Title

The impact of common variants on gene expression in the human brain: from RNA to protein to schizophrenia risk

Authors

Qiuman Liang[1†], Yi Jiang[1,2†], Annie W. Shieh[3], Dan Zhou[4,5], Rui Chen[5], Feiran Wang[1], Meng Xu[1], Mingming Niu[6], Xusheng Wang[7], Dalila Pinto[8,9,10,11], Yue Wang[12], Lijun Cheng[13], Ramu Vadukapuram[14], Chunling Zhang[15], Kay Grennan[16], Gina Giase[17], The PsychENCODE Consortium[‡], Kevin P White[18], Junmin Peng[6], Bingshan Li[5], Chunyu Liu[1,16,19*], Chao Chen[1,20,21,22*], Sidney H. Wang[3*]

Affiliations

1 MOE Key Laboratory of Rare Pediatric Diseases & Hunan Key Laboratory of Medical Genetics, School of Life Sciences, and Department of Psychiatry, The Second Xiangya Hospital, Central South University, Changsha, Hunan 410000, China.

2 Department of Epidemiology and Biostatistics, Ministry of Education Key Laboratory of Environment and Health, School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei 430000, China.

19 3 Center for Human Genetics, The Brown foundation Institute of Molecular Medicine,
20 The University of Texas Health Science Center at Houston, Houston, TX 77030, USA.

21 4 School of Public Health and the Second Affiliated Hospital, Zhejiang University School
22 of Medicine, Hangzhou, Zhejiang 310058, China.

23 5 Department of Molecular Physiology and Biophysics, Vanderbilt Genetics Institute,
24 Vanderbilt University, Nashville, TN 37232, USA.

25 6 Department of Structural Biology, Department of Developmental Neurobiology, Center
26 for Proteomics and Metabolomics, St. Jude Children's Research Hospital, Memphis, TN 38105,
27 USA.

28 7 Department of Genetics, Genomics, and Informatics, University of Tennessee Health
29 Science Center, Memphis, TN 38163, USA.

30 8 Department of Psychiatry, and Seaver Autism Center for Research and Treatment, Icahn
31 School of Medicine at Mount Sinai, New York, NY 10029, USA.

32 9 Department of Genetics and Genomic Sciences, and Icahn Institute for Data Science and
33 Genomic Technology, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA.

34 10 The Mindich Child Health and Development Institute, Icahn School of Medicine at
35 Mount Sinai, New York, NY 10029, USA.

36 11 Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY
37 10029, USA.

12 Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, Arlington, VA 22203, USA.

13 Institute for Genomics and Systems Biology, University of Chicago, Chicago, IL 60637, USA.

14 Department of Psychiatry, The University of Texas Rio Grande Valley, Harlingen, TX 78550, USA.

15 Department of Neuroscience and Physiology, SUNY Upstate Medical University, Syracuse, NY 13210, USA.

16 Department of Psychiatry, SUNY Upstate Medical University, Syracuse, NY 13210, USA.

17 The Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA.

18 Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 117596, Singapore.

19 School of Psychology, Shaanxi Normal University, Xi'an, Shaanxi 710062, China.

20 Furong Laboratory, Changsha, Hunan 410000, China.

21 National Clinical Research Center for Mental Disorders, The Second Xiangya Hospital, Central South University, Changsha, Hunan 410000, China.

22 National Engineering Research Center of Personalized Diagnostic and Therapeutic Technology, Central South University, Changsha, Hunan 410000, China.

† These authors contributed equally to this work.

‡ PsychENCODE Consortium authors and affiliations are listed in the supplementary materials.

* Corresponding author email: **Hsi.Ming.S.Wang@uth.tmc.edu** (S. H. Wang), chenchao@sklmg.edu.cn (C. Chen), LiuCh@upstate.edu (C. Liu)

Abstract

Background

The impact of genetic variants on gene expression has been intensely studied at the transcription level, yielding in valuable insights into the association between genes and the risk of complex disorders, such as schizophrenia (SCZ). However, the downstream impact of these variants and the molecular mechanisms connecting transcription variation to disease risk are not well understood.

Results

We quantitated ribosome occupancy in prefrontal cortex samples of the BrainGVEX cohort. Together with transcriptomics and proteomics data from the same cohort, we performed cis-Quantitative Trait Locus (QTL) mapping and identified 3,253 expression QTLs (eQTLs), 1,344 ribosome occupancy QTLs (rQTLs), and 657 protein QTLs (pQTLs) out of 7,458 genes quantitated in all three omics types from 185 samples. Of the eQTLs identified, only 34% have their effects propagated to the protein level. Further analysis on the effect size of eQTLs showed clear post-transcriptional attenuation of eQTL effects.

We then investigated the biological relevance of the attenuated eQTLs by identifying omics-specific QTLs, and identified 228 omics-specific QTLs (70 esQTL, 51 rsQTL, and 107 psQTL),

with five showing strong colocalization with SCZ GWAS signals and three of which were esQTL. Using both S-PrediXcan and two-sample Mendelian Randomization (MR), we identified a total 52 SCZ risk genes, 34% of which are novel. We found 67.3% of these SCZ-risk-driving eQTLs show little to no evidence of driving corresponding variations at the protein level.

Conclusion

The effect of eQTLs on gene expression in the prefrontal cortex is commonly attenuated post-transcriptionally. Many of the attenuated eQTLs still correlate with SCZ GWAS signal. Further investigation is needed to elucidate a mechanistic link between attenuated eQTLs and SCZ disease risk.

Introduction

Complex diseases such as neuropsychiatric disorders are multi-factorial with genetic components (*1*, *2*). Large scale Genome-Wide Association Studies (GWAS) have uncovered thousands of disease associated loci, signaling a promising era ahead for causal variant identification (*3*). However, efforts in fine mapping these disease risk loci often narrowed down the underlying causal signal to non-coding regions of the genome (*4–6*). Regulatory variants in such non-coding regions are therefore the prime candidates for driving the genetic risk of disease etiology. Consequently, integrating gene expression information to pinpoint causal variants or to identify risk genes has become a staple of genetic studies of complex diseases (*7*), with multiple consortia efforts facilitating large scale gene expression profiling and regulatory element mapping (*8–10*). Many powerful methods, such as *coloc*, PrediXcan, SMR/HEIDI, to name a few, have also been developed to leverage gene expression information for fine mapping GWAS signals or for identifying underlying risk genes (*11–13*).

103 Schizophrenia (SCZ) is a psychiatric disorder affecting ~1% of the world-wide population (*14*).
104 The heritability of SCZ has been estimated at between 60% to 80% indicating a strong genetic
105 component (*15*, *16*). Accordingly, recent schizophrenia GWAS study reported by Trubetskoy et
106 al. identified 287 significant risk loci and prioritized 120 risk genes using functional genomic data
107 (*17*). The use of RNA-Seq data from the brain was instrumental for risk gene prioritization by
108 Trubetskoy et al.; however, information from downstream gene regulation processes, such as
109 translation rate and protein abundance, was either not utilized or unavailable.

110 Measuring transcriptional changes as a proxy for gene activity has a long history in molecular
111 biology (*18*). In the context of human genetics, buffering of downstream effects of genetic
112 variants impacting gene expression (i.e. an eQTL) has been shown to be prevalent (*19*). In
113 addition, QTLs specific to protein level have also been reported (*19-21*). Together, these
114 observations indicated the importance and potential benefits of including downstream omics
115 types, such as proteomics data, as information sources for fine mapping disease regulatory
116 variants. Indeed, recent studies using multi-omics approaches have demonstrated increased power
117 for risk gene identification among other benefits (*22*, *23*). Of note, our recent work on genetic
118 variants associated with protein level in prefrontal cortex of the human brain indicated the extent
119 of contribution from non-synonymous coding variants to changes at the protein level, and the
120 utility of these protein QTL variants in prioritizing GWAS risk genes for psychiatric disorders
121 (*21*).

122 Another potential benefit of taking a multi-omics approach for identifying disease risk genes rests
123 in the potential to dissect the fine details of regulatory mechanisms driving the disease-genotype
124 association. Having relevant datasets to illuminate the origin and propagation of genetic impact
125 could potentially arrive at a conclusion of the driver regulatory process for a risk gene. Ribo-seq
126 is a technology that could be used to collect relevant data to fill in the gap between transcript and

127 protein expression. By adapting RNA-Seq to a ribosome footprinting method, ribo-seq provides
128 transcriptome-wide quantification of ribosome occupancy (*24*, *25*), which can serve as a proxy for
129 the amount of active translation synthesizing proteins from each mRNA transcript. When
130 analyzed in conjunction with RNA-Seq and quantitative proteomics data, ribo-seq enables
131 identification of translational and post-translational regulatory events (*19, 26*), both major steps
132 defining the human genetics aspect of the Central Dogma of molecular biology.

133 As a part of the consortium efforts to improve our understanding of the genetic basis of
134 neuropsychiatric disorders (*27*), we generated multiple data modalities that included SNP
135 genotyping, RNA-Seq, ribo-seq, and proteomics of postmortem cortical tissue samples of the
136 BrainGVEX cohort, which altogether covered multiple omics levels from DNA, transcript to
137 protein. In conjunction with the quantitative proteomics and transcriptome profiling results that
138 we previously published (*21*, *28*), here we integrated ribo-seq data as our operational definition
139 for protein translation based on the transcriptome to make it a true multi-omics investigation. Our
140 results reveal regulatory properties of common variants in the human brain and their utility in
141 identifying the regulatory processes driving disease risk for schizophrenia. It offers an opportunity
142 from a population angle to dissect and appreciate the regulatory information flow in the biological
143 processes associated with the Central Dogma. Additional "rules" including information
144 attenuation and modification regulation in the process are recognized.

145 **Results**

146 *Measuring transcriptome-wide ribosome occupancy level in prefrontal cortex of adult human*
147 *brain to quantify the level of protein translation*

148 To investigate regulatory impact of genetic variants on protein translation in the prefrontal cortex
149 of the human brain, we performed ribosome profiling on 211 prefrontal cortex samples from the

150 BrainGVEX collection. In total we collected ~62 billion ribosome profiling reads. Consistent with
151 the expected ribosome footprint size, we found the average insert size of our ribo-seq libraries to
152 range between 27.4 and 29.5 nucleotides. Similar to prior published studies (*29*), we found on
153 average 74 % of unwanted reads from ribosomal RNA, tRNA, and snoRNA, which contributed
154 no information to the translation of protein coding genes. After removing these unwanted reads,
155 we obtained an average of 30.3 million uniquely mapped informative reads per sample (inter-
156 quartile range: 20.5 ~ 37.6 million reads). When focusing our analysis on the informative reads,
157 we found our ribo-seq data to have ~3 times higher proportion of exon reads than that of the total
158 RNA-Seq data collected from the same individuals using an rRNA-depletion method (*28*). When
159 visualized in aggregate across annotated coding genes, we found our ribo-seq data to show strong
160 sub-codon periodicity at the expected positions (Fig. S1). High proportion of exon reads and
161 strong sub-codon periodicity reflect the enrichment of footprints from ribosomes actively engaged
162 in translating mature mRNA and indicates the quality of the dataset.

