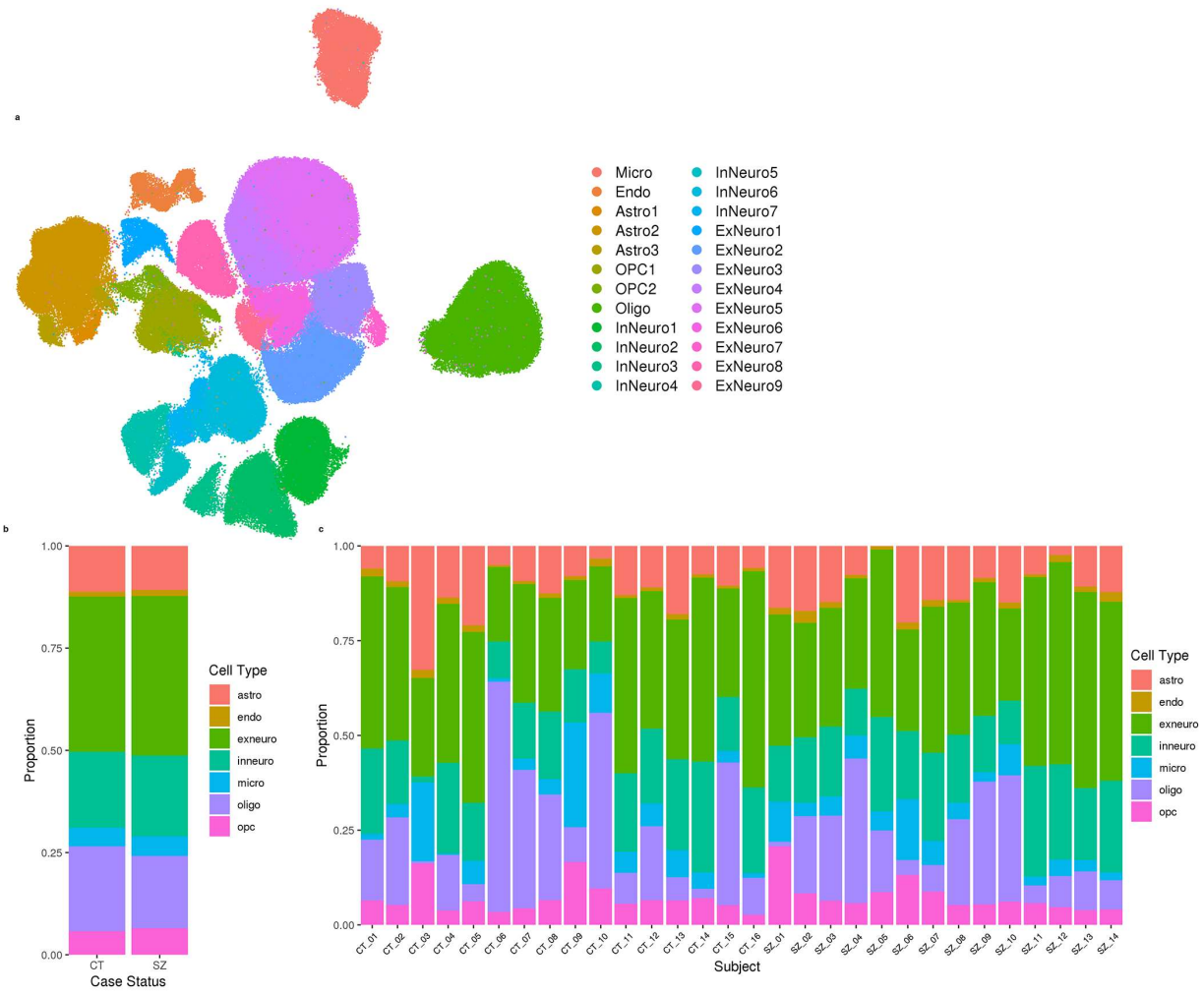
783



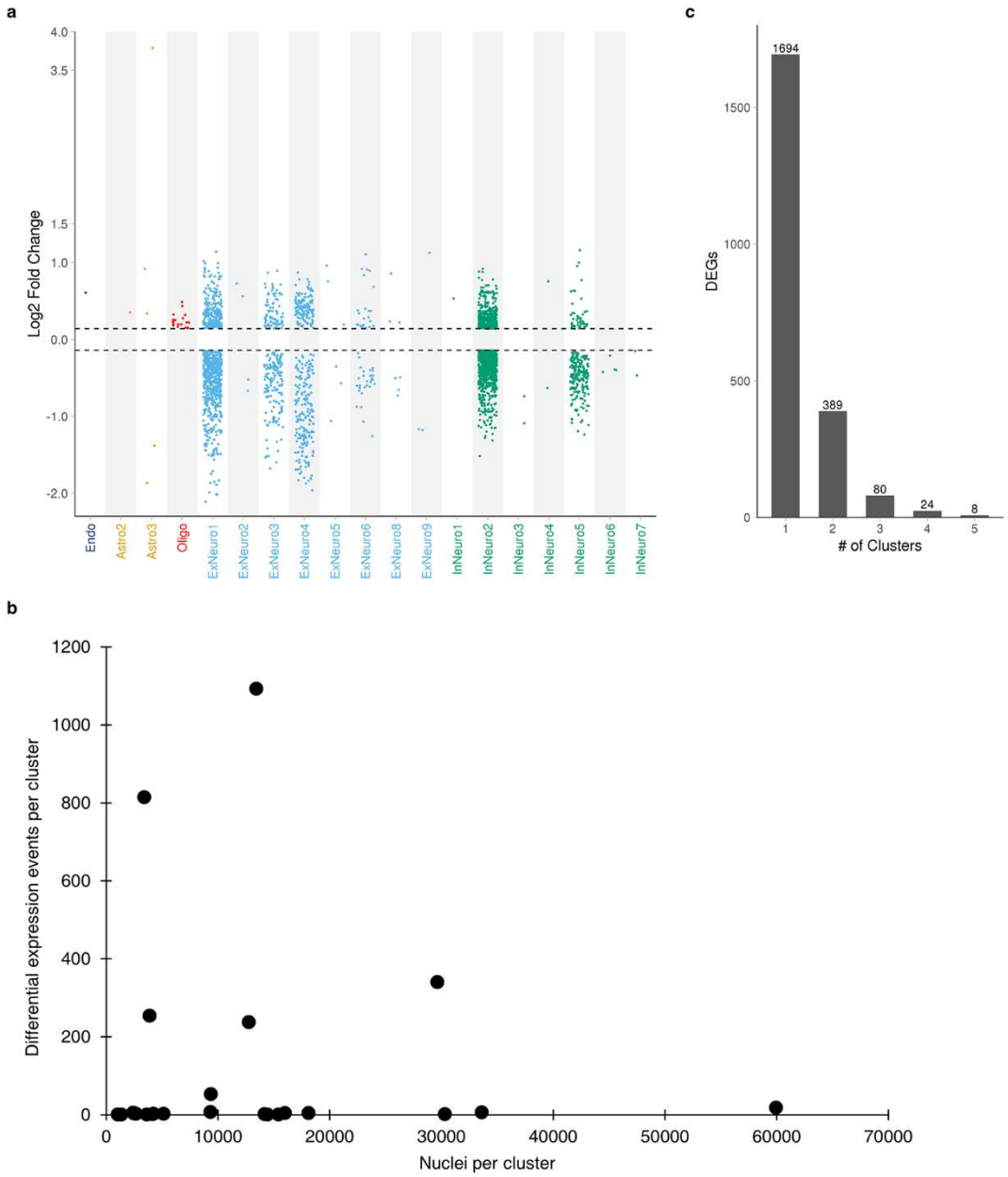
784

785 **Supplementary Figure 2: Sequencing statistics**. snRNAseq was performed on postmortem
786 brain samples from the schizophrenia and control individuals using a the 10x Genomics
787 Chromium platform. Sequencing data was initially analyzed using the 10x Genomics Cell
788 Ranger pipeline (see Methods). (a) The number of sequencing reads per individual did not vary
789 between the schizophrenia and control groups (control 475,224,180.6 ± 112,275,580.0;
790 schizophrenia 520,881,021.7 ± 196,101,237.3; Mann-Whitney U = 99, z = -0.52, p = 0.60306).
