

# Mapping the Cellular Etiology of Schizophrenia and Diverse Brain Phenotypes

Laramie E Duncan<sup>1,2,#</sup>, Tayden Li<sup>3</sup>, Madeleine Salem<sup>4</sup>, Will Li<sup>4</sup>, Leili Mortazavi<sup>5,2</sup>, Hazal Senturk<sup>6</sup>, Naghmeh Shargh<sup>1</sup>, Sam Vesuna<sup>1</sup>, Hanyang Shen<sup>7</sup>, Jong Yoon<sup>1,2,8</sup>, Gordon Wang<sup>1,2</sup>, Jacob Ballon<sup>1</sup>, Longzhi Tan<sup>9,2</sup>, Brandon Scott Pruett<sup>10</sup>, Brian Knutson<sup>5,2</sup>, Karl Deisseroth<sup>11,1,2</sup>, William J Giardino<sup>1,2</sup>

#Corresponding author

<sup>1</sup>Stanford University, Department of Psychiatry and Behavioral Sciences

<sup>2</sup>Stanford University, Wu Tsai Neurosciences Institute

<sup>3</sup>Stanford University, Department of Statistics

<sup>4</sup>Stanford University

<sup>5</sup>Stanford University, Department of Psychology

<sup>6</sup>University of San Francisco, Department of Computer Science

<sup>7</sup>Stanford University, Department of Epidemiology

<sup>8</sup>VA Palo Alto Health Care System

<sup>9</sup>Stanford University, Department of Neurobiology

<sup>10</sup>University of Alabama at Birmingham, School of Medicine

<sup>11</sup>Stanford University, Department of Bioengineering

## Summary paragraph

Psychiatric disorders account for a substantial fraction of the world's disease burden<sup>1</sup>, and yet the development of novel therapeutics has been notoriously slow<sup>2</sup>. Likely contributing factors include the complexity of the human brain and the high polygenicity of psychiatric disorders<sup>3-5</sup>, meaning that thousands of genetic factors contribute to disease risk. Fortunately, technological advances have enabled comprehensive surveys of human brain cell types using transcriptomes from single nuclei (snRNAseq)<sup>6-8</sup>. Additionally, genome-wide association studies (GWAS) have linked thousands of risk loci to psychiatric disorders<sup>9-11</sup>. Here, we combined these two landmark data resources to infer the cell types involved in the etiology of schizophrenia and comparison phenotypes. This work demonstrated: 1) cell types that are concordant with prior findings about schizophrenia, 2) novel cell type associations for schizophrenia, 3) greater molecular specificity regarding schizophrenia-associated cell types than was previously available, 4) evidence that well powered genome-wide and brain-wide datasets are required for these analyses, 5) distinct cellular profiles for five brain-related phenotypes, 6) a prototype for a cell-type based classification system for psychiatric and other brain disorders, and 7) a roadmap toward drug repurposing, novel drug development, and personalized treatment recommendations. Thus, this work formalizes a data-driven, cellular and molecular model of complex brain disorders.

## Introduction

Approximately one in five adults has a psychiatric disorder<sup>12</sup>. While these disorders may resolve on their own or with treatment, for many, they represent lifelong afflictions.

Schizophrenia is often considered the cardinal adult psychiatric disorder given its severity, chronicity, and enduring worldwide distribution<sup>13–15</sup>. Like all other major psychiatric disorders (depression, substance use disorders, PTSD, etc.), the etiology of schizophrenia depends both on genetic and environmental factors. All such disorders are moderately to highly heritable, and schizophrenia is one of the most heritable of these “complex polygenic” disorders, with heritability estimates from twin studies of approximately 80%<sup>16,17</sup>.

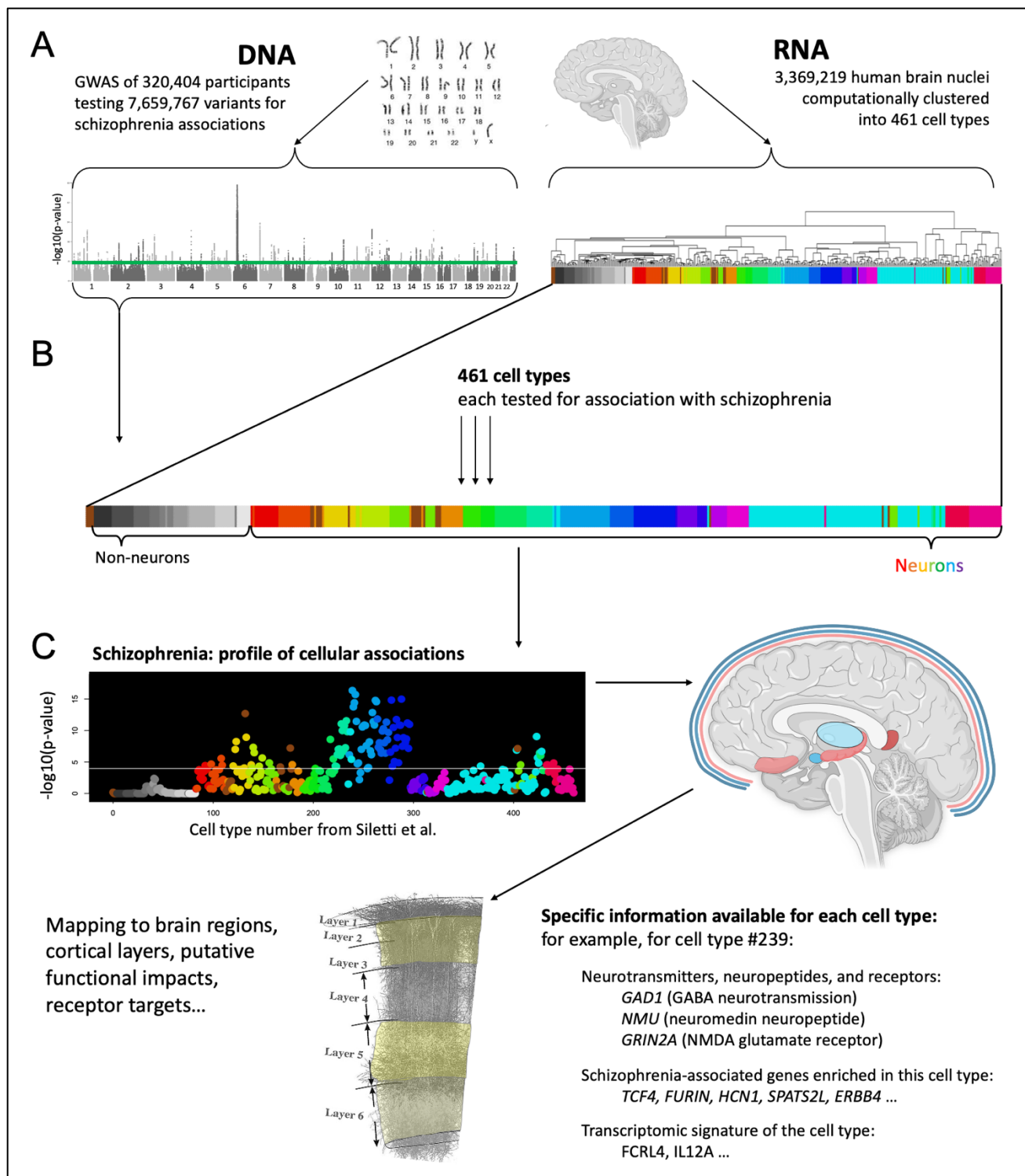
Accordingly, investment in research on schizophrenia has successfully generated knowledge about its genetic basis. Massive human genome-wide association studies (GWAS) encompassing millions of individuals, often led by the international Psychiatric Genomics Consortium (PGC), showed that common genetic variants account for a substantial fraction of population liability to schizophrenia as well as other psychiatric disorders<sup>3,9,18–20</sup>. For instance, the most recent GWAS of schizophrenia in a sample of 320,404 participants found 287 risk loci across the human genome (see **Figure 1A**). These loci exceeded a stringent, international threshold for statistical significance ( $p < 5 \times 10^{-8}$ ), which corrects for tests of multiple hypotheses. While these landmark studies have revealed novel discoveries about psychiatric disorders, they also raised a new challenge: determining the physiological relevance of associated genomic loci.

Polygenic influences on schizophrenia offer a powerful entry point for discovering disease etiology, when combined with data about how genes are used in specific brain cell types. Gene expression from a common nuclear genome is largely what generates cellular diversity in the body, and newly available technology allows measurement of gene expression (RNA molecule counts) from individual cells or the nuclei within those cells via single nucleus RNA sequencing (snRNAseq<sup>i</sup>). Moreover, computational clustering of cellular transcriptomes in snRNAseq datasets readily reveals known as well as novel cell types, and consequently this approach is being used to create comprehensive atlases of cell types in human and animal organs<sup>10,21,22</sup>. Recently, a landmark snRNAseq dataset for the human brain was released<sup>6</sup>. Previous human studies had surveyed only small numbers of brain regions. Thus, Siletti et al.’s<sup>6</sup> analysis of 106 brain regions and 3,369,219 individual nuclei is by far the most comprehensive human snRNAseq dataset available to date. This dataset afforded statistical clustering of 461 cell types in normal human brains, 378 of which are neuronal. The right side of **Figure 1A**, adapted from Siletti et al., shows these 461 cell types as the ends of the leaves of a dendrogram. Using this critical dataset which specifies gene expression in hundreds of healthy, adult human brain cell types, we were able to map the polygenic risk for schizophrenia<sup>9</sup> (as quantified in schizophrenia GWAS) to specific human brain cell types and their anatomical locations. This was accomplished in an entirely data-driven manner using the leading available genome-wide and brain-wide datasets and is therefore “unbiased” with respect to researcher-driven hypotheses about which cell types should be associated with schizophrenia.

---

<sup>i</sup> “scRNAseq” when single cells are used instead of snRNAseq when single nuclei within cells are used.

**Figure 1. Approach for systematically testing 461 human brain cell types for association with schizophrenia.** We tested whether genes associated with schizophrenia were preferentially expressed in one or more brain cell types. **A.** We used two types of human genome-wide data. **Left:** Genome wide association study (GWAS) results from the most recent (2022) schizophrenia GWAS which identified 287 loci for schizophrenia in a sample of 320,404 participants. GWAS results are depicted in a Manhattan plot with the chromosomal position on the x-axis and statistical significance ( $-\log_{10}(p)$ ) on the y-axis, for each genetic variant tested. **Right:** cell type data from the most comprehensive human brain snRNAseq study to date (3,369,219 cells from 106 brain regions, clustered into 461 statistical clusters, referred to here as cell types). For the 461 cell types, colors reflect superclusters as assigned in Siletti et al. Here we used grayscale to denote non-neuronal cell types and rainbow colors for neuronal cell types (see **Figure 2** for complete color coding). **B.** Each of the 461 cell types was tested for association with schizophrenia. As detailed in the methods, we first calculated ‘specificity’ scores using established procedures to quantify the fraction of each gene’s total expression that was found in each cell type (for 16,641 genes, see Methods). Specificity values ranged from 0-1, and for each gene the specificity scores summed to 1 over the 461 cell types (by definition). MAGMA was first used to quantify schizophrenia associations for each gene (as applied to GWAS data). Next, for each of the 461 cell types (analyzed separately), we used regression (implemented in MAGMA) to test for linear relationships between specificity scores and schizophrenia associations of those genes, while correcting for known potential confounding variables. **C.** Results are depicted first as a cell type profiles (here for schizophrenia), then associated cell types are described with functionally relevant details (e.g., brain structure, neurotransmitters, receptors, cortical layer localization).



## Present study

Based on the work of Bryois et al., Watanabe et al., and Skene et al. (who used rodent data and more limited human snRNAseq datasets<sup>23–25</sup> available at the time), we used MAGMA<sup>26,27</sup> software for analysis of GWAS data to analyze the schizophrenia GWAS results implementing

gene property analysis. We determined the cell types that had gene usage that was positively correlated with schizophrenia gene associations, while adjusting for potential confounders (e.g. gene size, gene density; see methods for details). We then used conditional analyses to empirically select representative cell types among all significant results for further discussion.