163 *Multi-omics cis-QTL mapping identified candidate regulatory variants and revealed translational*
164 *and post-translational attenuation of eQTL effects*

165 To identify variants associated with inter-individual expression differences, we perform cis-QTL
166 mapping for each data type independently. Using the full dataset (i.e. 416 RNA-Seq samples, 195
167 ribo-seq samples, and 268 proteomics samples, which are coupled with the corresponding
168 genotype data), we identified 12,411 eQTLs (out of 16,540 genes we deem sufficiently
169 quantitated), 2,849 rQTLs (out of 14,572 genes), and 1,036 pQTLs (out of 8,013 genes) at FDR <
170 0.1. The majority of the eQTLs identified here were replicated in the prefrontal cortex RNA-Seq
171 data from the GTEx consortium (Table S1). Intriguingly, we found drastic differences between
172 omics types in the numbers of QTLs mapped, suggesting that some of the eQTL effects do not
173 propagate all the way to the protein level. However, the differences in the number of genes tested

174 between omics types and the differences in sample size make the comparison challenging to
175 interpret. To better compare the effects of genetic regulation between multiple data modalities, we
176 identified 185 samples with 7,458 genes that were sufficiently quantitated across all three
177 datatypes. Using this unified dataset, we found 3,253 eQTLs, 1,344 rQTLs, and 657 pQTLs at
178 FDR <0.1 (Fig. 1A, Fig. S2, Table S2, S3, S4). Similar to the results derived from the full dataset,
179 using the unified dataset we found fewer significant QTLs as we moved downstream the Central
180 Dogma of molecular biology.

181 A challenge in comparing between the numbers of QTLs identified from each omics type rests in
182 the fact that not all true effects were identified. Tests replicating QTLs identified from one omics
183 type in the other omics types can better capture the proportion of genetic effects shared between
184 QTL types. We performed replication tests using $\pi 1$ estimates from the qvalue method (*30*).
185 Overall, we found high proportion of QTLs replicated in other omics types (Fig. 1B). However,
186 when considering the replication rates with the direction of genetic information flow, we found
187 asymmetric replication rates, with the downstream omics types to replicate less than the upstream
188 omics types. More specifically, we found 84.7% of the rQTLs were replicated at the transcript
189 level, but only 60.2% of the eQTLs were replicated at the ribosome occupancy level. Moreover,
190 while 75.9% of the pQTLs were replicated at the transcript level, only 34.0% of the eQTLs were
191 replicated at the protein level (Fig. 1B). The lower percentage of eQTLs and rQTLs replicated at
192 the protein level indicates potential effect attenuation (i.e. either the inter-individual variation in
193 gene expression becoming smaller and therefore harder to detect or a lack of such effect in the
194 downstream omics types). Interestingly, in addition to the effect attenuation at the protein level,
195 which was previously reported for lymphoblastoid cell lines (LCLs) (*19*), here, using brain
196 samples, we found further asymmetry between eQTL and rQTL replication, indicating effect
197 attenuation at the level of translation. A similar asymmetry in proportion replicated between

198 upstream and downstream omics types was observed when using a direction-aware cutoff-based
199 approach across a wide range of significance cutoffs used to define replication rates (Fig. S3).

200 While our replication tests revealed a trend of effect attenuation for eQTL variants in the
201 downstream phenotypes (Fig. 1B), these same observations could alternatively be explained by
202 differences in statistical power between technologies. An independent piece of evidence that is
203 not sensitive to measurement precision is needed to reach a solid conclusion. Using eQTLs
204 independently identified from prefrontal cortex samples by the CommonMind Consortium (CMC)
205 (*31*), we avoid the ascertainment bias for large effect size eQTLs identified from our dataset and
206 can therefore directly compare the effects size of eQTL variants between the three omics types in
207 our dataset. A similar approach was successfully implemented to address this power issue in prior
208 work in LCLs (*19*). Using 5,915 CMC eQTLs that were also quantitated in our dataset, we found
209 the eQTL effects on the transcript expression level to be significantly larger than their effects on
210 ribosome occupancy level (per allele log2 fold differences: mRNA 0.2433 [95% CI=
211 0.2381~0.2486] vs. ribosome occupancy 0.1836 [95% CI= 0.1794~0.1878]), which were in turn
212 significantly larger than their effects on protein level (per allele log2 fold differences: 0.1486
213 [95% CI= 0.1451~0.1522]) (Fig. 1C, t-test $P < 2.2e^{-16}$ for all pairwise comparisons). By focusing
214 on the effect sizes of independently identified eQTLs, our results strongly support the presence of
215 downstream mechanisms attenuating eQTL effects both at the ribosome occupancy level
216 (translationally) and at the protein level (post-translationally). Moreover, for the CMC eQTLs, we
217 found translational regulation to account for more effect size reduction than post-translational
218 regulation (Fig. 1C).

219 *Identifying omics specific QTLs and their signal colocalization with schizophrenia GWAS*

220 The prevalent effect size reduction of eQTLs raised the question of the relevance of these genetic
221 regulation at the organismal level. Because most cellular tasks are executed by proteins, the
222 genetic regulatory effects not reaching the protein level are less likely to have an impact on
223 organismal traits. To answer this question, we set out to investigate the relevance of different
224 QTL types in SCZ. More specifically, we aim to identify expression specific QTLs (i.e. genetic
225 variants that impact transcript level of the linked genes but not the downstream ribosome
226 occupancy level nor protein level) that colocalize with SCZ GWAS.

227 Consistent with prior reports (*32*), using our full dataset we found significant proportion of SCZ
228 heritability to be mediated by gene expression. By performing mediated expression score
229 regression (MESC) (*32*) on summary statistics from the Trubetskoy et al. SCZ GWAS (*17*), we
230 found our eQTLs to mediate 7.09%, rQTLs to mediate 4.06%, and pQTLs to mediate 2.17% of
231 SCZ heritability (Table S5). After establishing the relevance for each of the three QTL types in
232 SCZ, we next sought to identify omics-specific QTLs in order to further evaluate their relevance
233 in driving SCZ risk. Because regulation of a gene is often modulated by multiple genetic variants,
234 to evaluate the consequence of overall *cis*-QTL impact on gene expression, we use PrediXcan to
235 estimate aggregate genetic regulatory effects for each gene. To distinguish between genes that
236 have QTL effects shared across multiple omics types and genes that have omics-specific QTL
237 effects, for each omics type we built PrediXcan models with or without regressing out the other
238 omics types and computed the correlation between the imputed expression from the two models.
239 We termed this correlation R_c. To identify R_c^2 values that reflect significant sharing between
240 omics types, we permuted sample labels to emulate conditions of no real correlation between the
241 three omics types in order to generate empirical null distributions. Using both a false discovery
242 rate (FDR), which is calculated based on the empirical null, and an effect size cutoff based on R_c^2,
243 we defined a set of shared QTL genes and three sets of omics-specific QTL genes. At 10% FDR,
244 we defined shared QTL genes by further requiring the R_c^2 to be smaller than 0.5. Using these

245 criteria, from the 1,354 genes that passed the minimum PrediXcan criteria for being included in
246 this analysis (see details in methods), we identified 295 shared QTL genes that have QTL effects
247 shared between at least two omics types. For genes that failed to reject null at the 10% FDR cutoff
248 (i.e. potentially omics-specific) we further set a conservative $R_c^2 > 0.9$ cutoff to define omics-
249 specific QTLs for those QTLs that did not change after regressing out effects from other omics
250 types. Using these criteria, we found 70 esQTL (mRNA-specific QTL) genes, 51 rsQTL
251 (ribosome occupancy-specific QTL) genes, and 107 psQTL (protein-specific QTL) genes (Fig.
252 S4, Table S6).

253 To investigate the relevance of omics-specific QTL genes in SCZ, we performed summary
254 statistics-based signal colocalization between QTL signals and SCZ GWAS signals. Using *coloc*
255 with default prior (*11*), at a posterior probability cutoff of 70% for the signal colocalization
256 hypothesis, we found esQTLs of 3 genes, *CCDC117, GATAD2A,* and *JAKMIP2* to colocalize
257 with SCZ GWAS signals at loci *22q12.1*, *19p13.11,* and *5q32*, respectively (Fig. 2A, Fig. 2B). In
258 addition, we found rsQTLs of *UGGT2* to colocalize with SCZ GWAS signals at locus *13q32.1*
259 (Fig. 2C, Fig. 2D) and psQTLs of *P2RX7* to colocalize with SCZ GWAS signals at locus
260 *12q24.31*. On the other hand, for shared QTL genes, we found the eQTLs of 6 genes and the
261 rQTLs of 1 gene to colocalize with SCZ GWAS signals. In summary, we identified strong signal
262 colocalization with SCZ GWAS from both shared QTLs and omics-specific QTLs, at comparable
263 proportions (i.e. 7 from 295 of shared vs. 5 from 228 of omics-specific). This indicates that the
264 omics-specific QTLs are equally important in explaining SCZ GWAS signals.

265 *Functional genomics identification of Schizophrenia risk genes*

266 To further investigate the relevance of attenuated eQTLs in SCZ risk, we next took a
267 complementary approach by first identifying risk genes from each omics type separately and then

268 investigating the relevant regulatory processes driving SCZ risk. Following the observed
269 percentages of SCZ heritability mediated by gene regulation found in our full dataset, here using
270 the same GWAS we focused our effort to identify risk genes for SCZ based on our unified multi-
271 omic dataset using S-PrediXcan (*33*). At 5% family-wise error rate, we found 52, 29, and 16 SCZ
272 risk genes, respectively from RNA-Seq, ribo-seq, and proteomics data (Fig. 3A, Fig. S5; note that
273 the color and shape code of this figure will become relevant in the next results section). Among
274 them only four genes, *NEK4*, *KHK*, *CNNM2* and *DARS2* were consistently identified as SCZ risk
275 genes from all three omics types (Fig. 3B). The majority (74.3%) of the risk genes were identified
276 only from one of the three omics types. This limited sharing in risk gene identification between
277 omics types is in clear contrast to the amount of signal sharing found between QTL types (Fig.
278 1B).

279 Among the 74 risk genes we identified using S-PrediXcan (i.e. the union between the risk genes
280 identified from each of the three omics types), 44 have previously been reported in GWAS studies
281 as either the mapped genes or as one of the nearby genes under the GWAS peak (Table S7) (*17*,
282 *34–36*). Of these previously reported genes, 27 matched the mapped genes while the remaining 17
283 nominated an alternative candidate gene for each GWAS locus (Note that for some of these 17
284 loci, the original GWAS signal was mapped to an intergenic region). Comparing our results to
285 other published SCZ risk gene identification studies, we found 15 matched to the risk genes
286 identified by Giambartolomei et al., which used RNA-Seq and DNA methylation data from
287 prefrontal cortex of the human brain (*22*) and 12 matched to the 120 prioritized SCZ risk genes
288 from Trubetskoy et al., which used colocalization with eQTL and Hi-C data (*17*). On the other
289 hand, for 25 of our 74 risk genes (33.8%), we found no match to the known risk gene list (Table
290 S8), which we compiled based on previous GWAS and functional genomics risk gene
291 identification studies (*17*, *22*, *34–36*). These "no match" novel risk genes have relatively weak
292 SCZ GWAS signals and are therefore challenging to identify without the additional information

293 provided in our multi-omics QTL dataset. For example, we found strong colocalization between a
294 modest schizophrenia GWAS signal at *2p23.3* and all three types of molecular QTLs of the gene
295 *KHK* (Fig. 3D). Intriguingly, the *KHK* pQTL is in opposite direction of the *KHK* eQTL and rQTL
296 (Fig. 3C), suggesting a linked post-translational process regulating the protein level in addition to
297 the transcriptional regulation. *KHK*, known as Ketohexokinase, plays a pivotal role in fructose
298 metabolism and has been hypothesized to contribute significantly to sustaining neuronal function
299 (*37*). Another novel SCZ risk genes, *BTN3A2*, belongs to the Butyrophilin Subfamily 3.
300 Overexpression of *BTN3A2* has been observed to inhibit excitatory synaptic activity onto CA1
301 pyramidal neurons (*38*). *NSF* or N-Ethylmaleimide Sensitive Factor, which encodes a vesicle
302 fusing ATPase, has been identified as a causal factor in intelligence traits (*39*).