791 (b) The number of nuclei identified per sample (control 13,131.1 ± 3,326.8; schizophrenia
792 14,362.4 ± 2,116.2; Mann-Whitney U = 90, z = -0.89, p = 0.37346) and (c) the average number
793 of sequencing reads per nuclei per sample (control 37,571.1 ± 9,837.0; schizophrenia 36,475.6
794 ± 12,646.4; Mann-Whitney U = 98.5, z = 0.54, p = 0.5892) did not differ between the
795 schizophrenia and control groups. One sample from the schizophrenia group had an increased
796 number of total sequencing reads and mean sequencing reads per nuclei (Sample SZ_02;
797 Supplementary Table 2). This is likely the result of the sample being pooled at a higher

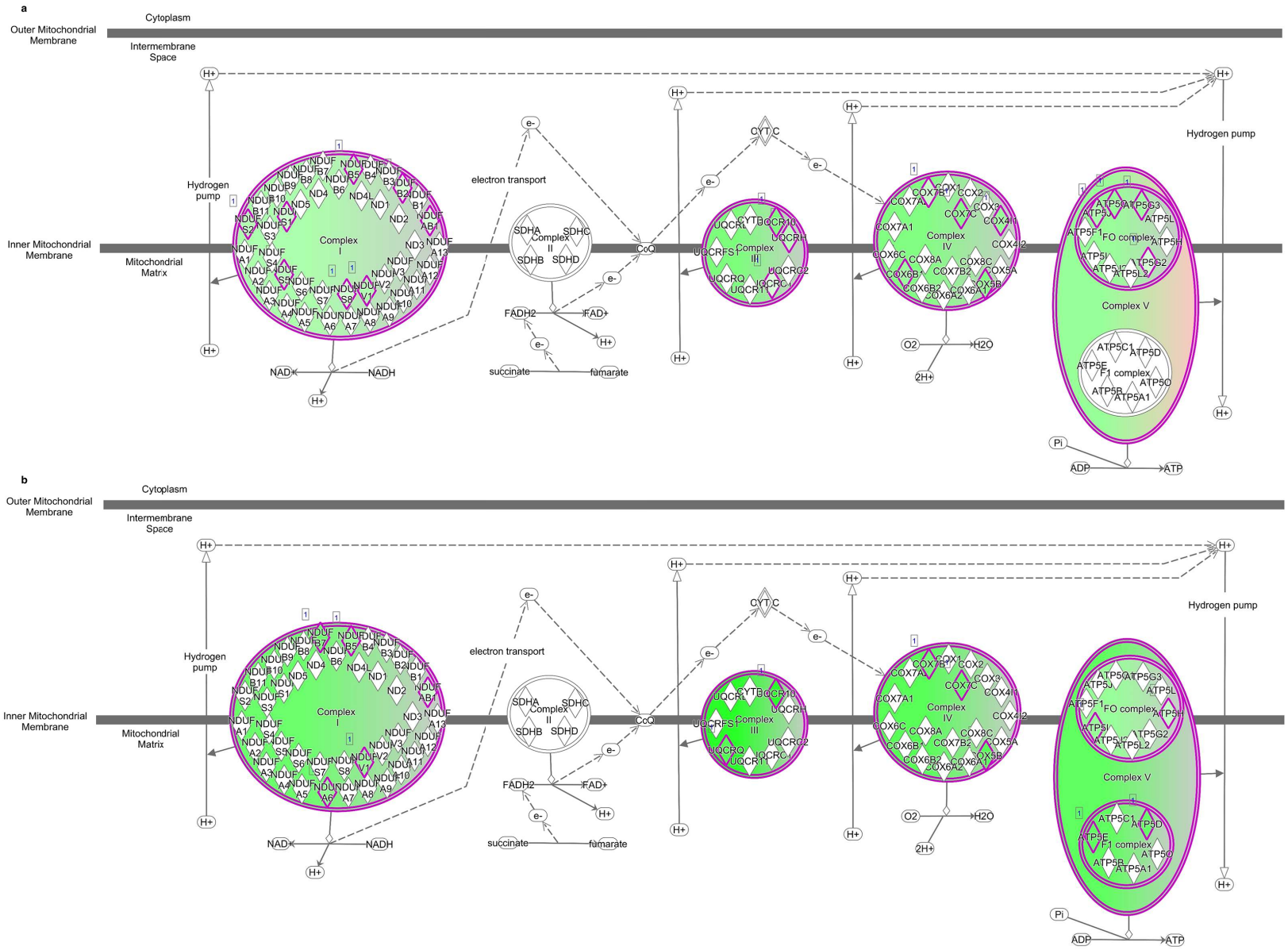
concentration than the other samples with which it sequenced. Given the normalization of the counts per nuclei (see Methods), this is unlikely to negatively affect the results of this study. Nuclei from the schizophrenia and control groups did not differ in (d) the median number of genes detected per nuclei (control 2,923.1 ± 962.0; schizophrenia 2,862.2 ± 662.7; Mann-Whitney U = 107, z = 0.19, p = 0.8493) or (e) the median number of UMI per nuclei (control 6,723.9 ± 3,092.8; schizophrenia 6,432.7 ± 2,237.9; Mann-Whitney U = 110, z = -0.06, p = 0.95216). Data presented as Mean ± Standard Deviation. Box and Whisker plots display the median (center line), first and third quartiles (box), and the maximum and minimum values within 1.5x the interquartile distance (whiskers). Circles denote data points in addition to maximum and minimum whiskers.