The 461 cell types tested here are hierarchically organized and numerically close cell types tend to have similar gene expression profiles. This means that there are not 461 statistically independent cell types and that Bonferroni correction is overly stringent. Nevertheless, we implemented Bonferroni correction to be conservative. This also means that significant cell types may be significant for two reasons: either they are driving a phenotypic association (truly associated), or they appear associated because they share overlapping gene expression with a cell type (or types) driving associations. The latter scenario (correlated statistical tests) is a common problem relevant to diverse domains of research, and both statistical and experimental procedures can clarify truly associated cell types. Here, we used conditional analysis (see Methods), the leading statistical approach for handling this issue in this type of analysis<sup>24,28</sup>. In brief, each cell type was tested for association while controlling for each of the other cell types. Interpretation-wise, it is most conservative to only focus on these 'significant independent' cell types, hence our primary focus on the significant independent cell types in this manuscript.

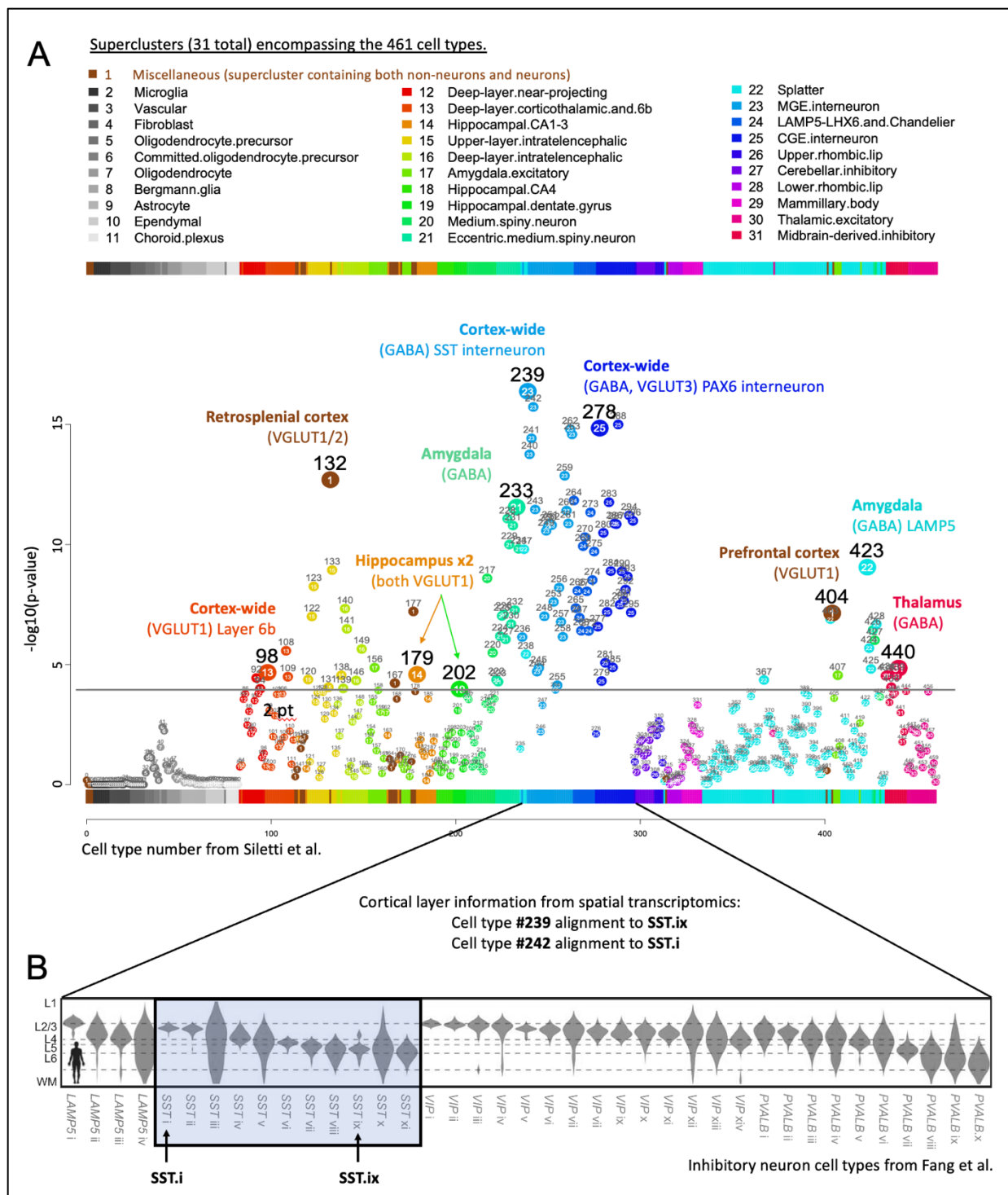
Summarizing, we used robust statistical procedures to combine two massive, unbiased, genome-wide and brain-wide datasets to systematically test which brain cell types were linked with schizophrenia. To validate this approach and determine disease-specificity, we also analyzed four comparison phenotypes (alcohol consumed per week<sup>29</sup>, sleep duration per night<sup>11</sup>, multiple sclerosis<sup>30</sup>, and Alzheimer's disease<sup>31</sup>). Results from these analyses matched prior expectations (e.g. of immune cell relevance to multiple sclerosis) and also revealed novel cell type associations. In sum, these findings yielded a data driven map of the cellular etiology of schizophrenia, demonstrated clear extensibility to other brain phenotypes, and suggested a tractable roadmap toward personalized psychiatric treatment using low-cost genotype data. Throughout this manuscript, please refer to **Supplementary Table S1** for more complete information about the 461 cell types analyzed in this report (numbered 0-460, per Siletti et al.<sup>6</sup>).

## Schizophrenia associated cell types

Of the 109 total cell types that were significantly associated with schizophrenia, there were ten significant independent cell types (see **Figure 2**), as follows. The most significant cell type was a subtype of somatostatin interneurons (#239,  $p=4.3 \times 10^{-17}$ ). The next two significant independent cell types were also cortical: PAX6 interneurons distributed widely across the cortex (#278,  $p=1.5 \times 10^{-15}$ ) and an excitatory cell type found almost exclusively in retrosplenial cortex (#132,  $p=2.1 \times 10^{-13}$ , 91% of cells from retrosplenial cortex). Note that the PAX6 interneurons (#239) were annotated as GABA/VGLUT3, indicating co-expression of GABA and glutamate, which is relatively uncommon. Fourth and fifth were two distinct inhibitory amygdala neuron types #233 ( $p=2.8 \times 10^{-12}$ ) and #423 ( $p=9.0 \times 10^{-10}$ ), see below for more details. The remaining five cell types had neurons primarily from the prefrontal cortex (#404,  $p=7.3 \times 10^{-8}$ , specifically from Brodmann area 14), thalamus (#440,  $p=1.4 \times 10^{-5}$ ), cortex-wide excitatory

neurons annotated to deep layer 6b (#98,  $p=2.2 \times 10^{-5}$ ), and two excitatory hippocampal neuron types (#179  $p=2.6 \times 10^{-5}$  and #202  $p=1.1 \times 10^{-4}$ ). In addition to these ten significant independent cell types, other notable significant cell types were medium spiny neurons in the striatum (i.e., caudate and putamen, #222), cortex-wide excitatory neurons in layer 2/3 (#123), as well as cell types preferentially located in visual cortex (#133), septal nuclei (#428), superior colliculus (#433, #367), and substantia innominata (#232).

**Figure 2. Cell type associations for schizophrenia reveal known and novel cell types for schizophrenia.** **A.** Schizophrenia results for 461 cell types are depicted in a scatterplot, with cell type number on the x-axis and statistical significance on the y-axis ( $-\log_{10}(p)$ ), such that higher values are more statistically significant). The horizontal gray line denotes Bonferroni correction for the 461 cell types ( $p < .0001$ ). Numbers above points are cell type numbers and numbers within points are supercluster numbers. Larger points denote the ten cell types that were ‘significant independent’ cell types after conditional analyses (see Methods). **B.** The top cell type for schizophrenia (#239) aligned to somatostatin interneuron subtype SST.ix (from Fang et al.) for which cortical layer localization had been determined using spatial transcriptomics. Thus, we inferred that cell type #239 was likely most abundant in cortical layer 5, but also found in cortical layers 6 and 2/3, whereas cell type #242 was likely most abundant in layers 2/3.



## Interneurons

For schizophrenia, arguably the most widely replicated postmortem tissue finding involves aberrant inhibitory neurons in the cortex (interneurons)<sup>32,33</sup>. While both the parvalbumin and somatostatin (SST) subclasses of interneurons have been implicated, the most recent evidence shows the greatest abnormalities in somatostatin interneurons<sup>32,33</sup>. Thus, the

present report of the strongest association with somatostatin (SST) interneurons, and many significantly associated interneurons (including parvalbumin), is concordant with leading prior findings. We extended these findings by specifying more subtle subtypes of somatostatin interneurons and by determining the probable cortical layers for these schizophrenia-associated somatostatin interneuron subtypes (#239 and #242<sup>ii</sup>) using additional snRNAseq and spatial transcriptomics datasets.

Specifically, recent studies have identified far more interneuron subtypes than were distinguishable using traditional markers. These transcriptomically-defined cell types reside preferentially in specific cortical layers in both humans and mice<sup>7</sup>. In **Figure 2B**, with a plot adapted from Fang et al.<sup>7</sup>, the cortical layer distribution of human interneuron subtypes is shown (i.e. cortical layers 1-6). We aligned somatostatin subtypes from the Linnarsson<sup>6</sup> and Zhuang<sup>7</sup> laboratories, and found that #239 and #242 (from Siletti et al.) best matched SST.ix and SST.i, respectively (from Fang et al.), and that these cell types differed in cortical layer distribution. #239 somatostatin interneurons were most likely to be localized to cortical layer 5 (and less so to layers 6 and 2/3), while #242 somatostatin interneurons were most likely to be in layers 2/3. Alignment to two mouse datasets<sup>8,34</sup> yielded similar cortical layer localization results, providing additional support for putative layer localizations. Contextualizing these findings, we note that schizophrenia postmortem tissue studies have primarily identified deficits in upper cortical layers, particularly deep layer 3<sup>35,36</sup>, which may be consistent with the primary localization of cell type #242 and, to a lesser extent, #239. Given known disruptions in sensory processing and integration in schizophrenia, it is conceivable that normal cortical layer 5 integration and output functions<sup>37</sup> are disrupted in schizophrenia due to abnormalities in layer 5 somatostatin interneurons, of the #239 type.

### *Amygdala*

Within the amygdala, which is a structure known to have diminished volume in schizophrenia<sup>38</sup>, we found 17 significant cell types (including both inhibitory and excitatory), two of which were independent significant cell types (#233 and #423, both inhibitory). Cell type #233 was annotated by Siletti et al as a medium spiny neuron of the “eccentric” subtype, a recently discovered subtype of medium spiny neurons<sup>10</sup>. This general class of neurons – medium spiny neurons – has been frequently linked to schizophrenia, and medium spiny neurons are the dominant inhibitory neuron type of the striatum (caudate and putamen). In addition to the two major subtypes of medium spiny neurons, which are well characterized based on their differential expression of dopamine receptors (D1 versus D2), recent transcriptomic studies made clear that a third subtype of medium spiny neurons also exists. These ‘eccentric’ medium spiny neurons had evaded detection because classical D1/D2 markers do not reliably differentiate the newly named ‘eccentric’ medium spiny neurons<sup>10</sup>. Thus, the linking of this eccentric medium spiny neuron type (#233) to schizophrenia demonstrates multiple benefits of large-scale transcriptomic studies and the unbiased approaches used here to link cell types to

---

<sup>ii</sup> We also examined cell type #242 because it was nearly as significant as the top cell type (#239). This second-most significant cell type (#242) was also associated with other psychiatric phenotypes, was larger than #239 (i.e., encompassed ~3x more cells in Siletti et al), and it was the most significant cell type in our analysis of the 2018 schizophrenia GWAS. Thus, we thought #242 was also a strong candidate for psychiatric relevance.



schizophrenia. First, scRNAseq studies made possible the detection of the relatively rare (~4%) but clearly transcriptomically distinct “eccentric” medium spiny neuron subtype<sup>10</sup>. Second, the brain-wide Siletti et al snRNAseq dataset used here shows that neurons that transcriptomically resemble all three subtypes of medium spiny neurons (D1, D2, and eccentric) are also found outside of the striatum. And third, this approach links these newly discovered, relatively rare, and extra-striatal eccentric medium spiny neurons to schizophrenia. Such a discovery would not have been possible without these large, comprehensive, and “unbiased” datasets that allow us to see more clearly the gaps in prior knowledge, and to discover which newly filled gaps might also provide critical information about schizophrenia etiology<sup>iii</sup>.