303 *Analyses of multi-omics dataset reveal regulatory mechanisms of schizophrenia risk genes*

304 While PrediXcan is a powerful tool for GWAS risk gene identification, it does not control for
305 potential horizontal pleiotropy (*40*). To this end, we performed two-sample Mendelian
306 Randomization (MR) with Egger regression to replicate the risk genes we identified using S-
307 PrediXcan. Egger regression includes an intercept term, which can be used to evaluate the level of
308 horizontal pleiotropy (*41*). For each risk gene we first used LD-clumping (*42*) to identify top
309 SNPs from independent signals as strong genetic instruments (*43*). We then tested for the causal
310 relationship between gene regulation (i.e. QTL signal) and SCZ (i.e. GWAS). We used an
311 operational definition of a causal effect based on the MR test results (see Methods). At 10% FDR,
312 of the 97 gene-by-omics combinations (i.e. a total of 74 risk genes including some discovered
313 from more than one omics type), 67.0% (65/97) passed the MR test. Of those 33.0% that failed
314 the MR test, 93.8% (30 out of 32) failed because of horizontal pleiotropy identified by the Egger
315 intercept test (Table S9). A total of 52 genes were replicated in at least one of the three omics
316 types. Similar, but stronger, causal effects were observed when using SuSiE (the Sum of Single

317 Effects) (*44*) as an alternative approach to define instrument SNPs (Fig. S6). However, when
318 using SuSiE for selecting instrument variables, 41 genes were tested because some of the risk
319 genes had no fine mapped QTL SNPs according to SuSiE.

320 A key strength of using a multi-omics QTL approach to identify GWAS risk genes rests in the
321 possibility of further narrowing down the potential regulatory mechanisms. To this end, we
322 further examined the likely causal mechanisms for the 52 replicated SCZ risk genes using one-
323 sample Mendelian Randomization to infer causality between QTL types. We focused our analysis
324 on independently testing for causal effects between neighboring QTL types following the
325 direction of information flow of the Central Dogma (i.e. mRNA -> ribosome occupancy, and
326 ribosome occupancy -> protein). Here we used fine-mapped QTL SNPs identified from the
327 exposure omics types (i.e. the upstream omics types) as instrument variables for one-sample MR
328 analysis. Among the 52 two-sample MR replicated SCZ risk genes, we found 17 genes with
329 significant causal effects both from eQTL to rQTL (i.e. the upstream pathway) and from rQTL to
330 pQTL (i.e. the downstream pathway) (both-passed risk genes; Fig. 3A dark red solid circle
331 datapoints). Significant causal effects detected from both pathways suggest transcriptionally
332 regulated protein level differences as the potential mechanisms for these risk genes in SCZ
333 etiology (see an example in Fig. 4A, Fig. 4B). On the other hand, 27 and 8 replicated risk genes
334 have either significant causal effects only in one of the two pathways (single-passed risk genes;
335 Fig. 3A blue triangle datapoints) or have no significant causal effects (none-passed risk genes;
336 Fig. 3A orange rectangle datapoints), respectively.

337 Of note, for the 27 single-passed risk genes, we found significant causal effects only in the
338 upstream pathway (mRNA -> ribosome occupancy). This asymmetry is reminiscent of the eQTL
339 effect attenuation described in the prior sections. A failed test could indicate either a true lack of
340 effect or a lack of statistical power. To take a closer look, we directly assessed the effect size, the

341 noise level, and instrument strength of the one-sample MR test results. When comparing to the 17
342 both-passed risk genes, we found significantly smaller effect sizes both for failed tests of the 27
343 single-passed risk genes (only the downstream MR test results included, average 0.367 vs 0.055,
344 t-test $P < 4.9e^{-5}$, Fig. S7A) and the 8 none-passed risk genes (both upstream and downstream MR
345 test results included, average 0.405 vs 0.033, t-test $P < 4e^{-12}$, Fig. S8A). Note that in both cases
346 the inter-quartile range of the estimated causal effects for the failed tests covered zero (Fig. S7A,
347 Fig. S8A). On the other hand, for these same comparison groups, we found no significant
348 difference in instrument strength (Fig. S7C, S8C, if anything slightly stronger instrument was
349 observed in failed tests) and found only slightly higher noise level in the failed MR tests (Fig.
350 S7B and S8B; t-test P = 0.026 and P= 0.013 respectively). The estimated small causal effects
351 from the failed tests indicate that either we are observing weak effects that are obscured by the
352 slightly elevated noise level (false negatives) or a lack of true causal effects, or a mixture of the
353 two. In other words, some of these risk genes are likely to be driven by specific QTL types. Case
354 in point, we found *SF3B1* to have similar patterns in p value distributions between eQTLs and
355 SCZ GWAS but no clear QTL signals in either ribo-seq or proteomics data (Fig. 4C, Fig. 4D).

356 **Discussion**

357 Using a panel of postmortem prefrontal cortex samples, we found clear evidence of post-
358 transcriptional attenuation of eQTL effects in the human brain. Many of the differences found
359 between individual brain transcriptomes were not present between individual brain proteomes.
360 This observation echoes earlier work in HapMap lymphoblastoid cell lines (*19*) and extends the
361 prior conclusion from *in vitro* cell lines to complex human tissues. Importantly, distinct from the
362 earlier work in lymphoblastoid cell lines, which found translation to mostly track with
363 transcription, we found clear attenuation of eQTL effect in ribo-seq data indicating that
364 translational regulatory processes are involved in eQTL effect attenuation in the human brain.

365 Prevalent translational attenuation of variant impact on transcriptional gene expression level has
366 previously been reported in budding yeast (*45*, *46*). However, results from follow up studies (*47*)
367 appear to present an inconsistent picture. Here, using replication tests for individual eQTLs and
368 testing for aggregate effect size of eQTLs independently identified from CMC, we provided
369 strong evidences supporting prevalent translational attenuation of eQTL effects in the human
370 brain. Although operating at the molecular level, our study remained observational. Omics-
371 specific features could potentially confound the results. Future work replicating this observation
372 and elucidating molecular mechanisms of translational attenuation of eQTL effects are needed to
373 provide a clear understanding of the phenomenon.

374 Following this observation, our current work focuses on the important question of whether the
375 attenuated eQTLs are functionally (biologically) relevant. We attempted to address this question
376 by exploring the relevance of attenuated eQTLs in SCZ, a neuropsychiatric disorder that is highly
377 heritable. We took two approaches to identify risk genes that have either omics-specific QTL
378 signals or attenuated eQTLs. In the first approach, we used PrediXcan to aggregate variant impact
379 at the gene level in order to identify omics-specific risk genes. Our method is distinct from
380 published work on omics-specific QTLs that took a SNP-based approach (*19*). Using a gene-
381 based approach, we aimed to increase the interpretability of the results and decrease the challenge
382 of the needle in the haystack problem. Indeed, our approach reduced our search space to 1,354
383 genes, and among these genes we identified 228 omics-specific QTL genes. By limiting our
384 search to the most confident set of omics-specific QTL genes, we identified three clear examples
385 of esQTLs that show strong colocalization with SCZ GWAS signals. At the same time, the
386 limited number of discoveries put certain constraints on our ability to investigate the properties of
387 esQTLs in the context of SCZ disease risk. It is expected that larger samples would reveal more
388 esQTLs and enable deep mechanistic investigation.

389 In the second approach, we expanded our search by first identified risk genes from all quantitated
390 genes using a TWAS approach, S-PrediXcan method. We then replicated the TWAS risk genes
391 using two-sample Mendelian Randomization with Egger intercept tests. While the majority of the
392 TWAS risk genes were replicated in two-sample Mendelian Randomization tests, of the minority
393 that failed the MR tests most failed because of the Egger intercept test. Our results, therefore,
394 clearly confirmed the need of controlling for potential horizontal pleiotropy in TWAS studies
395 (*41*). About 34% of the risk genes that we identified here have not been previously reported as
396 SCZ risk genes. These novel SCZ risk genes tend to have weak GWAS signals and are therefore
397 challenging to identify without the help of functional genomics data (see Fig. 3D for an example).
398 To identify causal regulatory mechanisms for each replicated risk gene, we further tested
399 causality between QTL types using one-sample MR. We found 17 risk genes that are likely
400 contributing to SCZ risk through transcriptionally regulated protein level. On the other hand, we
401 also identify 8 genes that show no significant MR tests (candidate omics-specific risk genes) and
402 27 genes that have significant MR test results only from transcription to translation (candidate
403 post-translationally-attenuated risk genes).

404 In essence, attempts to identify esQTLs (or any omics-specific QTLs) are dealing with the
405 challenge of separating true negatives from false negatives. While we applied very conservative
406 criteria to identify esQTLs and performed subsequent evaluations, such caveat is important to
407 keep in mind when interpreting our results. Similarly, the interpretation of the failed MR tests is
408 challenging. Our subsequent analyses looking at comparing instrument strength, noise level, and
409 effect size between passed and failed MR tests, indicated comparable instrument strength, slightly
410 elevated noise level in the failed tests, and clearly smaller effect size in the failed tests. In other
411 words, the failed tests are reflecting either a smaller effect size obscured by noise in the data, or a
412 true lack of causal effect, or possibly a mixture of both. In addition to the issues with false
413 negative results in replication omics types, some of the omics-specific QTL discoveries could be

414 false positives. Although we have good confidence with our FDR and FWER estimates, given
415 that our test statistics for QTL mapping and risk gene identification are well calibrated (Fig. S2,
416 S5), pleiotropy could introduce positive results from a different underlying cause (i.e. true effects
417 on SCZ risk but false positive risk genes).

418 Power issues, however, do not explain the whole story. As was consistently observed throughout
419 our study, when viewed in aggregate, we see clear effect size differences between omics types,
420 both for QTLs and for causal effects from MR tests. These effect size estimates are not influenced
421 by significance cutoffs and are not biased by power differences. Such general trends are also
422 unlikely to simply be a result of spurious associations or made up entirely of false positives. Our
423 results therefore bring forth an interesting mechanistic question: how do attenuated eQTL variants
424 impact SCZ without changing protein levels? One interesting possibility is that the biologically
425 relevant traits here are translation efficiency (i.e. protein synthesis rates) and protein turnover
426 rates. For example, an association between genotype and translation efficiency could manifest in
427 the form of an attenuated eQTL, where the differences in translation rate appears to offset the
428 differences introduced at the transcript level, which in turn resulted in a lack of association
429 between eQTL SNP genotype and ribosome occupancy level. In other words, the colocalization
430 between an attenuated eQTL and GWAS signal could be reflecting a colocalization of GWAS
431 signal with a translation efficiency QTL. Similarly a protein turnover QTL could also manifest in
432 the form of a rQTL attenuated at the protein level. This hypothesis would predict additional
433 linked regulatory variants (i.e. linked to the attenuated eQTLs) that impact translation efficiency
434 or protein turnover rates. Moreover, the effect at the protein level may be missed at the pQTL due
435 to technical issues specific to proteomics. We hope by presenting the current results, our findings
436 can inspire future studies on this topic to understand the detailed regulation processes from DNA
437 to RNA, protein and diseases.