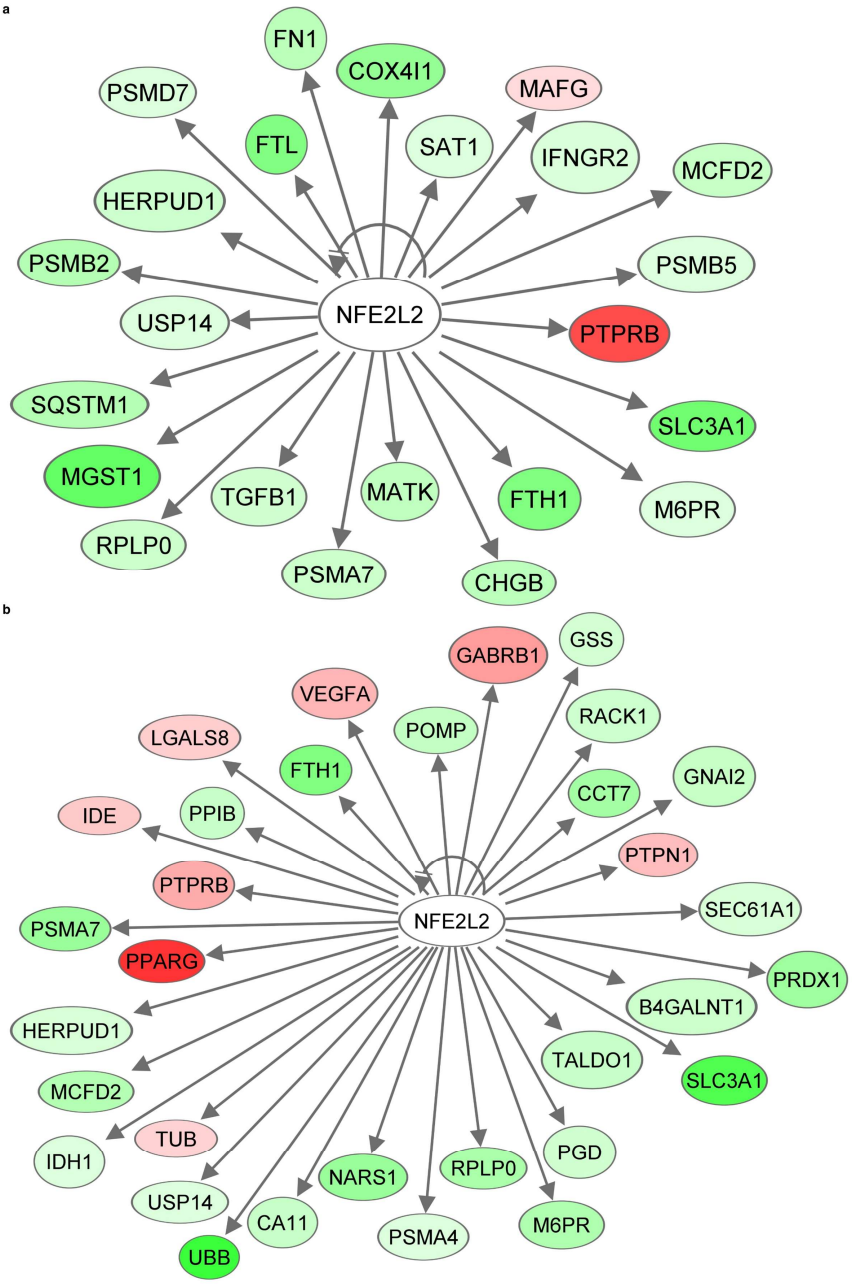
Supplementary Figure 3: Clustering and cell type contributions. (a) A UMAP dimension reduction plot depicting all 24 individual clusters. (b) A stacked bar graph depicting the proportion of nuclei in the control and schizophrenia cohorts that came from each of the major neural cell types. (c) Stacked bar graphs depicting the proportion of nuclei from each of the major cell types from all individuals.



Supplementary Figure 4: Differential expression events. (a) A categorical scatter plot depicting the $\log_2$ fold change of all differential expression events. Dashed lines indicate the minimum bioinformatic cutoff of 10% change in transcript expression. (b) The number of differential expression events detect per cellular cluster was not related to the number of the nuclei in the cluster (ANOVA, $F(1,22) = 0.031$, $p = 0.862$). (c) The number of cellular clusters in which any DEG was found to be differentially expressed.


a
Cytoplasm
Outer Mitochondrial Membrane
Intermembrane Space
Inner Mitochondrial Membrane
Mitochondrial Matrix
Hydrogen pump
Hydrogen pump
electron transport
Complex I
Complex II
Complex III
Complex IV
Complex V
F1 complex
FO complex
CYT C
CoQ
NAD
NADH
FADH2
FAD+
succinate
fumarate
O2
H2O
2H+
Pi
ADP
ATP
b
Cytoplasm
Outer Mitochondrial Membrane
Intermembrane Space