The second significant independent amygdala cell type (#423) is inhibitory and while found predominately in amygdala (51% of cells), is also present in the thalamus (20%), hypothalamus (15%), and other brain regions. Notable gene enrichments for this cell type include a cytochrome p450 gene (*CYP19A1*) responsible for the biosynthesis of estrogen, a sodium channel subunit gene (*SCN5A*), a GABA receptor A subunit gene (*GABRQ*), and neurotensin, a neuropeptide that modulates dopamine and other neurotransmitters relevant to psychiatric disorders, which has been investigated as a therapeutic target for schizophrenia<sup>39,40</sup>. Other amygdala cell types associated with schizophrenia, but not deemed independent after conditional analyses, include medium spiny neurons of the D1 (e.g. #220) and D2 (#217) subtypes.

### *Hippocampus*

Our hippocampal findings provide molecular details that can augment and extend schizophrenia imaging and postmortem tissue findings. Prior work showed lower average volume of the hippocampus and smaller individual hippocampal neurons, as reported in meta-analyses<sup>41</sup>. Here, we have identified specific neuron types that may underlie these hippocampal volume reductions in schizophrenia. Of the 86 neuron types located primarily in the hippocampus (from Siletti et al.’s data<sup>6</sup>), seven were significantly associated with schizophrenia in our analyses. Two of those seven significant hippocampal neuron types were significant independent cell types after conditional analyses (#179 and #202). Cell type #179 is an excitatory neuron type (VGLUT1) and contributing cells came from all major subregions of the hippocampus (subiculum, CA1-4, dentate gyrus). Cell type #202 is also excitatory (VGLUT1) and contributing dissections encompass all major hippocampal subregions except the subiculum. The present report may also be concordant with prior reports of increased hippocampal excitatory (glutamate) neurotransmitter metabolites in individuals with schizophrenia<sup>42</sup>, and based on these results future studies can target more specific subtypes of neurons.

In sum, the primary anatomical locations for the significant independent schizophrenia cell types were three widely distributed across the cortex, two from specific cortical regions

---

<sup>iii</sup>Note that we also found medium spiny neurons from the striatum to be associated with schizophrenia (#222 and #224).

(retrosplenial and prefrontal), two from amygdala, two from hippocampus, and one from the thalamus. Of note, the three subcortical structures linked to schizophrenia here (amygdala, hippocampus, and thalamus) are precisely the same subcortical structures found to have the largest volume decreases in schizophrenia patients as compared to controls (Hedges  $g$  of  $-.66$ ,  $-.46$ , and  $-.31$ , respectively) in a recent meta-analysis of brain volume studies in first-episode psychosis<sup>43</sup>. Thus, two completely independent, brain-wide, data-driven approaches to understanding schizophrenia – one using imaging data<sup>43</sup> and the present approach using genetic data – pointed to the same three subcortical regions: amygdala, hippocampus, and thalamus. In the next section we consider the cell types and brain regions linked to four other phenotypes using this approach.

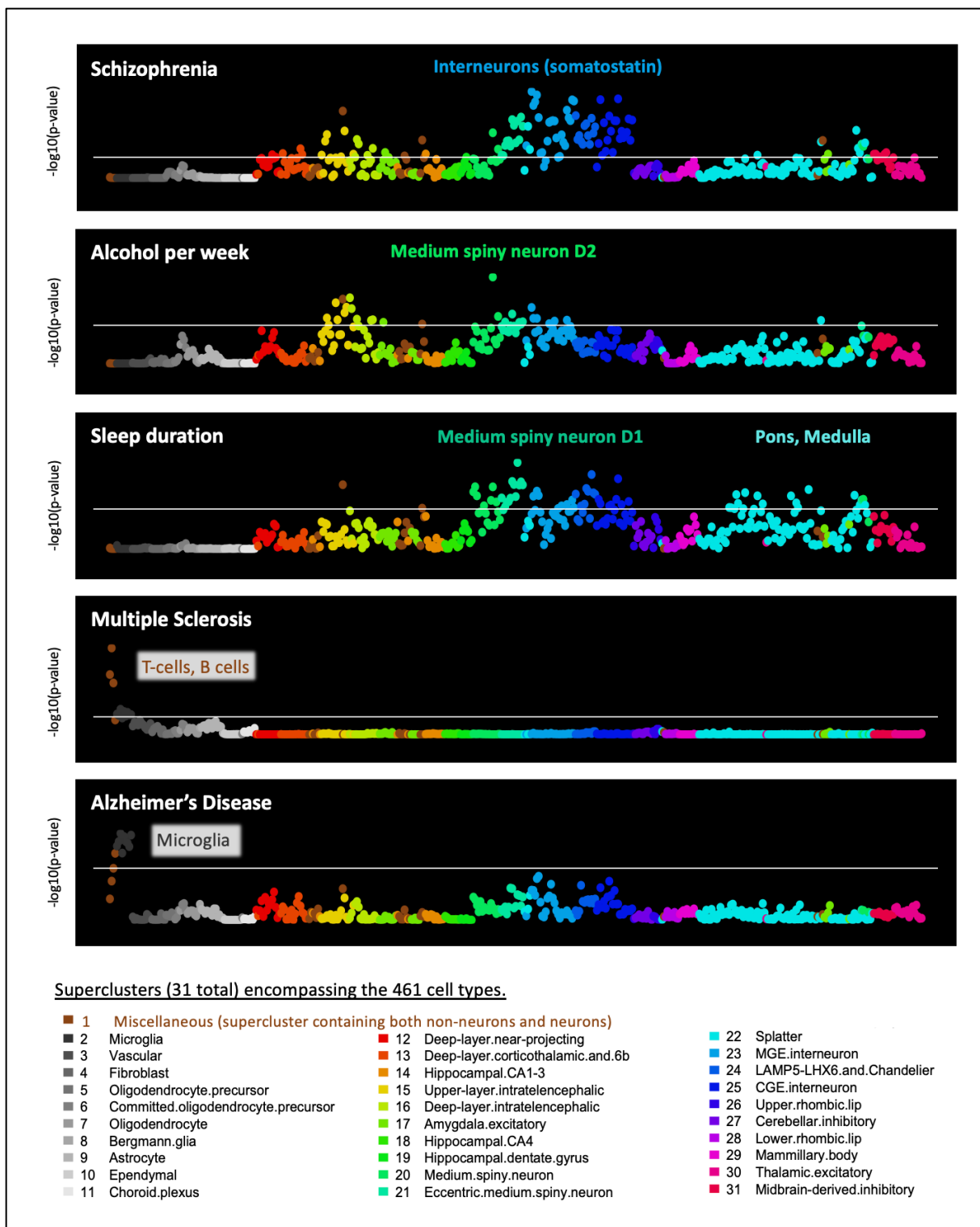
### Contrasting five brain phenotypes

Prior investigations across different branches of medicine have linked cell types to phenotypes using gene expression and GWAS data<sup>23–25,44</sup>. We sought to establish whether this new human snRNAseq dataset could also afford accurate identification of cell types linked to comparison brain-related phenotypes (alcohol consumed per week, sleep duration per night, multiple sclerosis, and Alzheimer’s disease; see the Methods for the rationale for selecting these phenotypes). As shown in **Figure 3**, expected cell type phenotype pairs were found and are discussed below (see **Supplementary Tables 2-5** for full results).

**Figure 3. Distinct cell type profiles for five phenotypes.** Cell-type associations for five brain-related phenotypes are depicted in scatterplots, with cell type number (0-460) on the x-axis and statistical significance on the y-axis (reported as  $-\log_{10}(p)$ , such that higher values are more statistically significant). Horizontal white lines reflect correction for 461 tests (i.e.  $p < 0.0001$ ). The first three phenotypes (schizophrenia, alcohol consumed per week, and sleep duration per night) yielded neuronal cell-type associations<sup>iv</sup>. The two neurological phenotypes (multiple sclerosis & Alzheimer’s disease) yielded non-neuronal cell type associations of immune and microglial cell types, respectively. See **Supplementary Figures 1-4** for larger annotated scatterplots for the comparison phenotypes.

---

<sup>iv</sup>Committed oligodendrocyte precursors (cOPC) were the most significant non-neuronal cell types for psychiatric phenotypes. Note that Bonferroni correction for 461 cell types is overly stringent, and we are not arguing that glia play no role in psychiatric phenotypes.



### *Schizophrenia, Alcohol consumed per week, and Sleep duration per night*

We discussed previously the notability of the somatostatin interneuron associations with schizophrenia (e.g., #239), given prior findings of somatostatin interneuron abnormalities in

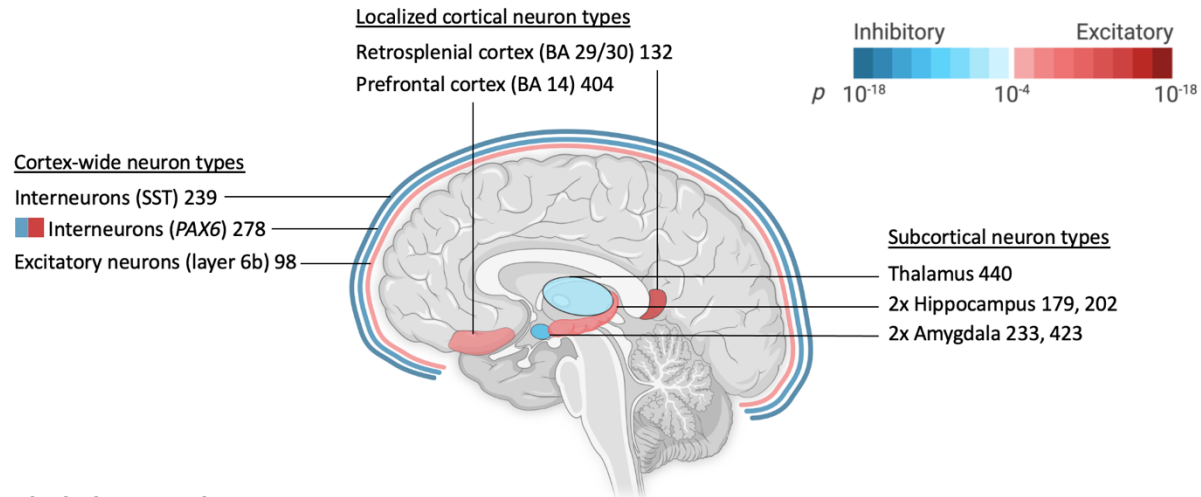
schizophrenia<sup>32,33</sup>. For alcohol consumed per week, the top cell type was a D2 medium spiny neuron (#217,  $p=1.3 \times 10^{-9}$ ), which also matched prior expectations given that D2 medium spiny neurons causally influence alcohol consumption<sup>45-49</sup>. For sleep duration per night, the top cell type was also a medium spiny neuron, but of the D1 type (#231,  $p=2.8 \times 10^{-9}$ ). This finding may be consistent with recent findings linking D1-MSNs and sleep, particularly rapid eye movement (REM) sleep<sup>50</sup>. For sleep duration per night, we also highlight associated cell types from the pons (#396,  $p=1.3 \times 10^{-6}$ ) and medulla (#386,  $p=5.9 \times 10^{-6}$ ) since these structures are key nodes of sleep regulatory circuits<sup>51</sup>. Note that these pons and medulla cell types were not associated with schizophrenia or alcohol consumption, but rather they are specific among these phenotypes for sleep. Note as well that the sleep phenotype had many associations with the newly named “Splatter” supercluster of neurons (described in Siletti et al.<sup>6</sup>), which includes a wide variety of subcortical neuron types. See **Figure 4** for a comparison of the brain regions harboring the significant independent cell types for these three psychiatrically relevant phenotypes.