438 **Materials and Methods**

439 **Data sources**

440 The SNP genotypes (*21*), RNA-Seq (*28*), and quantitative mass spectrometry (*21*) data generated
441 for prefrontal cortex tissue samples of the BrainGVEX cohort were downloaded from the
442 PsychENCODE Synapse portal (https://www.synapse.org/#!Synapse:syn5613798) (See Table
443 S10 for a summary on the number of samples in each dataset and their respective overlap with the
444 samples in the genotype data). The BrainGVEX cohort includes 420 Caucasians, 2 Hispanics, 1
445 African American, 3 Asian Americans, and 14 unknown-origin individuals. We also used PGC3
446 SCZ GWAS summary statistics data obtained from the Psychiatric Genomics Consortium (*17*).

447 **Ribosome profiling**

448 Ribosome Profiling experiments were performed using Illumina's TrueSeq Ribo Profile
449 (Mammalian) Kit. TrueSeq Ribo Profile (Mammalian) Kit was developed for cell lines. We
450 adapted the protocol to frozen tissue samples with MP Biomedicals™ FastPrep-24™ Classic
451 Bead Beating Grinder and Lysis System. Specifically, 60-80 mg of frozen brain tissue was
452 homogenized in Lysing Matrix D tubes containing 800 µl polysome lysis buffer. The Lysing
453 tubes were placed on the FastPrep-24™ homogenizer set at 4.0 m/s with 20 s increments until no
454 visible chunks of tissue remained. Tissue lysate was incubated on ice for 10 min followed by
455 centrifugation at 20,000g at 4 °C for 10 minutes to pellet cell debris. Clarified lysate was
456 subsequently used for ribo-seq library preparation following TrueSeq Ribo Profile (Mammalian)
457 Kit instructions. Indexed libraries were pooled and then sequenced on an Illumina HiSeq 4000.
458 Note that the experimental protocol for TrueSeq Ribo Profile (Mammalian) Kit that we followed
459 is a modified version of the previous ribo-seq protocol published by Ingolia and colleagues (*25*),
460 and it has the following key modifications: Monosome isolation was performed using Sephacryl

461 S400 spin columns (GE; 27–5140-01) on a tabletop centrifuge instead of ultra-high speed
462 centrifugation in sucrose cushion. Ribosomal RNA depletion was carried out by using Ribo-Zero
463 Magnetic Kits and this step is moved up to right after the monosomal RNA isolation step and
464 before the Ribosome Protected Fragment gel isolation step.

465 **Data processing, gene expression quantification, and normalization**

466 For RNA-Seq data, we obtained the FASTQ raw data from the PsychENCODE BrainGVEX
467 project (https://www.synapse.org/#!Synapse:syn5613798). Then we used cutadapt (v1.12) to trim
468 adapter for raw reads with code "cutadapt –minimum-length=5 –quality-base=33 -q 30 -a
469 AGATCGGAAGAGCACACGTCTGAACTCCAGTCA -A
470 AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT". Then we mapped trimmed reads onto
471 GENCODE Release 19 (GRCh37.p13) genome with same version's GTF by STAR (v2.4.2a).
472 Then we used RSEM software (v1.2.29) to quantify the read counts for each gene (*48*). The cpm
473 function in the R package "limma" was used to calculate the log-transformed counts per million
474 (CPM). We filtered out the genes with CPM < 1 in more than 75% samples and the samples with
475 network connectivity (*49*) z score < -5 (Fig. S9), which resulted in a total of 17,207 genes from
476 426 samples in the quantification table. We then used normalize.quantiles function in the R
477 package "preprocessCore" (*50*) to normalize expression level for each sample. We used DRAMS
478 software to detect and correct mixed-up samples (*41*), which resulted in final 423 samples.

479 For ribo-seq data, we used cutadapt (v1.12) to trim adapter for raw reads with code "cutadapt -a
480 AGATCGGAAGAGCACACGTCT –quality-base=33 –minimum-length=25 –discard-
481 untrimmed". The trimmed reads were then mapped against a FASTA file of rRNA, tRNA, and
482 snoRNA sequence using bowtie2 (*52*) to filter out uninformative reads. The filtered reads were
483 mapped to GENCODE Release 19 (GRCh37.p13) genome with corresponding transcript model

484 GTF file using STAR (v2.4.2a). The uniquely mapped reads, as defined by the "NH:i:1" flag of
485 the alignment files, were kept for subsequent analysis. We used the featureCounts function in the
486 R package "subreads" (*53*) to calculate gene level read counts for ribosome occupancy. The cpm
487 function in the R package "limma" was used to calculate log-transformed CPM value. We filtered
488 out genes with CPM < 1 in more than 75% samples and the samples with network connectivity
489 (*49*) z score < -3.5 (Fig. S10), which resulted in a total of 15,171 genes quantitated from 209
490 samples in the quantification table. We then used the normalize.quantiles function in the R
491 package "preprocessCore" (*50*) to normalize ribosome occupancy level for each sample. We used
492 DRAMS software to detect and correct mixed-up samples (*51*), which resulted in 199 samples.

493 For quantitative mass spectrometry data, we obtained protein quantification table from the
494 PsychENCODE BrainGVEX Synapse portal (https://www.synapse.org/#!Synapse:syn5613798).
495 This table includes abundance quantification for 11,572 proteins from 268 samples. The data
496 processing steps for producing the mass spectrometry quantification table is detailed in Luo et al.
497 (*21*). We further log-transformed protein abundance for each sample. We filtered out genes with
498 log-transformed protein abundance < 1 in more than 75% samples and the samples with network
499 connectivity (*49*) z score < -6, and found no gene and sample were filtered out. We then used the
500 normalize.quantiles function in the R package "preprocessCore" (*50*) to normalize protein level
501 for each sample. We used DRAMS software to detect and correct mixed-up samples (*51*), which
502 found no mixed-up samples. We matched the protein ID to gene ID according to the UCSC
503 database of hg19 version, which resulted in 8,330 genes.

504 **QTL mapping**

505 *Estimating and adjusting for unwanted factors*

506 We used the R package "PEER" (*54*) to estimate hidden factors for RNA-Seq, ribo-seq, and
507 quantitative mass spectrometry data separately. The principle of selecting unwanted hidden
508 factors was to maximize the variance explained with the least number of factors. We identified
509 30, 29, and 19 hidden factors to remove from RNA-Seq, ribo-seq, and mass spectrometry data,
510 respectively (Fig. S11). For each gene from each omics type, we adjusted the expression level by
511 fitting the selected hidden factors as predictors in a linear model and taking the residuals as the
512 adjusted expression level. The adjusted expression levels were then further centered by mean and
513 scaled by standard deviation.

514 *Genotype association tests*

515 We identified *cis*-region expression QTLs (*cis*-eQTLs), ribo-seq QTLs (*cis*-rQTLs), and protein
516 QTLs (*cis*-pQTLs) separately using QTLTools (1.3.1) (*55*). Because each gene can encode
517 several protein isoforms, we selected the protein isoform with the highest median abundance as
518 the representative protein. For each gene, we defined cis-region as the region ranging from 1Mb
519 upstream of the 5' end of the gene to 1Mb downstream of the 3' end of the gene (i.e. the length of
520 the gene body plus 2 Mb in size). We tested all common SNPs (MAF > 0.05) within the cis-
521 region using 10,000 permutations to create empirical null distributions and used the beta-
522 approximation approach implemented in QTLTools to estimate the empirical p values. For each
523 gene, we selected the SNP with the most significant empirical p value from the genotype-
524 expression level association tests to represent the QTL signal. To calculate a genome-wide FDR,
525 we used the qvalue function of the R package "qvalue" for multiple testing correction and set a
526 qvalue < 0.1 (i.e. 10% FDR) cutoff to identify significant QTLs.

527 **Estimating mediated SCZ heritability of each omics**

528 We used MESC to estimate the proportion of SCZ heritability mediated by different omics (*32*).
529 In the first step, we calculated the overall expression score using the unwanted-factor-adjusted-
530 expression data (see QTL mapping section) as the individual-level gene expression data with the
531 corresponding BrainGVEX genotype data and the 1000 genome phase 3 genotype data as the
532 ancestry matched genotype data for GWAS. In the second step, we used the overall expression
533 score calculated in the previous step and PGC3 SCZ GWAS summary statistics to estimate SCZ
534 heritability mediated by each omics.

535 **Identifying omics-shared and omics-specific QTL genes**

536 *Building PrediXcan gene expression prediction models*

537 We used the PrediXcan software (*13*) to separately build gene expression prediction models for
538 RNA-Seq, ribo-seq, and quantitative mass spectrometry. For each gene, an Elastic Net algorithm
539 was used for feature selection from SNPs located within the cis-region defined for each gene (i.e.
540 gene body +/- 1Mb flanking regions) based on results from a ten-fold cross-validation. After that,
541 weights were produced for every selected SNP of each gene, which were used in the prediction
542 models. For each gene, we calculated the Pearson correlation between predicted and observed
543 gene expression (Cross-validation R, R_{cv}), which was considered as a metric for prediction
544 accuracy. Only the genes with $R_{cv} > 0.1$ and $P < 0.05$ were retained. We produced prediction
545 models for the unified set of 7,458 genes and 185 samples shared across omics.

546 *Building conditionally independent prediction models*

547 To identify omics-shared and omics-specific QTL genes among different omics, we built
548 conditionally independent prediction models for each omics. For each gene, we regressed from
549 the data the genetic regulation signals of all other omics types (i.e. the imputed quantification

level of all other omics types). The models were built based on the regressed expression (i.e. the residual from the regression). Note that because in many cases multiple protein isoform quantifications match to one gene, when building model for gene-based quantification omics types, such as mRNA and ribosome occupancy, we aggregate data across all protein isoforms for each gene. In contrast, when building model for protein data, quantifications between isoforms were kept separate.

More specifically, assuming a total of *p* omics types, we took the following steps to identify shared-QTL genes and to define omics-specific QTL genes:

Step 1: For each gene and each omics type, we built original prediction models using equation 1

$$E_k = \sum_{i=0}^{n} w_i S_i + \varepsilon, \quad (1)$$

where E_k denotes the observed expression of omics type *k*; w_i denotes the weight of the *i*-th SNP S_i; n denotes the number of SNPs selected by the elastic net.

We can calculate the imputed expression with genotype data: $\widehat{E_k} = \sum_{i=0}^{n} w_i S_i$.

Step 2: For each gene in omics type *k*, we regressed out the imputed expression on the same gene of all the other omics types from the observed expression using equation 2 and kept the residual

$$E_k = \beta_1 \widehat{E_1} + \beta_2 \widehat{E_2} + \cdots + \beta_{l-1} \widehat{E_{l-1}} + e_k, \quad (2)$$

where β_l denotes slope of the *l*-th other omics types, $\widehat{E_l}$ denotes the imputed expression of the *l*-th other omics types, and e_k denotes the residual for omics type *k*. Then, the expression value conditioning on the genetic regulation of all the other omics types (R_k) can be represented as the

569 sum of mean expression across genes ($\overline{E_k}$) and the residual (e_k). Step 3: After that, we built the
570 conditional prediction models and estimated the SNP-weights parameters) using equation 3.

571 $$R_k = \sum_{i=0}^{n} m_i S_i + \varepsilon, \quad (3)$$

572 where m_i is the weight of i-th SNP S_i for the conditional prediction model.