Inner Mitochondrial Membrane
Mitochondrial Matrix
Hydrogen pump
Hydrogen pump
electron transport
Complex I
Complex II
Complex III
Complex IV
Complex V
F1 complex
FO complex
CYT C
CoQ
NAD
NADH
FADH2
FAD+
succinate
fumarate
O2
H2O
2H+
Pi
ADP
ATP

Supplementary Figure 5: Oxidative phosphorylation pathway. Diagrams of the Oxidative Phosphorylation conical signaling pathway in (a) ExNeuro1 and (b) InNeuro2. The large purple outlined mitochondrial complexes with green tinting are predicted to have downregulated function. Mitochondrial complexes I, III, IV, and V are predicted to have downregulated function in both ExNeuro1 and InNeuro2 neuronal clusters. Purple diamonds represent DEGs in each cluster. While similar patterns of overall disruption to the oxidative phosphorylation are predicted in ExNeuro1 and InNeuro2, the underlying DEGs are a combination of shared and unique DEGs.



Supplementary Figure 6: Upstream regulation of DEGs by NFE2L2. The function of NFE2L2 (center) was predicted to be downregulated in both the (a) ExNeuro1 and (b) InNeuro2 cellular clusters. Green and red DEGs indicate observed downregulation and upregulation, respectively. While the level of NFE2L2 transcripts was not altered between the schizophrenia and control groups in the ExNeuro1 and InNeuro2 cellular clusters, these data suggest that function of NFE2L2 was downregulated in the context of schizophrenia, leading to the altered transcription of the target DEGs.