### *Multiple sclerosis and Alzheimer’s disease*

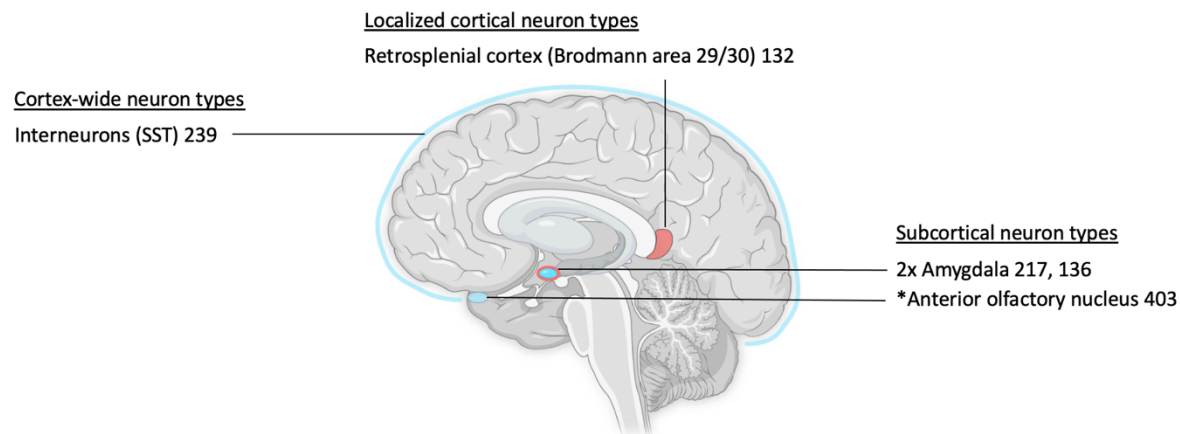
For multiple sclerosis, an autoimmune condition, the strongest association was a T cell type (cell type #1,  $p=6.0 \times 10^{-20}$ ). The next most significant cell types were B cells (cell type #0,  $p=4.7 \times 10^{-14}$ ) and natural killer cells (cell type #2,  $p=3.5 \times 10^{-12}$ ). These results are consistent with longstanding understanding of pathogenic T cell involvement in multiple sclerosis, best-available treatments for multiple sclerosis, and also with more recent findings linking B cells to multiple sclerosis<sup>52-54</sup>. For Alzheimer’s disease, the most significant cell types were microglial (most associated cell type: #6,  $p=2.4 \times 10^{-7}$ ). Again, these results are consistent with longstanding (>100 years ago) and newer findings. Specifically, glial alterations were reported by Alzheimer in his 1907 neuropathological description of the disease<sup>55</sup>, prior work established that many Alzheimer’s disease genes are preferentially expressed in microglia<sup>30</sup>, and newer therapeutics targeting microglia are being investigated for Alzheimer’s disease<sup>44,56</sup>.

**Figure 4. Brain locales of origin for the independent significant cell types associated with schizophrenia, alcohol consumed per week, and sleep duration per night.** Inhibitory cell types are depicted in blue, and excitatory cell types are in red, with saturation of color denoting statistical significance. Cell type numbers are given after brief descriptions. Unless an asterisk is present, the regions depicted were the source of >50% (and usually much >50%) of cells for a particular transcriptomic cell type. If there are two excitatory or two inhibitory cell types for a single brain structure, color saturation corresponds to the more significant cell type. If both excitatory and inhibitory cell types are associated with a single structure, then the fill of the structure denotes the more significant association (as opposed to the border). For example, there are two amygdala cell types associated with alcohol per week and the inhibitory cell type #217 is more significant (hence the blue/inhibitory fill). BA=Brodmann area, SST=somatostatin, PAX6=paired box 6 gene, VIP=vasoactive intestinal peptide, Inf=inferior. Cell type #278 is annotated as GABA/VGLUT3.

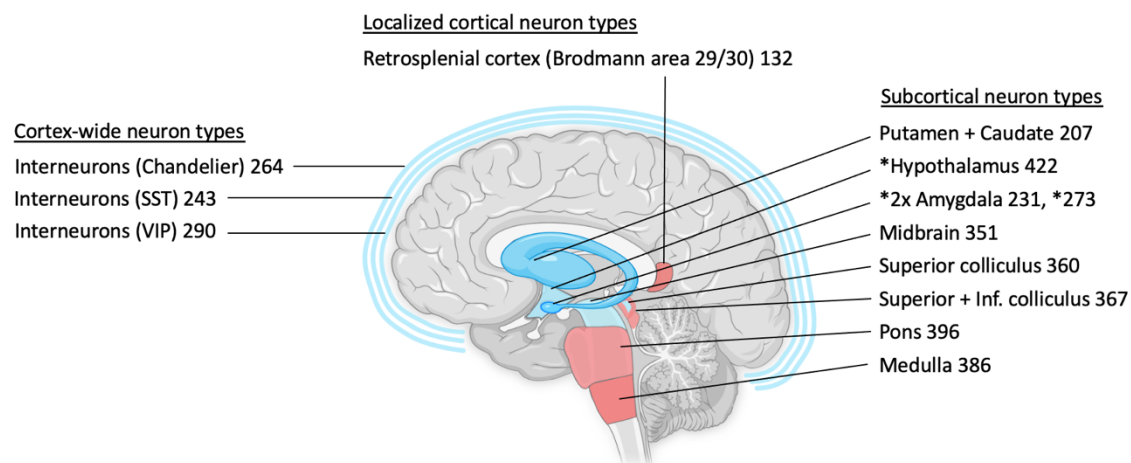
## Schizophrenia



## Alcohol per week



## Sleep per night

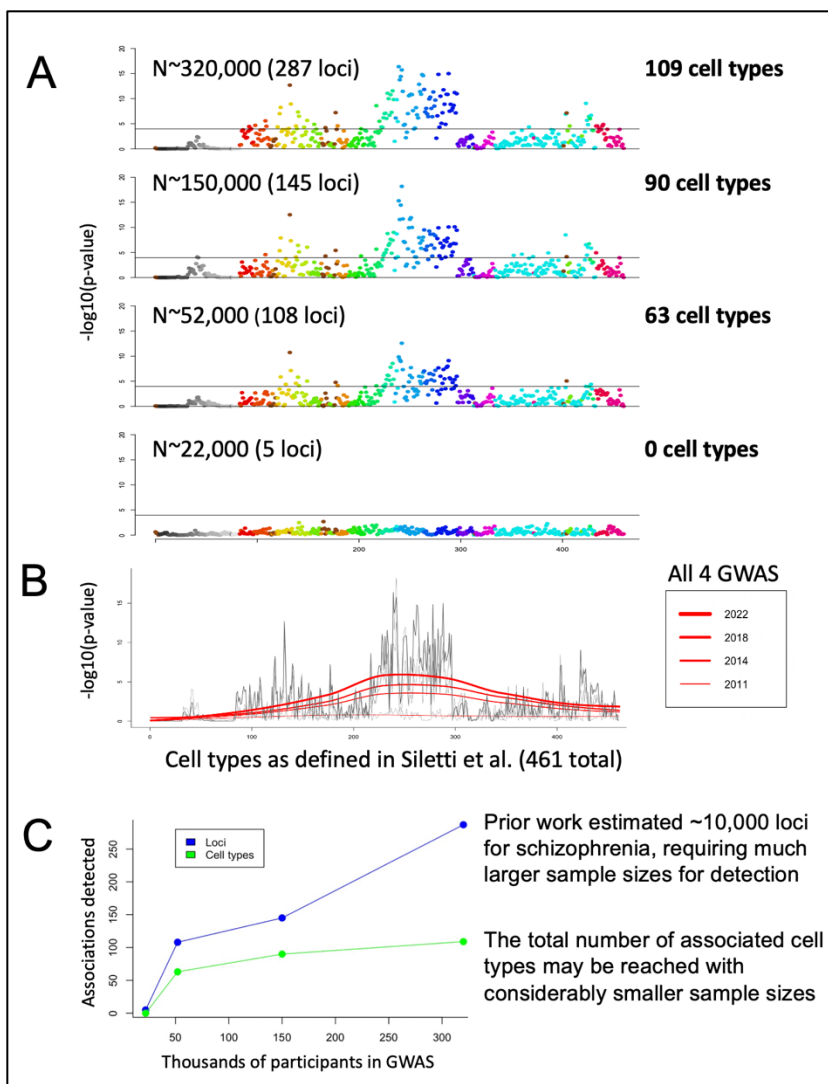


## Statistical power requirements

We also sought to determine how the statistical power of individual GWAS influenced our detection of cell type associations, hypothesizing that poorly powered GWAS might not afford discovery of cell type associations. We further hypothesized that successive GWAS with increasing statistical power (as evidenced by increasing numbers of loci detected) would also afford increasing numbers of associated cell types until most of the relevant cell types were associated (or until “saturation” was afforded by adequately powered GWAS). Saturation would be indicated by a leveling off in the number of associated cell types. Thus, we re-ran our primary analysis on four schizophrenia GWAS with progressively increasing sample sizes (and statistical power), ranging from a 2011 GWAS (N=21,865) to the most recent GWAS (N=320,404)<sup>9,19,20,57</sup>. **Figure 5** shows results consistent with our hypotheses. The smallest GWAS revealed no significant cell type associations in our analysis, though it was adequately powered to detect five GWAS loci. By contrast, the subsequent schizophrenia GWAS afforded detection of 63, 90, and then 109 cell types, corresponding to discovery of 108, 145, and then 287 loci as reported in the primary publications<sup>9,19,20</sup>. Thus, we demonstrated that cell types can be linked to psychiatric phenotypes with this approach *when adequately powered GWAS are available*, and that

“saturation”, of cell type discovery appears to occur at much smaller sample sizes than “saturation” of loci (predicted to saturate at >10,000 associated variants for psychiatric disorders<sup>4,5</sup>). The plateauing of associated cell types suggests that relevant biological features are being captured at the cell type level, via aggregation of polygenic signal into cell types.

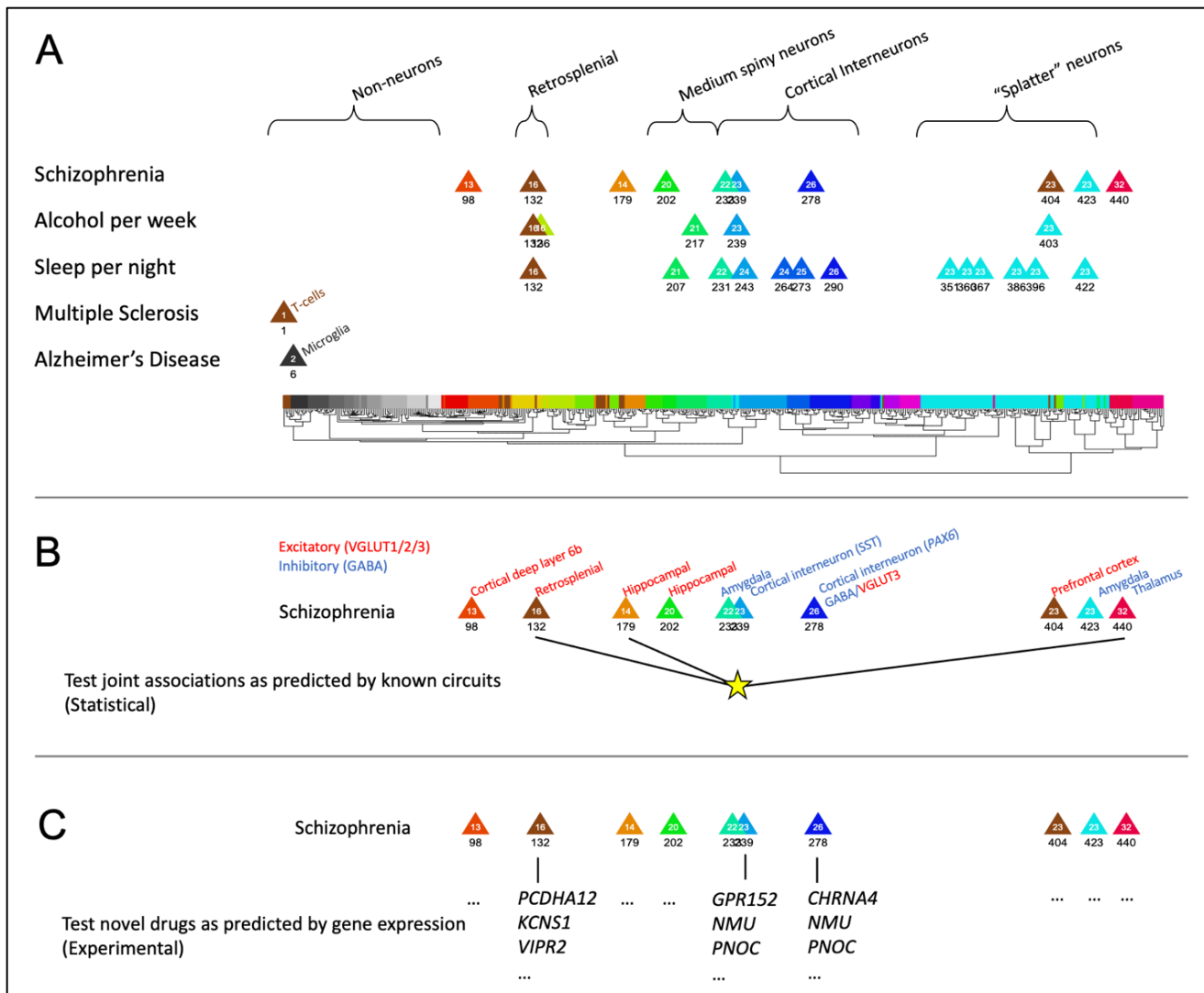
**Figure 5. Demonstration of required statistical power of the GWAS for detecting cell type associations.** **A.** Results for four successive, increasingly larger, and better powered, schizophrenia GWAS. Horizontal lines in all plots denote Bonferroni corrected statistical significance for 461 cell types ( $p < .0001$ ). **B.** All four sets of results superimposed with loess lines in red. **C.** In this range of sample sizes for schizophrenia GWAS, the number of loci detected continues to increase steeply, but the number of cell types plateaus.