573 Step 3: We calculated the square of correlation of imputed gene expression between the original
574 model and conditional model and named it as observed R_c^2.

575 Step 4: To determine whether a gene share genetic regulation signals with other omics types, we
576 used permutation to create a null distribution of R_c^2. We permuted the sample label for imputed
577 expression of the other omics types, and repeated step 2to build the conditional prediction model
578 using permuted data, which we named conditional permutation model. We then calculated
579 permutation R_c^2 of imputed gene expression between the original model and conditional
580 permutation model.

581 Step 5: We performed 50 permutations and calculated empirical p values of the observed R_c^2
582 based on its rank among the permutation R_c^2 (i.e. ordered from the smallest to the greatest).

583 Step 6: We used Benjamini-Hochberg control procedure to adjust the empirical p values to
584 calculate FDR for shared-QTL genes.

585 Step 7: We defined omics-specific QTL genes as genes with FDR > 0.1 and R_c^2 >0.9, and the
586 omics-shared QTL genes as genes with FDR <0.1 and R_c^2 <0.5.

587 **Colocalization**

588 We used *coloc* (*11*) to detect signal colocalization between SCZ GWAS and each QTL type at the
589 cis-region of each QTL gene. For each QTL gene, for all common SNPs (MAF > 0.05) within the
590 cis-region, we use QTL effects and GWAS summary statistics as input for coloc analysis. For
591 QTL effects, we calculated slope and square of slope' standard error from a linear model fit using
592 SNP genotype as the predictor of gene quantifications. We then used the coloc.abf function in the
593 R package "*coloc*" (*11*) to calculate the posterior possibility of each hypothesis using the default
594 prior. We use posterior probability of 70% for the colocalization hypothesis (i.e. PPH4) as the
595 cutoff for reporting our colocalization findings.

596 **TWAS identification of SCZ risk genes**

597 Gene-level association tests were performed for SCZ using S-PrediXcan (*33*) based on the
598 prediction models built using our omics data, which is described in the "building prediction
599 models" section, and the SCZ GWAS summary statistics data from PGC3 (*17*). The association
600 tests were done separately for each omics. For protein data, we performed omnibus test to
601 incorporate p values of all protein isoforms together to calculate a single p value for the
602 corresponding gene. The family wise error rates for SCZ risk genes were calculated using
603 Bonferroni correction of nominal p values.

604 **Two-sample Mendelian randomization (MR)**

605 To identify causal relationships between each omics type and SCZ we used MR analysis. Here we
606 used fine mapped QTL SNPs as instruments, gene expression quantification at each omics type as
607 exposure, and SCZ GWAS signal as outcome.

608 More specifically, we took the following steps to test for causal relationship between gene
609 regulation at each omics level and SCZ:

610 Step 1: Here we used two methods of selecting instrument SNPs. For each gene of each QTL
611 type, we performed LD clumping by PLINK with "–clump-kb 1000 –clump-r2 0.5" parameters to
612 select instrument SNPs with p value < 0.05. Another method is fine-mapping by susie_rss()
613 function with default parameters in R package "SuSiE" (*44*).

614 Step 2: For each gene, we used harmonise_data() function in R package "TwoSampleMR" (*56*) to
615 harmonize QTLs of each omics type and SCZ GWAS SNPs to be in the same direction (i.e. effect
616 relative to the same allele).

617 Step 3: We then performed two-sample MR for each gene in each omics type separately. Two-
618 sample MR analysis was done using mr() function, which includes IVW and Egger methods, in
619 the R package "TwoSampleMR" (*56*).

620 Step 4: We used the intercept test (i.e. the Egger method) to test for horizontal pleiotropy, and
621 used predictor coefficient and its corresponding p value from IVW to determine the effect size
622 and significance of causal effects for each omics type on SCZ.

623 Step 5: We used Benjamini-Hochberg to adjust for multiple testing.

624 Step 6: We define a gene by omics type combination as causal for SCZ if the causal effect test
625 passed the FDR < 0.1 cutoff and the Egger intercept test passed the intercept p value > 0.05
626 cutoff.

627 **One-sample Mendelian randomization**

628 For one-sample MR we used two-stage least squares (2SLS) approach to find causal relationships
629 between omics types. We performed the following analysis in two iterations, both following the
630 direction of genetic information flow. In the first iteration, we tested causal relationships between

transcript level and ribosome occupancy level (i.e. mRNA -> ribosome occupancy). In the second iteration, we tested causal relationships between ribosome occupancy level and protein level (i.e. ribosome occupancy -> protein).

Step 1: We used the same two methods for instrument SNP selection as described in the previous section. Here we tested two pathways (mRNA -> ribosome occupancy and ribosome occupancy -> protein). For mRNA -> ribosome occupancy, we used the SNPs of eQTL with p value < 0.05 and LD clumping by PLINK with "–clump-kb 1000 –clump-r2 0.5" parameters. For ribosome occupancy -> protein, we used the SNPs of rQTL with p value < 0.05 and LD clumping by PLINK with "–clump-kb 1000 –clump-r2 0.5".

Step 2: We then combined genotype and quantification data of the relevant omics types into a dataframe: dat_{mr} for mRNA -> ribosome occupancy pathway, dat_{rp} for ribosome occupancy -> protein pathway.

Step 3: For each gene, we then set formula in R package "ivreg": ivreg("ribosome occupancy ~ mRNA | SNP_1+SNP_2+SNP_3+…+SNP_n", data = dat_{mr}) and ivreg("protein ~ ribosome occupancy | SNP_1+SNP_2+SNP_3+…+SNP_m", data = dat_{rp}).

Step 4: We used the summary() function to get slope, p value of slope, intercept, p value of intercept, F-statistic and p value of F-statistic of each ivreg object. P value of Intercept was used to test horizontal pleiotropy, F-statistic was used to check instrument strength.

Step 5: We used Benjamini-Hochberg control procedure to adjust for multiple testing.

Step 6: We defined a causal relationship for each gene by pathway combination as the causal effect test passed the FDR < 0.1 cutoff and the Egger intercept test passed the intercept p value > 0.05 cutoff.

653 **References**

654 1. A. Fialkowski, T. M. Beasley, H. K. Tiwari, "[11 – Multifactorial Inheritance and Complex Diseases]" in
655 *Emery and Rimoin's Principles and Practice of Medical Genetics and Genomics (Seventh Edition)*, R. E.
656 Pyeritz, B. R. Korf, W. W. Grody, Eds. (Academic Press, 2019), pp. 323–358.

657 2. A. S. Brown, The environment and susceptibility to schizophrenia. *Prog. Neurobiol*. 93, 23–58 (2011).

658 3. N. Solovieff, C. Cotsapas, P. H. Lee, S. M. Purcell, J. W. Smoller, Pleiotropy in complex traits: challenges and
659 strategies. *Nat. Rev. Genet.* 14, 483–495 (2013).

660 4. K. Watanabe, E. Taskesen, A. van Bochoven, D. Posthuma, Functional mapping and annotation of genetic
661 associations with FUMA. *Nat. Commun.* 8, 1826 (2017).

662 5. M. G. Heckman, C. Labbé, A. L. Kolicheski, A. I. Soto-Beasley, R. L. Walton, R. R. Valentino, E. R. Brennan,
663 P. W. Johnson, S. Baheti, V. Sarangi, Y. Ren, R. J. Uitti, Z. K. Wszolek, O. A. Ross, Fine-mapping of the non-
664 coding variation driving the Caucasian LRRK2 GWAS signal in Parkinson's disease. *Parkinsonism Relat.*
665 *Disord*. 83, 22–30 (2021).

666 6. D. J. Schaid, W. Chen, N. B. Larson, From genome-wide associations to candidate causal variants by statistical
667 fine-mapping. *Nat. Rev. Genet.* 19, 491–504 (2018).

668 7. A. Gusev, A. Ko, H. Shi, G. Bhatia, W. Chung, B. W. J. H. Penninx, R. Jansen, E. J. C. de Geus, D. I.
669 Boomsma, F. A. Wright, P. F. Sullivan, E. Nikkola, M. Alvarez, M. Civelek, A. J. Lusis, T. Lehtimäki, E.
670 Raitoharju, M. Kähönen, I. Seppälä, O. T. Raitakari, J. Kuusisto, M. Laakso, A. L. Price, P. Pajukanta, B.
671 Pasaniuc, Integrative approaches for large-scale transcriptome-wide association studies. *Nat. Genet.* 48, 245–
672 252 (2016).

673 8. M. Fromer, P. Roussos, S. K. Sieberts, J. S. Johnson, D. H. Kavanagh, T. M. Perumal, D. M. Ruderfer, E. C.
674 Oh, A. Topol, H. R. Shah, L. L. Klei, R. Kramer, D. Pinto, Z. H. Gümüş, A. E. Cicek, K. K. Dang, A. Browne,

675 C. Lu, L. Xie, B. Readhead, E. A. Stahl, J. Xiao, M. Parvizi, T. Hamamsy, J. F. Fullard, Y.-C. Wang, M. C.
676 Mahajan, J. M. J. Derry, J. T. Dudley, S. E. Hemby, B. A. Logsdon, K. Talbot, T. Raj, D. A. Bennett, P. L. De
677 Jager, J. Zhu, B. Zhang, P. F. Sullivan, A. Chess, S. M. Purcell, L. A. Shinobu, L. M. Mangravite, H.
678 Toyoshiba, R. E. Gur, C.-G. Hahn, D. A. Lewis, V. Haroutunian, M. A. Peters, B. K. Lipska, J. D. Buxbaum,
679 E. E. Schadt, K. Hirai, K. Roeder, K. J. Brennand, N. Katsanis, E. Domenici, B. Devlin, P. Sklar, Gene
680 expression elucidates functional impact of polygenic risk for schizophrenia. *Nat. Neurosci*. 19, 1442–1453
681 (2016).

682 9. ENCODE Project Consortium, J. E. Moore, M. J. Purcaro, H. E. Pratt, C. B. Epstein, N. Shoresh, J. Adrian, T.
683 Kawli, C. A. Davis, A. Dobin, R. Kaul, J. Halow, E. L. Van Nostrand, P. Freese, D. U. Gorkin, Y. Shen, Y. He,
684 M. Mackiewicz, F. Pauli-Behn, B. A. Williams, A. Mortazavi, C. A. Keller, X.-O. Zhang, S. I. Elhajjajy, J.
685 Huey, D. E. Dickel, V. Snetkova, X. Wei, X. Wang, J. C. Rivera-Mulia, J. Rozowsky, J. Zhang, S. B. Chhetri,
686 J. Zhang, A. Victorsen, K. P. White, A. Visel, G. W. Yeo, C. B. Burge, E. Lécuyer, D. M. Gilbert, J. Dekker, J.
687 Rinn, E. M. Mendenhall, J. R. Ecker, M. Kellis, R. J. Klein, W. S. Noble, A. Kundaje, R. Guigó, P. J. Farnham,
688 J. M. Cherry, R. M. Myers, B. Ren, B. R. Graveley, M. B. Gerstein, L. A. Pennacchio, M. P. Snyder, B. E.
689 Bernstein, B. Wold, R. C. Hardison, T. R. Gingeras, J. A. Stamatoyannopoulos, Z. Weng, Expanded
690 encyclopaedias of DNA elements in the human and mouse genomes. *Nature*. 583, 699–710 (2020).

691 10. GTEx Consortium, The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science*.
692 369, 1318–1330 (2020).

693 11. C. Giambartolomei, D. Vukcevic, E. E. Schadt, L. Franke, A. D. Hingorani, C. Wallace, V. Plagnol, Bayesian
694 test for 31etwork31renia31 between pairs of genetic association studies using summary statistics. *PloS Genet*.
695 10, e1004383 (2014).

696 12. E. R. Gamazon, H. E. Wheeler, K. P. Shah, S. V. Mozaffari, K. Aquino-Michaels, R. J. Carroll, A. E. Eyler, J.
697 C. Denny, D. L. Nicolae, N. J. Cox, H. K. Im, A gene-based association method for mapping traits using
698 reference transcriptome data. *Nat. Genet*. 47, 1091–1098 (2015).

699 13. Z. Zhu, F. Zhang, H. Hu, A. Bakshi, M. R. Robinson, J. E. Powell, G. W. Montgomery, M. E. Goddard, N. R.
700 Wray, P. M. Visscher, J. Yang, Integration of summary data from GWAS and eQTL studies predicts complex
701 trait gene targets. *Nat. Genet*. 48, 481–487 (2016).

702 14. R. S. Kahn, I. E. Sommer, R. M. Murray, A. Meyer-Lindenberg, D. R. Weinberger, T. D. Cannon, M. O
703 'Donovan, C. U. Correll, J. M. Kane, J. van Os, T. R. Insel, Schizophrenia. *Nat. Rev. Dis. Primers*. 1, 15067
704 (2015).

705 15. P. Lichtenstein, B. H. Yip, C. Björk, Y. Pawitan, T. D. Cannon, P. F. Sullivan, C. M. Hultman, Common
706 genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study.
707 *Lancet*. 373, 234–239 (2009).

708 16. P. F. Sullivan, K. S. Kendler, M. C. Neale, Schizophrenia as a complex trait: evidence from a meta-analysis of
709 twin studies. *Arch Gen Psychiatry*. 60, 1187–1192 (2003).

710 17. V. Trubetskoy, A. F. Pardiñas, T. Qi, G. Panagiotaropoulou, S. Awasthi, T. B. Bigdeli, J. Bryois, C.-Y. Chen,
711 C. A. Dennison, L. S. Hall, M. Lam, K. Watanabe, O. Frei, T. Ge, J. C. Harwood, F. Koopmans, S.
712 Magnusson, A. L. Richards, J. Sidorenko, Y. Wu, J. Zeng, J. Grove, M. Kim, Z. Li, G. Voloudakis, W. Zhang,
713 M. Adams, I. Agartz, E. G. Atkinson, E. Agerbo, M. Al Eissa, M. Albus, M. Alexander, B. Z. Alizadeh, K.
714 Alptekin, T. D. Als, F. Amin, V. Arolt, M. Arrojo, L. Athanasiu, M. H. Azevedo, S. A. Bacanu, N. J. Bass, M.
715 Begemann, R. A. Belliveau, J. Bene, B. Benyamin, S. E. Bergen, G. Blasi, J. Bobes, S. Bonassi, A. Braun, R.
716 A. Bressan, E. J. Bromet, R. Bruggeman, P. F. Buckley, R. L. Buckner, J. Bybjerg-Grauholm, W. Cahn, M. J.
717 Cairns, M. E. Calkins, V. J. Carr, D. Castle, S. V. Catts, K. D. Chambert, R. C. K. Chan, B. Chaumette, W.
718 Cheng, E. F. C. Cheung, S. A. Chong, D. Cohen, A. Consoli, Q. Cordeiro, J. Costas, C. Curtis, M. Davidson,
719 K. L. Davis, L. de Haan, F. Degenhardt, L. E. DeLisi, D. Demontis, F. Dickerson, D. Dikeos, T. Dinan, S.
720 Djurovic, J. Duan, G. Ducci, F. Dudbridge, J. G. Eriksson, L. Fañanás, S. V. Faraone, A. Fiorentino, A.
721 Forstner, J. Frank, N. B. Freimer, M. Fromer, A. Frustaci, A. Gadelha, G. Genovese, E. S. Gershon, M.
722 Giannitelli, I. Giegling, P. Giusti-Rodríguez, S. Godard, J. I. Goldstein, J. González Peñas, A. González-Pinto,
723 S. Gopal, J. Gratten, M. F. Green, T. A. Greenwood, O. Guillin, S. Gülöksüz, R. E. Gur, R. C. Gur, B. Gutié
724 rrez, E. Hahn, H. Hakonarson, V. Haroutunian, A. M. Hartmann, C. Harvey, C. Hayward, F. A. Henskens, S.

Herms, P. Hoffmann, D. P. Howrigan, M. Ikeda, C. Iyegbe, I. Joa, A. Julià, A. K. Kähler, T. Kam-Thong, Y. Kamatani, S. Karachanak-Yankova, O. Kebir, M. C. Keller, B. J. Kelly, A. Khrunin, S.-W. Kim, J. Klovins, N. Kondratiev, B. Konte, J. Kraft, M. Kubo, V. Kučinskas, Z. A. Kučinskiene, A. Kusumawardhani, H. Kuzelova-Ptackova, S. Landi, L. C. Lazzeroni, P. H. Lee, S. E. Legge, D. S. Lehrer, R. Lencer, B. Lerer, M. Li, J. Lieberman, G. A. Light, S. Limborska, C.-M. Liu, J. Lönnqvist, C. M. Loughland, J. Lubinski, J. J. Luykx, A. Lynham, M. Macek, A. Mackinnon, P. K. E. Magnusson, B. S. Maher, W. Maier, D. Malaspina, J. Mallet, S. R. Marder, S. marsal, A. R. Martin, L. Martorell, M. Mattheisen, R. W. McCarley, C. McDonald, J. J. McGrath, H. Medeiros, S. Meier, B. Melegh, I. Melle, R. I. Mesholam-Gately, A. Metspalu, P. T. Michie, L. Milani, V. Milanova, M. Mitjans, E. Molden, E. Molina, M. D. Molto, V. Mondelli, C. Moreno, C. P. Morley, G. Muntané, K. C. Murphy, I. Myin-Germeys, I. Nenadić, G. Nestadt, L. Nikitina-Zake, C. Noto, K. H. Nuechterlein, N. L. O'Brien, F. A. O'Neill, S.-Y. Oh, A. Olincy, V. K. Ota, C. Pantelis, G. N. Papadimitriou, M. Parellada, T. Paunio, R. Pellegrino, S. Periyasamy, D. O. Perkins, B. Pfuhlmann, O. Pietiläinen, J. Pimm, D. Porteous, J. Powell, D. Quattrone, D. Quested, A. D. Radant, A. Rampino, M. H. Rapaport, A. Rautanen, A. Reichenberg, C. Roe, J. L. Roffman, J. Roth, M. Rothermundt, B. P. F. Rutten, S. Saker-Delye, V. Salomaa, J. Sanjuan, M. L. Santoro, A. Savitz, U. Schall, R. J. Scott, L. J. Seidman, S. I. Sharp, J. Shi, L. J. Siever, E. Sigurdsson, K. Sim, N. Skarabis, P. Slominsky, H.-C. So, J. L. Sobell, E. Söderman, H. J. Stain, N. E. Steen, A. A. Steixner-Kumar, E. Stögmann, W. S. Stone, R. E. Straub, F. Streit, E. Strengman, T. S. Stroup, M. Subramaniam, C. A. Sugar, J. Suvisaari, D. M. Svrakic, N. R. Swerdlow, J. P. Szatkiewicz, T. M. T. Ta, A. Takahashi, C. Terao, F. Thibaut, D. Toncheva, P. A. Tooney, S. Torretta, S. Tosato, G. B. Tura, B. I. Turetsky, A. Üçok, A. Vaaler, T. van Amelsvoort, R. van Winkel, J. Veijola, J. Waddington, H. Walter, A. Waterreus, B. T. Webb, M. Weiser, N. M. Williams, S. H. Witt, B. K. Wormley, J. Q. Wu, Z. Xu, R. Yolken, C. C. Zai, W. Zhou, F. Zhu, F. Zimprich, E. C. Atbaşoğlu, M. Ayub, C. Benner, A. Bertolino, D. W. Black, N. J. Bray, G. Breen, N. G. Buccola, W. F. Byerley, W. J. Chen, C. R. Cloninger, B. Crespo-Facorro, G. Donohoe, R. Freedman, C. Galletly, M. J. Gandal, M. Gennarelli, D. M. Hougaard, H.-G. Hwu, A. V. Jablensky, S. A. McCarroll, J. L. Moran, O. Mors, P. B. Mortensen, B. Müller-Myhsok, A. L. Neil, M. Nordentoft, M. T. Pato, T. L. Petryshen, M. Pirinen, A. E. Pulver, T. G. Schulze, J. M. Silverman, J. W. Smoller, E. A. Stahl, D. W. Tsuang, E. Vilella, S.-H. Wang, S. Xu, R. Adolfsson, C. Arango, B. T. Baune, S. I. Belangero, A. D. Børglum, D. Braff, E. Bramon, J. D. Buxbaum, D. Campion, J. A. Cervilla, S. Cichon, D. A. Collier, A. Corvin, D. Curtis, M. D. Forti, E. Domenici, H. Ehrenreich, V. Escott-Price, T. Esko, A. H. Fanous, A. Gareeva, M. Gawlik, P. V. Gejman, M. Gill, S. J. Glatt, V. Golimbet, K. S. Hong, C. M. Hultman, S. E. Hyman, N. Iwata,

755 E. G. Jönsson, R. S. Kahn, J. L. Kennedy, E. Khusnutdinova, G. Kirov, J. A. Knowles, M.-O. Krebs, C.
756 Laurent-Levinson, J. Lee, T. Lencz, D. F. Levinson, Q. S. Li, J. Liu, A. K. Malhotra, D. Malhotra, A.
757 McIntosh, A. McQuillin, P. R. Menezes, V. A. Morgan, D. W. Morris, B. J. Mowry, R. M. Murray, V.
758 Nimgaonkar, M. M. Nöthen, R. A. Ophoff, S. A. Paciga, A. Palotie, C. N. Pato, S. Qin, M. Rietschel, B. P.
759 Riley, M. Rivera, D. Rujescu, M. C. Saka, A. R. Sanders, S. G. Schwab, A. Serretti, P. C. Sham, Y. Shi, D. St
760 Clair, H. Stefánsson, K. Stefansson, M. T. Tsuang, J. van Os, M. P. Vawter, D. R. Weinberger, T. Werge, D.
761 B. Wildenauer, X. Yu, W. Yue, P. A. Holmans, A. J. Pocklington, P. Roussos, E. Vassos, M. Verhage, P. M.
762 Visscher, J. Yang, D. Posthuma, O. A. Andreassen, K. S. Kendler, M. J. Owen, N. R. Wray, M. J. Daly, H.
763 Huang, B. M. Neale, P. F. Sullivan, S. Ripke, J. T. R. Walters, M. C. O'Donovan, Mapping genomic loci
764 implicates genes and synaptic biology in schizophrenia. *Nature*. 604, 502–508 (2022).

765 18. J. B. W. Wolf, Principles of transcriptome analysis and gene expression quantification: an RNA-seq tutorial.
766 *Mol Ecol Resour*. 13, 559–572 (2013).

767 19. A. Battle, Z. Khan, S. H. Wang, A. Mitrano, M. J. Ford, J. K. Pritchard, Y. Gilad, Impact of regulatory
768 variation from RNA to protein. *Science*. 347, 664–667 (2015).