## Towards a cellular psychiatric taxonomy

This work demonstrates the possibility of building a taxonomy for psychiatric disorders based on quantitative evaluation of cell types. **Figure 6A** illustrates this idea for the five examined phenotypes: schizophrenia, alcohol consumed per week, sleep per night, multiple sclerosis, and Alzheimer's disease. The proposed cellular taxonomy naturally structures future exploration, both statistical and experimental. As shown in **Figure 6B**, future statistical investigations can test for evidence consistent with schizophrenia-linked cell types acting together. For example, particular combinations of cell types act together in small local circuits, while others function together through longer-range projections. Pairs, triplets, or any other combination of cell types may be tested together for associations with schizophrenia. **Figure 6C** depicts another way that cell-type associations with schizophrenia may ultimately improve schizophrenia treatment. By examining the molecular properties of schizophrenia-associated neurons (e.g. neurotransmitter usage, receptor expression, etc.), scientists might repurpose currently available drugs, or design new drugs, to target these cell types, thus modulating specific dysfunctional brain circuits. By examining other cell types for the expression of relevant receptors, it may also be possible to predict side effects. Perhaps most excitingly, future work might even parse genetic risk for schizophrenia *at an individual level* using cell-type or circuit-level polygenic scores. The point of this endeavor is not to explain the most phenotypic variance (as is the goal of full-genome polygenic scores), but rather to predict the most vulnerable cell types for a given individual, and to accordingly tailor treatment. Note that the cost of these individualized predictions should be minimal, requiring only GWAS genotyping (<\$50 USD) and analysis (extremely inexpensive after development of algorithms; i.e., ~\$150 USD and subsequently decreasing).

**Figure 6. A cellular taxonomy for brain phenotypes maps shared and non-shared genetic influences in a mechanistically informative manner, and implies next steps toward deeper mechanistic understanding of schizophrenia as well as novel, personalized treatments.** The 31 superclusters are color coded as in Figure 2. **A.** Significant independent cell types are depicted for the five analyzed phenotypes. Given that psychiatric disorders have shared and non-shared genetic influences; this approach may reveal the cellular (and therefore mechanistic) relevance of these shared and non-shared cell type associations. For example, shared or 'pleiotropic' cell types include cortical interneurons (#239) and excitatory retrosplenial cortex neurons (#132). Regarding non-shared cell types, sleep-specific associations of pons (#396) and medulla (#386) "splatter" neurons are notable. Future work may extend to additional psychiatric disorders or subtypes of psychiatric disorders (e.g. positive, negative, and/or disorganized symptoms of schizophrenia). **B.** Statistical follow-up approaches include testing for joint association of cell types known to function together in circuits. **C.** Experimental follow up approaches include testing of drugs as predicted by gene expression in associated neuron types. Examples of predicted targets, in three schizophrenia associated cell types (#132, #239, #278), are given for illustration purposes, but data afford much more extensive predictions, for all cell types.



## Discussion

Major scientific advances often depend on newly available data obtained either in greater detail or scale, or via combined datasets capable of yielding novel insights. The vision and perseverance of leaders of the Psychiatric Genomics Consortium (PGC) enabled discovery of thousands of specific risk loci for schizophrenia and other psychiatric disorders, using unprecedented sample sizes (>1 million participants)<sup>9,18,19,29</sup>. Combining the PGC schizophrenia dataset with another landmark data resource, a single nucleus RNA sequencing (snRNAseq) dataset of 3,369,219 unique cells from 106 human brain regions<sup>6</sup>, we were able to use robust analytical approaches<sup>26</sup> to discover cell types likely to be directly involved in the etiology of schizophrenia based on their expression of schizophrenia associated genes. To our knowledge, this is the first time that such a comprehensive catalog of human brain cell types has been systematically tested for associations with schizophrenia.



Cell type associations consistent with prior findings (discussed throughout the manuscript) lend confidence to our results, but novel findings can broaden our understanding of these disorders. Highlighting just two examples, we note that fear and the sense of self are two symptom domains that are clinically important for many patients with schizophrenia. Indeed, fear is one of the defining emotional features for many who suffer from schizophrenia and the amygdala is critical for fear processing. Thus, the amygdala cell types reported here (#233, #423) may be causally linked to the maladaptive fear that causes so many problems for individuals with schizophrenia. A more amorphous, but nevertheless clinically relevant and historically notable feature of schizophrenia concerns distortions in the sense of self. Consequently, it is noteworthy that we discovered a highly associated cell type (#132) in a brain region (retrosplenial cortex) that is critical for the sense of self. Specifically, recent findings showed that a particular rhythm in the retrosplenial cortex causes dissociation<sup>58</sup>, meaning alterations in one's sense of self. In other words, we found that neurons in a brain region critical for an integrated sense of self are associated with schizophrenia, a disorder noted for alterations in the sense of self. Moreover, this retrosplenial neuron type is significantly associated with many psychiatric phenotypes, and thus may represent a common node of dysfunctional processing in psychiatric disorders.

Historically, classifying psychiatric disorders has proven challenging, generating enduring debate about psychiatric nosology. Our final figure depicts cell type profiles for the five phenotypes analyzed here, providing a proto-taxonomy for these phenotypes. Here we suggest that **cell typologies may prove the right level of analysis for a stable taxonomy of brain disorders; making available the appropriate measurements and organizing principles**. What remains to be determined is the relevance of developmental versus adult cell types to different disorders, the identifiability of subtypes of disorders with this approach, and the eventual impact of this research on classification of psychiatric disorders.

A cellular classification approach might be further refined using GWAS of different subtypes of psychiatric disorders. Appealingly, the long-postulated heterogeneity within psychiatric disorders (e.g., the 'schizophrenias') may be wrangled into comprehensibility with these approaches. Different cell type associations might be identified, for example, for major symptom domains of schizophrenia, including positive, negative, and disorganized symptoms. Further, by specifying molecular properties of etiologically relevant cell types for different disorders (e.g. receptors and neurotransmitters) this classification system should naturally nominate treatment targets, as well as relevant therapeutics. By analogy, quantifying and organizing electrons and protons allowed classification of elements into the periodic table and thus generated novel predictions about elemental interactions. A cell-typology for psychiatric disorders may similarly afford novel predictions about which medications and treatments are most likely to be effective for each disorder. One such approach involves developing polygenic risk scores for cell types or circuits. Individual level metrics of genetic risk, polygenic risk scores, are already available for multiple psychiatric disorders. Typically, polygenic scores use variants from across the genome to maximize accurate prediction of risk. To determine

the best medications for individuals, however, cell-type restricted polygenic scores may eventually prove more useful.

### *Limitations*

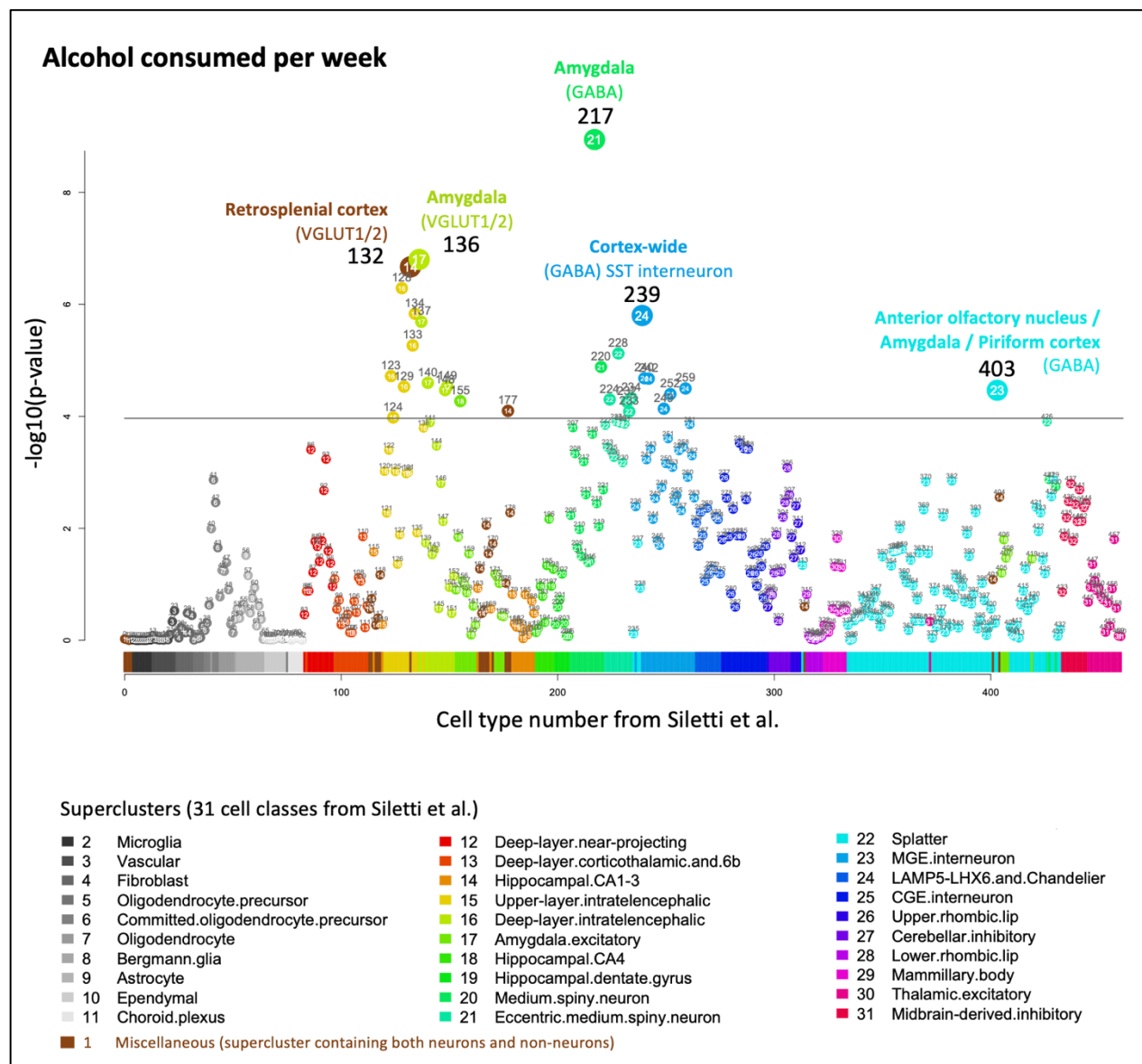
There are important limitations to consider regarding this approach and currently available datasets. First, this approach cannot discover cellular abnormalities resulting exclusively from environmental influences. Conversely, this “limitation” also implies that reported cellular associations do not result from environmental risk factors, treatment, or consequences of living with schizophrenia. Second, although these datasets represent landmark achievements in GWAS and transcriptomically-derived cell types, they have important limitations, which diminish the ultimate power of this approach. Briefly, this snRNAseq dataset<sup>6</sup> did not include all brain regions, the female sex, or non-European ancestry individuals. Specific brain regions not sampled in Siletti et al. which may be particularly important for schizophrenia include the dorsolateral prefrontal cortex (DLPFC Brodmann area 9) and the ventral tegmental area of the midbrain (VTA). Further, since this snRNAseq dataset came from adult humans, an important next step will be conducting the same analyses using developmental datasets. For example, another dataset from the Linnarsson group<sup>59</sup> includes human brains from post-conception weeks 5-14, a time period critical for neurogenesis and potentially of importance for psychiatric disorders.