769 20. E. Ferkingstad, P. Sulem, B. A. Atlason, G. Sveinbjornsson, M. I. Magnusson, E. L. Styrmisdottir, K.
770 Gunnarsdottir, A. Helgason, A. Oddsson, B. V. Halldorsson, B. O. Jensson, F. Zink, G. H. Halldorsson, G.
771 Masson, G. A. Arnadottir, H. Katrinardottir, K. Juliusson, M. K. Magnusson, O. T. Magnusson, R.
772 Fridriksdottir, S. Saevarsdottir, S. A. Gudjonsson, S. N. Stacey, S. Rognvaldsson, T. Eiriksdottir, T. A.
773 Olafsdottir, V. Steinthorsdottir, V. Tragante, M. O. Ulfarsson, H. Stefansson, I. Jonsdottir, H. Holm, T. Rafnar,
774 P. Melsted, J. Saemundsdottir, G. L. Norddahl, S. H. Lund, D. F. Gudbjartsson, U. Thorsteinsdottir, K.
775 Stefansson, Large-scale integration of the plasma proteome with genetics and disease. *Nat. Genet*. 53, 1712–
776 1721 (2021).

777 21. J. Luo, M. Niu, L. Li, D. Kong, Y. Jiang, A. Shieh, L. Cheng, G. Giase, K. Grennan, K. White, C. Chen, S.
778 Wang, D. Pinto, Y. Wang, C. Liu, J. Peng, X. Wang, https://doi.org/10.21203/rs.3.rs-1633422/v1 (2022).

779 22. C. Giambartolomei, J. Zhenli Liu, W. Zhang, M. Hauberg, H. Shi, J. Boocock, J. Pickrell, A. E. Jaffe,
780 CommonMind Consortium, B. Pasaniuc, P. Roussos, A Bayesian framework for multiple trait colocalization
781 from summary association statistics. *Bioinformatics*. 34, 2538–2545 (2018).

782 23. C. N. Foley, J. R. Staley, P. G. Breen, B. B. Sun, P. D. W. Kirk, S. Burgess, J. M. M. Howson, A fast and
783 efficient colocalization algorithm for identifying shared genetic risk factors across multiple traits. *Nat.*
784 *Commun*. 12, 764 (2021).

785 24. N. T. Ingolia, S. Ghaemmaghami, J. R. S. Newman, J. S. Weissman, Genome-wide analysis in vivo of
786 translation with nucleotide resolution using ribosome profiling. *Science*. 324, 218–223 (2009).

787 25. N. T. Ingolia, G. A. Brar, S. Rouskin, A. M. McGeachy, J. S. Weissman, The ribosome profiling strategy for
788 monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nat. Protoc*. 7,
789 1534–1550 (2012).

790 26. Cenik C, Cenik ES, Byeon GW, Grubert F, Candille SI, Spacek D, Alsallakh B, Tilgner H, Araya CL, Tang H,
791 Ricci E, Snyder MP. Integrative analysis of RNA, translation, and protein levels reveals distinct regulatory
792 variation across humans. *Genome Res.* 25, 1610-1621 (2015).

793 27. S. Akbarian, C. Liu, J. A. Knowles, F. M. Vaccarino, P. J. Farnham, G. E. Crawford, A. E. Jaffe, D. Pinto, S.
794 Dracheva, D. H. Geschwind, J. Mill, A. C. Nairn, A. Abyzov, S. Pochareddy, S. Prabhakar, S. Weissman, P. F.
795 Sullivan, M. W. State, Z. Weng, M. A. Peters, K. P. White, M. B. Gerstein, A. Amiri, C. Armoskus, A. E.
796 Ashley-Koch, T. Bae, A. Beckel-Mitchener, B. P. Berman, G. A. Coetzee, G. Coppola, N. Francoeur, M.
797 Fromer, R. Gao, K. Grennan, J. Herstein, D. H. Kavanagh, N. A. Ivanov, Y. Jiang, R. R. Kitchen, A.
798 Kozlenkov, M. Kundakovic, M. Li, Z. Li, S. Liu, L. M. Mangravite, E. Mattei, E. Markenscoff-Papadimitriou,
799 F. C. P. Navarro, N. North, L. Omberg, D. Panchision, N. Parikshak, J. Poschmann, A. J. Price, M. Purcaro, T.
800 E. Reddy, P. Roussos, S. Schreiner, S. Scuderi, R. Sebra, M. Shibata, A. W. Shieh, M. Skarica, W. Sun, V.
801 Swarup, A. Thomas, J. Tsuji, H. van Bakel, D. Wang, Y. Wang, K. Wang, D. M. Werling, A. J. Willsey, H.
802 Witt, H. Won, C. C. Y. Wong, G. A. Wray, E. Y. Wu, X. Xu, L. Yao, G. Senthil, T. Lehner, P. Sklar, N.
803 Sestan, The PsychENCODE project. *Nat. Neurosci*. 18, 1707–1712 (2015).

804 28. M. J. Gandal, P. Zhang, E. Hadjimichael, R. L. Walker, C. Chen, S. Liu, H. Won, H. van Bakel, M. Varghese,
805 Y. Wang, A. W. Shieh, J. Haney, S. Parhami, J. Belmont, M. Kim, P. Moran Losada, Z. Khan, J. Mleczko, Y.
806 Xia, R. Dai, D. Wang, Y. T. Yang, M. Xu, K. Fish, P. R. Hof, J. Warrell, D. Fitzgerald, K. White, A. E. Jaffe,
807 PsychENCODE Consortium, M. A. Peters, M. Gerstein, C. Liu, L. M. Iakoucheva, D. Pinto, D. H. Geschwind,
808 Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. *Science*. 362,
809 eaat8127 (2018).

810 29. G. A. Brar, J. S. Weissman, Ribosome profiling reveals the what, when, where and how of protein synthesis.
811 *Nat. Rev. Mol. Cell Biol*. 16, 651–664 (2015).

812 30. J. D. Storey, R. Tibshirani, Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. U.S.A.* 100,
813 9440–9445 (2003).

814 31. G. E. Hoffman, J. Bendl, G. Voloudakis, K. S. Montgomery, L. Sloofman, Y.-C. Wang, H. R. Shah, M. E.
815 Hauberg, J. S. Johnson, K. Girdhar, L. Song, J. F. Fullard, R. Kramer, C.-G. Hahn, R. Gur, S. Marenco, B. K.
816 Lipska, D. A. Lewis, V. Haroutunian, S. Hemby, P. Sullivan, S. Akbarian, A. Chess, J. D. Buxbaum, G. E.
817 Crawford, E. Domenici, B. Devlin, S. K. Sieberts, M. A. Peters, P. Roussos, CommonMind Consortium
818 provides transcriptomic and epigenomic data for Schizophrenia and Bipolar Disorder. *Sci. Data*. 6, 180
819 (2019).

820 32. D. W. Yao, L. J. O'Connor, A. L. Price, A. Gusev, Quantifying genetic effects on disease mediated by assayed
821 gene expression levels. *Nat. Genet*. 52, 626–633 (2020).

822 33. A. N. Barbeira, S. P. Dickinson, R. Bonazzola, J. Zheng, H. E. Wheeler, J. M. Torres, E. S. Torstenson, K. P.
823 Shah, T. Garcia, T. L. Edwards, E. A. Stahl, L. M. Huckins, D. L. Nicolae, N. J. Cox, H. K. Im, Exploring the
824 phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics.
825 *Nat. Commun*. 9, 1825 (2018).

826 34. S. Ripke, B. M. Neale, A. Corvin, J. T. R. Walters, K.-H. Farh, P. A. Holmans, P. Lee, B. Bulik-Sullivan, D. A.
827 Collier, H. Huang, T. H. Pers, I. Agartz, E. Agerbo, M. Albus, M. Alexander, F. Amin, S. A. Bacanu, M.
828 Begemann, R. A. Belliveau Jr, J. Bene, S. E. Bergen, E. Bevilacqua, T. B. Bigdeli, D. W. Black, R.

Bruggeman, N. G. Buccola, R. L. Buckner, W. Byerley, W. Cahn, G. Cai, D. Campion, R. M. Cantor, V. J. Carr, N. Carrera, S. V. Catts, K. D. Chambert, R. C. K. Chan, R. Y. L. Chen, E. Y. H. Chen, W. Cheng, E. F. C. Cheung, S. Ann Chong, C. Robert Cloninger, D. Cohen, N. Cohen, P. Cormican, N. Craddock, J. J. Crowley, D. Curtis, M. Davidson, K. L. Davis, F. Degenhardt, J. Del Favero, D. Demontis, D. Dikeos, T. Dinan, S. Djurovic, G. Donohoe, E. Drapeau, J. Duan, F. Dudbridge, N. Durmishi, P. Eichhammer, J. Eriksson, V. Escott-Price, L. Essioux, A. H. Fanous, M. S. Farrell, J. Frank, L. Franke, R. Freedman, N. B. Freimer, M. Friedl, J. I. Friedman, M. Fromer, G. Genovese, L. Georgieva, I. Giegling, P. Giusti-Rodríguez, S. Godard, J. I. Goldstein, V. Golimbet, S. Gopal, J. Gratten, L. de Haan, C. Hammer, M. L. Hamshere, M. Hansen, T. Hansen, V. Haroutunian, A. M. Hartmann, F. A. Henskens, S. Herms, J. N. Hirschhorn, P. Hoffmann, A. Hofman, M. V. Hollegaard, D. M. Hougaard, M. Ikeda, I. Joa, A. Julià, R. S. Kahn, L. Kalaydjieva, S. Karachanak-Yankova, J. Karjalainen, D. Kavanagh, M. C. Keller, J. L. Kennedy, A. Khrunin, Y. Kim, J. Klovins, J. A. Knowles, B. Konte, V. Kucinskas, Z. Ausrele Kucinskiene, H. Kuzelova-Ptackova, A. K. Kähler, C. Laurent, J. Lee Chee Keong, S. Hong Lee, S. E. Legge, B. Lerer, M. Li, T. Li, K.-Y. Liang, J. Lieberman, S. Limborska, C. M. Loughland, J. Lubinski, J. Lönnqvist, M. Macek Jr, P. K. E. Magnusson, B. S. Maher, W. Maier, J. Mallet, S. Marsal, M. Mattheisen, M. Mattingsdal, R. W. McCarley, C. McDonald, A. M. McIntosh, S. Meier, C. J. Meijer, B. Melegh, I. Melle, R. I. Mesholam-Gately, A. Metspalu, P. T. Michie, L. Milani, V. Milanova, Y. Mokrab, D. W. Morris, O. Mors, K. C. Murphy, R. M. Murray, I. Myin-Germeys, B. Müller-Myhsok, M. Nelis, I. Nenadic, D. A. Nertney, G. Nestadt, K. K. Nicodemus, L. Nikitina-Zake, L. Nisenbaum, A. Nordin, E. O'Callaghan, C. O'Dushlaine, F. A. O'Neill, S.-Y. Oh, A. Olincy, L. Olsen, J. Van Os, C. Pantelis, G. N. Papadimitriou, S. Papiol, E. Parkhomenko, M. T. Pato, T. Paunio, M. Pejovic-Milovancevic, D. O. Perkins, O. Pietiläinen, J. Pimm, A. J. Pocklington, J. Powell, A. Price, A. E. Pulver, S. M. Purcell, D. Quested, H. B. Rasmussen, A. Reichenberg, M. A. Reimers, A. L. Richards, J. L. Roffman, P. Roussos, D. M. Ruderfer, V. Salomaa, A. R. Sanders, U. Schall, C. R. Schubert, T. G. Schulze, S. G. Schwab, E. M. Scolnick, R. J. Scott, L. J. Seidman, J. Shi, E. Sigurdsson, T. Silagadze, J. M. Silverman, K. Sim, P. Slominsky, J. W. Smoller, H.-C. So, ChrisC. A. Spencer, E. A. Stahl, H. Stefansson, S. Steinberg, E. Stogmann, R. E. Straub, E. Strengman, J. Strohmaier, T. Scott Stroup, M. Subramaniam, J. Suvisaari, D. M. Svrakic, J. P. Szatkiewicz, E. Söderman, S. Thirumalai, D. Toncheva, S. Tosato, J. Veijola, J. Waddington, D. Walsh, D. Wang, Q. Wang, B. T. Webb, M. Weiser, D. B. Wildenauer, N. M. Williams, S. Williams, S. H. Witt, A. R. Wolen, E. H. M. Wong, B. K. Wormley, H. Simon Xi, C. C. Zai, X. Zheng, F. Zimprich, N. R. Wray, K. Stefansson, P. M. Visscher, W. Trust Case-Control Consortium, R. Adolfsson, O. A. Andreassen, D. H. R. Blackwood, E. Bramon, J. D.