The statistical methods used here were selected for their robustness to false positives and prior use with transcriptomic datasets, but such models may be further optimized, for example via inclusion of data from exome sequencing studies of schizophrenia. Further, we used a straightforward method of linking genetic variants to genes based only on proximity, and thus longer-range effects of variants on genes were omitted from our analyses. Finally, to the extent that the central assumption of our analysis method – that cell types that make greater relative use of schizophrenia-associated genes than other cell types – is incorrect, cell type associations reported may be incorrect.

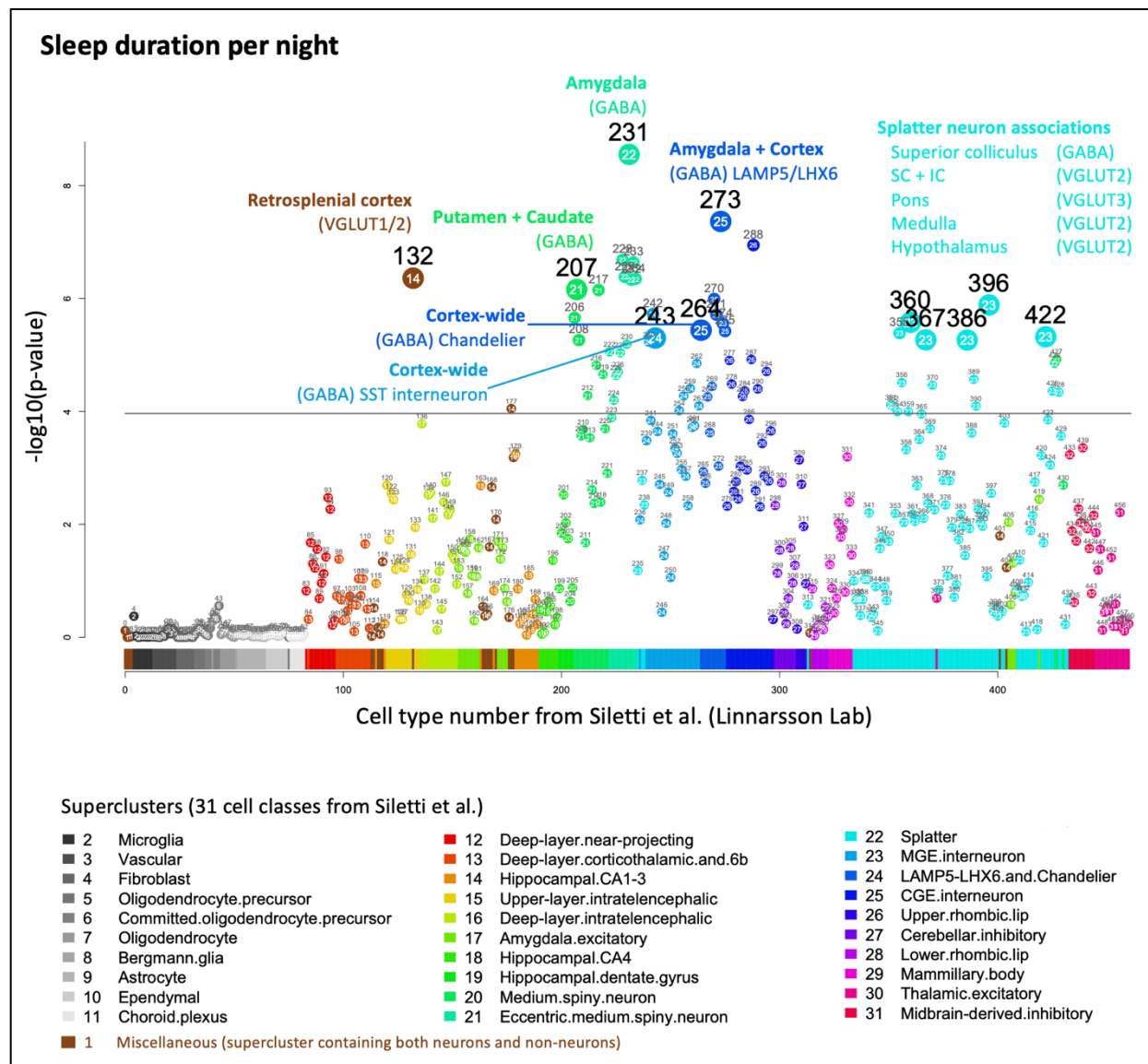
In conclusion, the results reported here are possible thanks to a combination of natural endowment (the genome as a catalog of biological elements) and human technology (i.e., modern advances which allow us to query the genome – in both RNA and DNA forms – in great detail). Together, these results provide a new framework for understanding the cellular basis of psychiatric disorders. Given the detailed molecular information available about each cell type (e.g. receptors, neurotransmitters, neuropeptides), this work also provides a roadmap toward novel therapeutics, better use of existing medications, and truly personalized medicine.

## SUPPLEMENTARY FIGURES

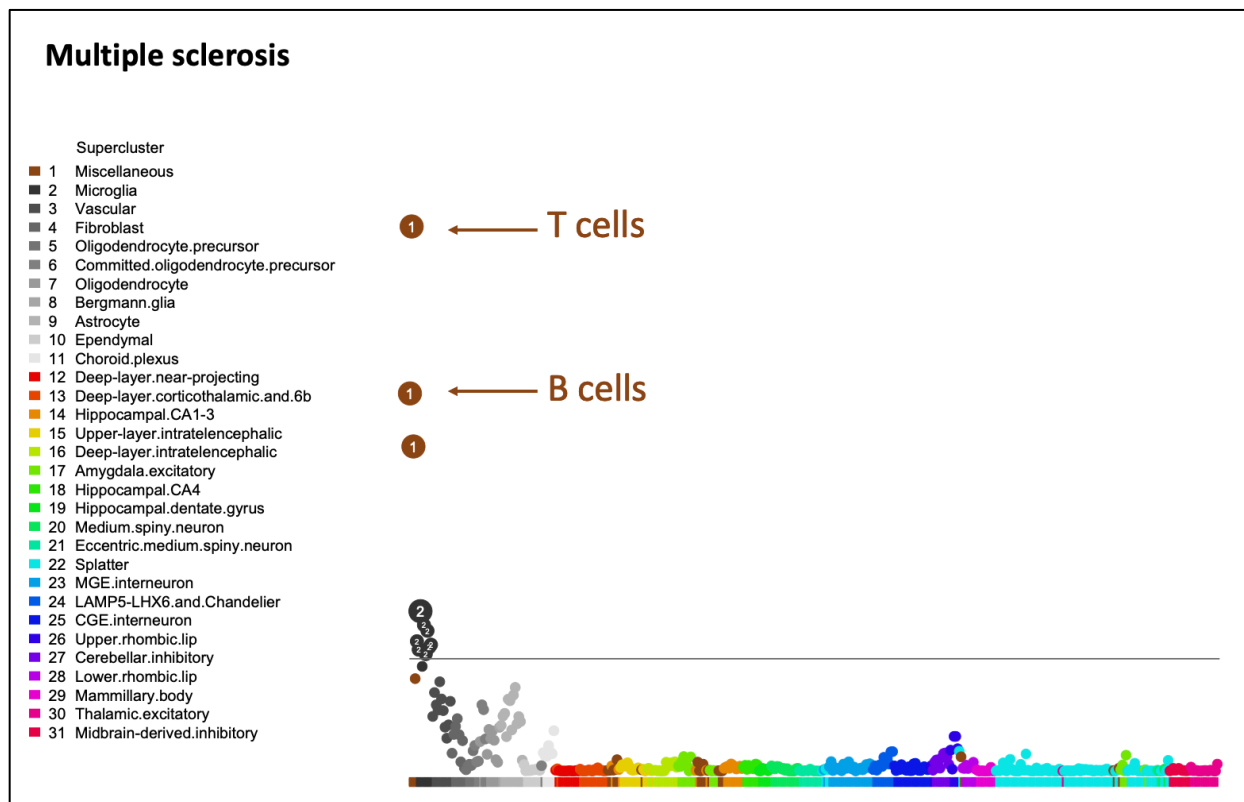
**Supplementary Figure 1. Cell type associations for alcoholic drinks consumed per week.**



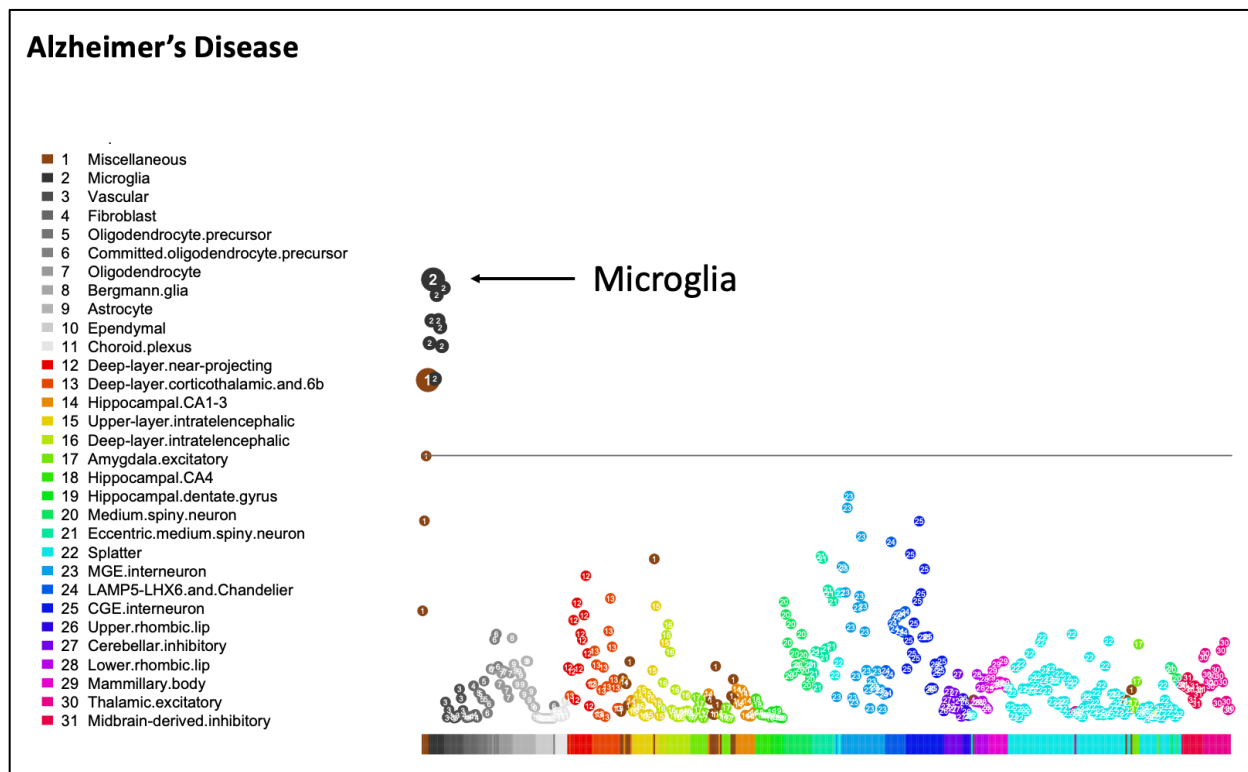
## Supplementary Figure 2. Cell type associations for sleep duration per night.



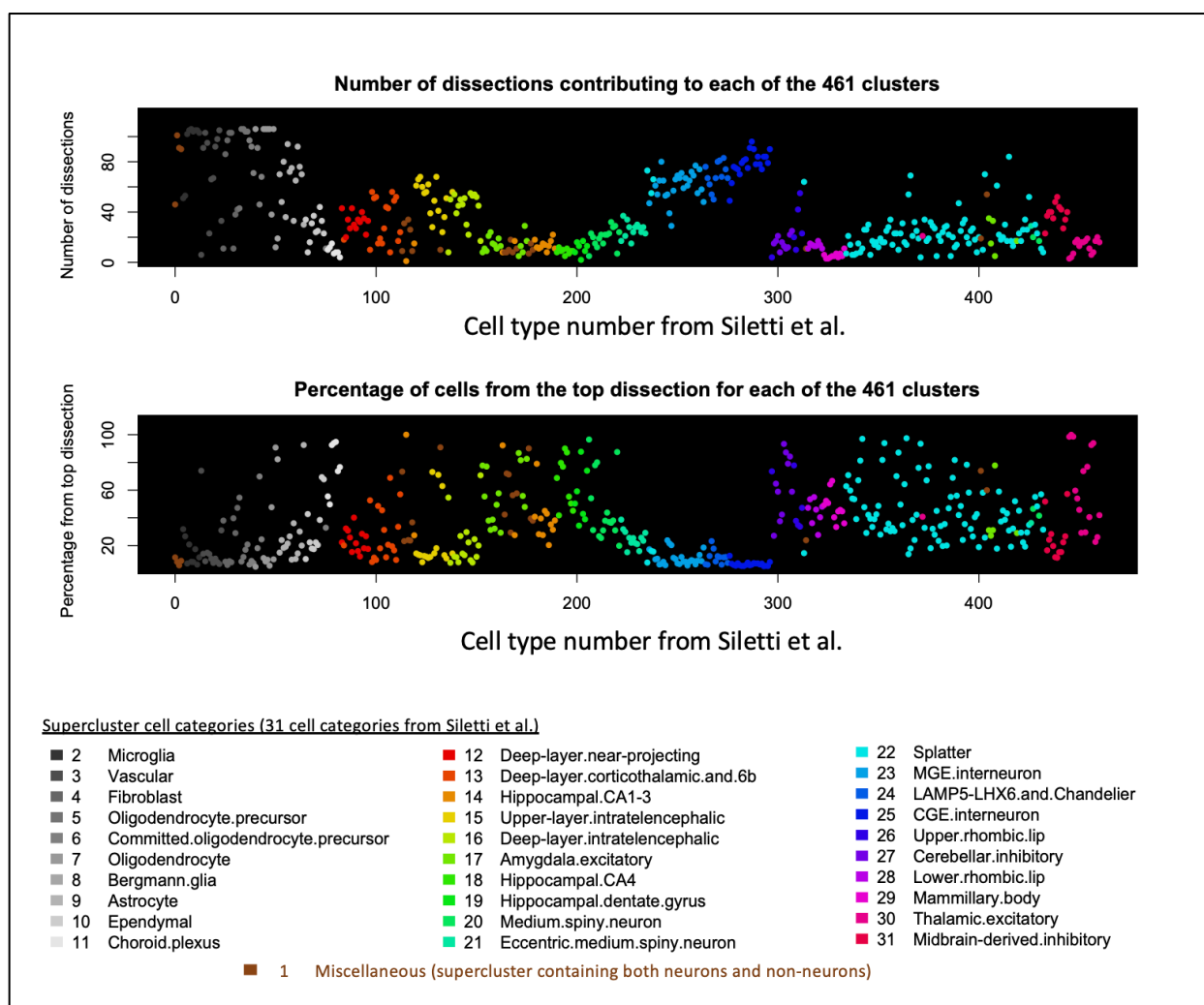
### Supplementary Figure 3. Cell type associations for multiple sclerosis



## Supplementary Figure 4. Cell type associations for Alzheimer's disease



**Supplemental Figure 5. Information about the number of dissections contributing to each of the 461 cell types.** For example, interneurons (cell types #236-296), shown in blue shades, are among the cell types that are found in the largest number of dissections (top) and no single dissection accounted for >25% of any interneuron type, and for most the top contributing dissection accounted for <10% of the cells in that cell type (bottom). The number of contributing dissections for these widely-distributed interneuron cell types was 56-84 out of 106 possible dissections.



## METHODS

### *Selecting comparison phenotypes*

In addition to schizophrenia, we selected four comparison phenotypes (alcohol consumed per week, sleep duration per night, multiple sclerosis, and Alzheimer's Disease). These phenotypes met the following criteria: 1) they were polygenic, and 2) well powered GWAS of those phenotypes were available (both conditions are required for this method). Regarding psychiatric comparison phenotypes, we wanted to mitigate concerns about the reliability of psychiatric diagnoses<sup>v</sup>, so we chose alcohol consumed per week and sleep duration per night because they are readily quantifiable (albeit noisily). The two neurological phenotypes (multiple sclerosis and Alzheimer's disease) were selected because have considerably different etiology from psychiatric disorders<sup>60</sup>, and consequently they provided an opportunity to observe greater contrasts among this set of five brain-related phenotypes. Genome-wide association studies (GWAS) results used here were from the best powered GWAS of these phenotypes available at the time of analysis (Nov 2022-July 2023). The relevant sample sizes and source publications were as follows: schizophrenia<sup>9</sup> (N=320,404), alcohol per week (N=2,428,851), sleep duration per night<sup>11</sup>(N=446,118), multiple sclerosis<sup>30</sup> (N=41,505), and Alzheimer's disease<sup>31</sup> (N=788,989). Given that well powered GWAS are required for these analyses, we used only European ancestry samples since adequately powered GWAS are not available for any other ancestry, for all these phenotypes.

### *Gene expression data*

The gene expression data from Siletti et al. is publicly available. Briefly, Siletti et al. sampled 3,369,219 nuclei from 106 different dissections in three postmortem brain donors using high-throughput 10X Chromium single-nucleus RNA sequencing. The nuclei were then computationally clustered into 31 superclusters and 461 cluster. The 461 clusters are referred to as "cell types" in our paper. For data processing, we first applied a transformation to the single-cell expression data to compress the scale and reduce outliers (as is typical for such data, here we used  $\ln(1 + x)$ ), and obtained the mean transformed expression for each gene in each cell type. We only kept protein-coding genes from the National Center for Biotechnology Information (NCBI) and removed unexpressed genes, genes with non-unique names, and genes within the MHC region (chromosome 6 base positions 25,000,000-34,000,000). For MAGMA use, we mapped Ensemble IDs to Entrez gene IDs with the Genome-Wide Association for Human package ("org.Hs.eg.db") in Bioconductor (v3.12.0). Genes without unique Ensemble-Entrez mappings were removed. A metric of gene expression specificity was then calculated by dividing the transformed expression of a gene in a cell type by the sum of the transformed expression of that gene across all cell types, yielding a value between 0 and 1 that characterizes the extent to which a particular gene is expressed in a particular cell type. For example, in hippocampal neuron type 202, 12% of the total transformed gene expression of the *GUCA2A* gene is in this cell type, so the specificity score for *GUCA2A* in cell type 202 is 0.12. If a hypothetical gene were completely evenly expressed

---

<sup>v</sup> Limitations notwithstanding, psychiatric diagnoses are moderately reliably diagnosed across time and raters. Further, unbiased clustering approaches applied to symptom level and genetic data support the utility and validity of such diagnoses.



across all cell types, then the specificity score for that gene, in all cell types, would be  $1/461 = 0.0022$ .

### *MAGMA overview*

MAGMA (v1.10) is software designed for gene and gene set analysis of GWAS data, and it has been extensively tested to ensure appropriate control of type I errors and adjustment for potentially confounding variables<sup>26,61</sup>. MAGMA uses a regression framework and employs a two-stage procedure to test for associations, first calculating gene level p-values and then using those gene level p-values to compute p-values for collections of genes. Collections of genes can either be analyzed as gene sets (using binary coding of genes that are in or out of the set) or as 'gene properties' meaning quantitative values assigned to all genes, as we have here (i.e., specificity values for each gene, in each cell type). We employed the gene property analysis rather than arbitrarily imposing a threshold on the specificity scores to define a gene sets for each cell type. We also used the optional third stage, conditional analysis, to specify likely independent associations among all significant associations.

### *MAGMA gene level analysis*

We first used MAGMA to map each SNP to a gene if the SNP was located within 35 kilobases (kb) upstream to 10 kb downstream of that gene. We then used MAGMA's SNP-wise mean (snp-wise=mean) model to conduct gene analysis while adjusting for linkage disequilibrium (LD). LD data was from the European ancestry panel of 1000 Genomes phase 3<sup>62</sup>. In gene analysis, the test statistic of a gene was calculated as the sum of squared SNP z-statistics, where z-statistics were the probit transformation of SNP p-values from GWAS. Because the test statistic for each gene followed a mixture of independent  $\chi_1^2$  distributions under the null hypothesis, we calculated gene p-values (each representing the association between a phenotype and a gene) accordingly.

### *MAGMA gene property analysis*

MAGMA's gene property analysis represents the association that a gene has with a given phenotype as a z-score  $Z_g = \text{probit}(1 - P_g)$  where  $P_g$  is the p-value of a given gene from the gene analysis step in MAGMA. Per MAGMA default, we truncated z-scores that were 3 standard deviations below or 6 standard deviations above the mean to prevent outliers from biasing analysis results. We then conducted the gene property analysis via a linear regression model  $Z = \beta_0 + P_c \beta_1 + C \beta_2 + \varepsilon$  (eq. 1), where  $Z$  is the aforementioned z-scores of each gene,  $P_c$  is the specificity of each gene in a given cell  $c$ ,  $C$  represents the covariates, and  $\varepsilon$  is modeled as a multivariate normal accounting for the LD between genes. Per MAGMA default, specificity values were truncated if they were 5 standard deviations from the mean. In our analysis, covariates were gene size, gene density, sample size, inverse mean minor allele count, and their log values. Lastly, we conducted a one-directional test of the coefficient  $\beta_1$  as a test of the association of each cell type with each phenotype. For each phenotype (e.g., schizophrenia), this analysis was run 461 times (i.e., once for each cell type).

## MAGMA conditional analysis

To specify likely independent signals from among all significant results (i.e. all significant cell types for each phenotype), we conducted pairwise conditional analyses using MAGMA. Here, we used the linear regression model  $Z = \beta'_0 + P_{c_1}\beta'_1 + P_{c_2}\beta'_2 + C\beta'_3 + \varepsilon$  (eq. 2), which is the same as the one from gene property analysis except that the model here includes two cell types of interest. We then conducted forward stepwise selection as detailed in Watanabe et al.<sup>23</sup> to arrive at a set of “independent significant” cell types. For cell type  $c_1$  and  $c_2$ , let us denote the p-value associated with  $\beta'_1$  in eq. 2 be  $p_{c_1,c_2}$  and the one associated with  $\beta'_2$  be  $p_{c_2,c_1}$ . We also denote the marginal p-values associated with the respective cell types from gene property analysis be  $p_{c_1}$  and  $p_{c_2}$ . We define proportional significance, which portrays the remaining significance of a cell type  $c_1$  after conditioning on  $c_2$ , as  $PS_{c_1,c_2}$  such that  $PS_{c_1,c_2} = \frac{-\log(p_{c_1,c_2})}{-\log(p_{c_1})}$ . In forward stepwise selection, the set of independent significant cell types (denoted as  $S$ ) initially only contained the most marginally significant cell type. The next most significant cell type  $c$  was added to the set  $S$  in succession only if it satisfied two scenarios: First, if both  $PS_{c,s}, PS_{s,c} \geq 0.8$  for all  $s \in S$ , the associations of cell type  $c$  and any  $s \in S$  with a given phenotype were considered independent. Second, if  $0.5 \leq PS_{c,s} < 0.8$ ,  $0.5 \leq PS_{s,c} < 0.8$ , and  $p_c \leq 0.05$  for all  $s \in S$ , the associations of cell type  $c$  and any  $s \in S$  were only partially explained by each other, while the majority of the signals were independent. Cell types not included in the set  $S$  can still play an important role in the etiology of a phenotype; however, the selection procedure excluded them because their association cannot be distinguished from the association of cell types in  $S$ . Note that in some rare cases, cell type with a lower marginal significance can have a higher conditional significance. When  $PS_{c_1,c_2} < 0.2$  yet  $PS_{c_2,c_1} \geq 0.2$  for cell type  $c_1$  and  $c_2$  where  $p_{c_1} < p_{c_2}$ , the order of the selection process was reversed for the two cell types.

## Considerations regarding anatomical annotations

Throughout this manuscript, we refer to the anatomical locations of cell types. These locations were reported in Siletti et al., whose procedures involved sampling from of 106 separate human brain regions (“dissections”), and to which each of the 3,369,219 cells in their study can be traced. Note, however, that it is rare for any of the 461 cell types to be exclusively from any one dissection or brain region. Indeed, one goal of such transcriptomic surveys is to reveal the ways in which transcriptomically similar cells are found in different areas of the brain.

**Supplementary Figure 5** summarizes information about the dissections contributing to each cell type. For example, interneurons, which are inhibitory neurons found widely across the cortex (cell types #236-296 in Siletti et al.) derive from dozens of cortical (and even some subcortical) regions. Accordingly, the single dissection contributing the most cells to any of the interneuron clusters typically accounts for <10% of cells in the cluster. Conversely, cell type #202 is from just two dissections, both in the hippocampus. For brevity in the manuscript, we refer to cell types as “hippocampal” or “amygdala”, etc. when cell types derive primarily from that brain region (i.e. >50%, but typically much higher). Precise information about the dissections accounting cells in each cell type is given in **Supplementary Table 1**, which

provides the total number of cells in each cluster and the percentage of cells in that cluster that came from each contributing dissection.

### *Note about cell types / cell nomenclature*

Nomenclature for cell types will continue to evolve as the naming and categorization of cell types continues. For clarity, we have not addressed the issue of continuous gradients of difference across cell types, but this will be an important area of investigation for future studies.

### *Acknowledgements*

This work was supported by the Jaswa Innovator Award to LD and by the National Institute of Mental Health (NIMH) to LD (R01 MH123486 & R21 MH125358).

### *Author contributions*

L.D. designed the experiment, reviewed all data and analyses, drafted the manuscript, made the figures; T.L., M.S, W.L, H.S., N.S, S.V, & H.S. analyzed the data; J.Y., G.W., J.B., L.T., B.S.P., B.K., K.D, & W.J.G provided essential expertise; all authors edited and approved the manuscript.

### *Competing interests*

The authors have no competing interests.

### *Materials & Correspondence*

Correspondence should be directed to L.D. (LaramieD@Stanford.edu)

### *Data availability*

All datasets used here are publicly available from the primary publications as noted throughout the manuscript.

## REFERENCES

1. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet Psychiatry* **9**, 137–150 (2022).
2. Hyman, S. E. Revolution stalled. *Sci. Transl. Med.* **4**, 155cm11 (2012).
3. Purcell, S. M. *et al.* Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460**, 748–752 (2009).
4. Visscher, P. M., Yengo, L., Cox, N. J. & Wray, N. R. Discovery and implications of polygenicity of common disease. *Science* **373**, 1468–1473 (2021).
5. O’Connor, L. J. *et al.* Extreme Polygenicity of Complex Traits Is Explained by Negative Selection. *Am. J. Hum. Genet.* **105**, 456–476 (2019).
6. Siletti, K. *et al.* Transcriptomic diversity of cell types across the adult human brain. 2022.10.12.511898 Preprint at <https://doi.org/10.1101/2022.10.12.511898> (2022).
7. Fang, R. *et al.* Conservation and divergence of cortical cell organization in human and mouse revealed by MERFISH. *Science* **377**, 56–62 (2022).
8. Tasic, B. *et al.* Shared and distinct transcriptomic cell types across neocortical areas. *Nature* **563**, 72–78 (2018).
9. Trubetskov, V. *et al.* Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature* **604**, 502–508 (2022).
10. Saunders, A. *et al.* Molecular Diversity and Specializations among the Cells of the Adult Mouse Brain. *Cell* **174**, 1015–1030.e16 (2018).
11. Dashti, H. S. *et al.* Genome-wide association study identifies genetic loci for self-reported habitual sleep duration supported by accelerometer-derived estimates. *Nat. Commun.* **10**, 1100 (2019).
12. Substance Abuse and Mental Health Services Administration, 2021 National Survey on Drug Use and Health. Key Substance Use and Mental Health Indicators in the United States. <https://www.samhsa.gov/data/report/2021-nsduh-annual-national-report> (2021).
13. Tandon, R., Keshavan, M. S. & Nasrallah, H. A. Schizophrenia, “Just the Facts” What we know in 2008. 2. Epidemiology and etiology. *Schizophr. Res.* **102**, 1–18 (2008).
14. Tandon, R., Nasrallah, H. A. & Keshavan, M. S. Schizophrenia, ‘just the facts’ 4. Clinical features and conceptualization. *Schizophr. Res.* **110**, 1–23 (2009).
15. Kadakia, A. *et al.* The Economic Burden of Schizophrenia in the United States. *J. Clin. Psychiatry* **83**, 22m14458 (2022).
16. Hilker, R. *et al.* Heritability of Schizophrenia and Schizophrenia Spectrum Based on the Nationwide Danish Twin Register. *Biol. Psychiatry* **83**, 492–498 (2018).
17. Cardno, A. G. *et al.* Heritability Estimates for Psychotic Disorders: The Maudsley Twin Psychosis Series. *Arch. Gen. Psychiatry* **56**, 162–168 (1999).
18. Sullivan, P. F. *et al.* Psychiatric Genomics: An Update and an Agenda. *Am. J. Psychiatry* **175**, 15–27 (2018).
19. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427 (2014).
20. Pardiñas, A. F. *et al.* Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nat. Genet.* **50**, 381 (2018).
21. THE TABULA SAPIENS CONSORTIUM. The Tabula Sapiens: A multiple-organ, single-cell transcriptomic atlas of humans. *Science* **376**, eabl4896 (2022).

22. Zeisel, A. *et al.* Molecular Architecture of the Mouse Nervous System. *Cell* **174**, 999-1014.e22 (2018).
23. Watanabe, K., Umićević Mirkov, M., de Leeuw, C. A., van den Heuvel, M. P. & Posthuma, D. Genetic mapping of cell type specificity for complex traits. *Nat. Commun.* **10**, 3222 (2019).
24. Skene, N. G. *et al.* Genetic identification of brain cell types underlying schizophrenia. *Nat. Genet.* **50**, 825–833 (2018).
25. Bryois, J. *et al.* Genetic identification of cell types underlying brain complex traits yields insights into the etiology of Parkinson’s disease. *Nat. Genet.* **52**, 482–493 (2020).
26. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput. Biol.* **11**, e1004219 (2015).
27. Watanabe, K., Taskesen, E., Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* **8**, 1826 (2017).
28. Bryois, J. *et al.* Genetic Identification of Cell Types Underlying Brain Complex Traits Yields Novel Insights Into the Etiology of Parkinson’s Disease. *bioRxiv* 528463 (2019) doi:10.1101/528463.
29. Saunders, G. R. B. *et al.* Genetic diversity fuels gene discovery for tobacco and alcohol use. *Nature* **612**, 720–724 (2022).
30. International Multiple Sclerosis Genetics Consortium. Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. *Science* **365**, eaav7188 (2019).
31. Bellenguez, C. *et al.* New insights into the genetic etiology of Alzheimer’s disease and related dementias. *Nat. Genet.* **54**, 412–436 (2022).
32. Lewis, D. A., Curley, A. A., Glausier, J. & Volk, D. W. Cortical Parvalbumin Interneurons and Cognitive Dysfunction in Schizophrenia. *Trends Neurosci.* **35**, 57–67 (2012).
33. Diemel, S. J., Fish, K. N. & Lewis, D. A. The Nature of Prefrontal Cortical GABA Neuron Alterations in Schizophrenia: Markedly Lower Somatostatin and Parvalbumin Gene Expression Without Missing Neurons. *Am. J. Psychiatry* appiajp20220676 (2023) doi:10.1176/appi.ajp.20220676.
34. Yao, Z. *et al.* A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation. *Cell* **184**, 3222-3241.e26 (2021).
35. Batiuk, M. Y. *et al.* Upper cortical layer–driven network impairment in schizophrenia. *Sci. Adv.* **8**, eabn8367 (2022).
36. Lewis, D. A. & González-Burgos, G. Neuroplasticity of Neocortical Circuits in Schizophrenia. *Neuropsychopharmacology* **33**, 141–165 (2008).
37. Moberg, S. & Takahashi, N. Neocortical layer 5 subclasses: From cellular properties to roles in behavior. *Front. Synaptic Neurosci.* **14**, (2022).
38. Ho, N. F. *et al.* The Amygdala in Schizophrenia and Bipolar Disorder: A Synthesis of Structural MRI, Diffusion Tensor Imaging, and Resting-State Functional Connectivity Findings. *Harv. Rev. Psychiatry* **27**, 150–164 (2019).
39. Boules, M. M., Fredrickson, P., Muehlmann, A. M. & Richelson, E. Elucidating the Role of Neurotensin in the Pathophysiology and Management of Major Mental Disorders. *Behav. Sci.* **4**, 125–153 (2014).
40. Binder, E. B., Kinkead, B., Owens, M. J. & Nemeroff, C. B. The role of neurotensin in the pathophysiology of schizophrenia and the mechanism of action of antipsychotic drugs. *Biol. Psychiatry* **50**, 856–872 (2001).

41. Roeske, M. J., Konradi, C., Heckers, S. & Lewis, A. S. Hippocampal volume and hippocampal neuron density, number and size in schizophrenia: a systematic review and meta-analysis of postmortem studies. *Mol. Psychiatry* **26**, 3524–3535 (2021).
42. Kraguljac, N. V. *et al.* A longitudinal magnetic resonance spectroscopy study investigating effects of risperidone in the anterior cingulate cortex and hippocampus in schizophrenia. *Schizophr. Res.* **210**, 239–244 (2019).
43. Brugger, S. P. & Howes, O. D. Heterogeneity and Homogeneity of Regional Brain Structure in Schizophrenia: A Meta-analysis. *JAMA Psychiatry* **74**, 1104–1111 (2017).
44. Hansen, D. V., Hanson, J. E. & Sheng, M. Microglia in Alzheimer’s disease. *J. Cell Biol.* **217**, 459–472 (2018).
45. Cheng, Y. *et al.* Distinct Synaptic Strengthening of the Striatal Direct and Indirect Pathways Drives Alcohol Consumption. *Biol. Psychiatry* **81**, 918–929 (2017).
46. Thanos, P. K. *et al.* DRD2 gene transfer into the nucleus accumbens core of the alcohol preferring and nonpreferring rats attenuates alcohol drinking. *Alcohol. Clin. Exp. Res.* **28**, 720–728 (2004).
47. Bocarsly, M. E. *et al.* A Mechanism Linking Two Known Vulnerability Factors for Alcohol Abuse: Heightened Alcohol Stimulation and Low Striatal Dopamine D2 Receptors. *Cell Rep.* **29**, 1147–1163.e5 (2019).
48. Heinz, A. *et al.* Correlation between dopamine D(2) receptors in the ventral striatum and central processing of alcohol cues and craving. *Am. J. Psychiatry* **161**, 1783–1789 (2004).
49. Pati, D. *et al.* Dopamine D2 receptors in the bed nucleus of the stria terminalis modulate alcohol-related behaviors. 2023.06.13.544820 Preprint at <https://doi.org/10.1101/2023.06.13.544820> (2023).
50. McCullough, K. M. *et al.* Nucleus Accumbens Medium Spiny Neuron Subtypes Differentially Regulate Stress-Associated Alterations in Sleep Architecture. *Biol. Psychiatry* **89**, 1138–1149 (2021).
51. Weber, F. & Dan, Y. Circuit-based interrogation of sleep control. *Nature* **538**, 51–59 (2016).
52. van Langelaar, J., Rijvers, L., Smolders, J. & van Luijn, M. M. B and T Cells Driving Multiple Sclerosis: Identity, Mechanisms and Potential Triggers. *Front. Immunol.* **11**, (2020).
53. Kaskow, B. J. & Baecher-Allan, C. Effector T Cells in Multiple Sclerosis. *Cold Spring Harb. Perspect. Med.* **8**, a029025 (2018).
54. Ochs, J. *et al.* Proinflammatory CD20+ T cells contribute to CNS-directed autoimmunity. *Sci. Transl. Med.* **14**, eabi4632 (2022).
55. Alzheimer, A., Stelzmann, R. A., Schnitzlein, H. N. & Murtagh, F. R. An English translation of Alzheimer’s 1907 paper, ‘Über eine eigenartige Erkrankung der Hirnrinde’. *Clin. Anat. N. Y. N* **8**, 429–431 (1995).
56. Leng, F. & Edison, P. Neuroinflammation and microglial activation in Alzheimer disease: where do we go from here? *Nat. Rev. Neurol.* **17**, 157–172 (2021).
57. Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium. Genome-wide association study identifies five new schizophrenia loci. *Nat. Genet.* **43**, 969–976 (2011).
58. Vesuna, S. *et al.* Deep posteromedial cortical rhythm in dissociation. *Nature* **586**, 87–94 (2020).
59. Braun, E. *et al.* Comprehensive cell atlas of the first-trimester developing human brain. 2022.10.24.513487 Preprint at <https://doi.org/10.1101/2022.10.24.513487> (2022).
60. Consortium, T. B. *et al.* Analysis of shared heritability in common disorders of the brain. *Science* **360**, eaap8757 (2018).

61. Sey, N. Y. A. *et al.* A computational tool (H-MAGMA) for improved prediction of brain-disorder risk genes by incorporating brain chromatin interaction profiles. *Nat. Neurosci.* **23**, 583–593 (2020).
62. 1000 Genomes Project Consortium *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56–65 (2012).