859 Buxbaum, A. D. Børglum, S. Cichon, A. Darvasi, E. Domenici, H. Ehrenreich, T. Esko, P. V. Gejman, M. Gill,
860 H. Gurling, C. M. Hultman, N. Iwata, A. V. Jablensky, E. G. Jönsson, K. S. Kendler, G. Kirov, J. Knight, T.
861 Lencz, D. F. Levinson, Q. S. Li, J. Liu, A. K. Malhotra, S. A. McCarroll, A. McQuillin, J. L. Moran, P. B.
862 Mortensen, B. J. Mowry, M. M. Nöthen, R. A. Ophoff, M. J. Owen, A. Palotie, C. N. Pato, T. L. Petryshen, D.
863 Posthuma, M. Rietschel, B. P. Riley, D. Rujescu, P. C. Sham, P. Sklar, D. St Clair, D. R. Weinberger, J. R.
864 Wendland, T. Werge, Schizophrenia Working Group of the Psychiatric Genomics Consortium, Psychosis
865 Endophenotypes International Consortium, Biological insights from 108 schizophrenia-associated genetic loci.
866 *Nature*. 511, 421–427 (2014).

867 35. A. F. Pardiñas, P. Holmans, A. J. Pocklington, V. Escott-Price, S. Ripke, N. Carrera, S. E. Legge, S. Bishop, D.
868 Cameron, M. L. Hamshere, J. Han, L. Hubbard, A. Lynham, K. Mantripragada, E. Rees, J. H. MacCabe, S. A.
869 McCarroll, B. T. Baune, G. Breen, E. M. Byrne, U. Dannlowski, T. C. Eley, C. Hayward, N. G. Martin, A. M.
870 McIntosh, R. Plomin, D. J. Porteous, N. R. Wray, A. Caballero, D. H. Geschwind, L. M. Huckins, D. M.
871 Ruderfer, E. Santiago, P. Sklar, E. A. Stahl, H. Won, E. Agerbo, T. D. Als, O. A. Andreassen, M. Bækvad-
872 Hansen, P. B. Mortensen, C. B. Pedersen, A. D. Børglum, J. Bybjerg-Grauholm, S. Djurovic, N. Durmishi, M.
873 G. Pedersen, V. Golimbet, J. Grove, D. M. Hougaard, M. Mattheisen, E. Molden, O. Mors, M. Nordentoft, M.
874 Pejovic-Milovancevic, E. Sigurdsson, T. Silagadze, C. S. Hansen, K. Stefansson, H. Stefansson, S. Steinberg,
875 S. Tosato, T. Werge, GERAD1 Consortium, CRESTAR Consortium, D. A. Collier, D. Rujescu, G. Kirov, M. J.
876 Owen, M. C. O'Donovan, J. T. R. Walters, Common schizophrenia alleles are enriched in mutation-intolerant
877 genes and in regions under strong background selection. *Nat. Genet*. 50, 381–389 (2018).

878 36. M. Lam, C.-Y. Chen, Z. Li, A. R. Martin, J. Bryois, X. Ma, H. Gaspar, M. Ikeda, B. Benyamin, B. C. Brown,
879 R. Liu, W. Zhou, L. Guan, Y. Kamatani, S.-W. Kim, M. Kubo, A. A. A. A. Kusumawardhani, C.-M. Liu, H.
880 Ma, S. Periyasamy, A. Takahashi, Z. Xu, H. Yu, F. Zhu, Schizophrenia Working Group of the Psychiatric
881 Genomics Consortium, Indonesia Schizophrenia Consortium, Genetic Research on 38etwork38renia 38etwork-
882 China and the Netherlands (GREAT-CN), W. J. Chen, S. Faraone, S. J. Glatt, L. He, S. E. Hyman, H.-G. Hwu,
883 S. A. McCarroll, B. M. Neale, P. Sklar, D. B. Wildenauer, X. Yu, D. Zhang, B. J. Mowry, J. Lee, P. Holmans,
884 S. Xu, P. F. Sullivan, S. Ripke, M. C. O'Donovan, M. J. Daly, S. Qin, P. Sham, N. Iwata, K. S. Hong, S. G.
885 Schwab, W. Yue, M. Tsuang, J. Liu, X. Ma, R. S. Kahn, Y. Shi, H. Huang, Comparative genetic architectures
886 of schizophrenia in East Asian and European populations. *Nat. Genet*. 51, 1670–1678 (2019).

887 37. Johnson RJ, Gomez-Pinilla F, Nagel M, Nakagawa T, Rodriguez-Iturbe B, Sanchez-Lozada LG, Tolan DR,
888 Lanaspa MA. Cerebral Fructose Metabolism as a Potential Mechanism Driving Alzheimer's Disease. *Front*
889 *Aging Neurosci*. 2, 560865 (2020).

890 38. Wu Y, Bi R, Zeng C, Ma C, Sun C, Li J, Xiao X, Li M, Zhang DF, Zheng P, Sheng N, Luo XJ, Yao YG.
891 Identification of the primate-specific gene BTN3A2 as an additional schizophrenia risk gene in the MHC loci.
892 *EBioMedicine*. 44:530-541 (2019).

893 39. Korologou-Linden R, Leyden GM, Relton CL, Richmond RC, Richardson TG. Multi-omics analyses of
894 cognitive traits and psychiatric disorders highlights brain-dependent mechanisms. *Hum Mol Genet.* 32(6):885–
895 96 (2021).

896 40. D. Zhou, Y. Jiang, X. Zhong, N. J. Cox, C. Liu, E. R. Gamazon, A unified framework for joint-tissue
897 transcriptome-wide association and Mendelian randomization analysis. *Nat. Genet*. 52, 1239–1246 (2020).

898 41. S. Burgess, S. G. Thompson, Interpreting findings from Mendelian randomization using the MR-Egger
899 method. *Eur. J. Epidemiol*. 32, 377–389 (2017).

900 42. S. Purcell, B. Neale, K. Todd-Brown, L. Thomas, M. A. R. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. W. de
901 Bakker, M. J. Daly, P. C. Sham, PLINK: A Tool Set for Whole-Genome Association and Population-Based
902 Linkage Analyses. *Am. J. Hum. Genet.* 81, 559–575 (2007).

903 43. A. Gkatzionis, S. Burgess, P. J. Newcombe, Statistical methods for cis-Mendelian randomization with two-
904 sample summary-level data. *Genet. Epidemiol*. 47, 3–25 (2023).

905 44. Zou Y, Carbonetto P, Wang G, Stephens M. Fine-mapping from summary data with the "Sum of Single
906 Effects" model. *PloS Genet*. 18(7): e1010299 (2022).

907 45. C. G. Artieri, H. B. Fraser, Evolution at two levels of gene expression in yeast. *Genome Res.* 24 (2014).

908 46. C. J. McManus, G. E. May, P. Spealman, A. Shteyman, Ribosome profiling reveals post-transcriptional
909 buffering of divergent gene expression in yeast. *Genome Res*. 24, 422–430 (2014).

910 47. F. W. Albert, D. Muzzey, J. S. Weissman, L. Kruglyak, Genetic influences on translation in yeast. *PloS Genet*.
911 10, e1004692 (2014).

912 48. B. Li, C. N. Dewey, RSEM: accurate transcript quantification from RNA-Seq data with or without a reference
913 genome. *BMC Bioinformatics*. 12, 323 (2011).

914 49. M. C. Oldham, P. Langfelder, S. Horvath, Network methods for describing sample relationships in genomic
915 datasets: application to Huntington's disease. *BMC Syst. Biol.* 6, 63 (2012).

916 50. B. M. Bolstad, R. A. Irizarry, M. Astrand, T. P. Speed, A comparison of normalization methods for high
917 density oligonucleotide array data based on variance and bias. *Bioinformatics*. 19, 185–193 (2003).

918 51. Y. Jiang, G. Giase, K. Grennan, A. W. Shieh, Y. Xia, L. Han, Q. Wang, Q. Wei, R. Chen, S. Liu, K. P. White,
919 C. Chen, B. Li, C. Liu, DRAMS: A tool to detect and re-align mixed-up samples for integrative studies of
920 multi-omics data. *PloS Comput. Biol.* 16, e1007522 (2020).

921 52. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods*. 9, 357–359 (2012).

922 53. Y. Liao, G. K. Smyth, W. Shi, featureCounts: an efficient general purpose program for assigning sequence
923 reads to genomic features. *Bioinformatics*. 30, 923–930 (2014).

924 54. O. Stegle, L. Parts, M. Piipari, J. Winn, R. Durbin, Using probabilistic estimation of expression residuals
925 (PEER) to obtain increased power and interpretability of gene expression analyses. *Nat. Protoc*. 7, 500–507
926 (2012).

927 55. O. Delaneau, H. Ongen, A. A. Brown, A. Fort, N. I. Panousis, E. T. Dermitzakis, A complete tool set for
928 molecular QTL discovery and analysis. *Nat. Commun*. 8, 15452 (2017).

929 56. B. L. Pierce, S. Burgess, Efficient Design for Mendelian Randomization Studies: Subsample and 2-Sample
930 Instrumental Variable Estimators. *Am. J. Epidemiol.* 178, 1177–1184 (2013).

931 **Acknowledgments**

932 We thank Majd Alsayed, Miguel Brown, Dominic Fitzgerald, Amber Thomas, and Mimi
933 Brown for their assistance with RNA-Seq data production and processing. We thank Richard
934 Kopp at SUNY Upstate Medical University for his help with wordsmithing. We thank all of the
935 donors and their relatives participated in the brain collections at Stanley Medical Research
936 Institute and Banner Sun Health Research Institute. We thank Drs. Maree J. Webster and Thomas
937 Beach for their supports to make the tissue available. We thank the High Performance Computing
938 Center of Central South University for partial support of calculation.

939 **Funding:**

940 SUNY Empire Innovation Program
941 National Institutes of Health grants U01MH122591
942 National Institutes of Health grants U01MH116489
943 National Institutes of Health grants R01MH110920
944 National Institutes of Health grants R01MH126459
945 National Institutes of Health grants U01MH103340
946 National Institutes of Health grants R01MH109715
947 National Institutes of Health grants R01MH110555
948 National Institutes of Health grants R21MH129817
949 National Institutes of Health grants R01GM139980
950 National Natural Science Foundation of China (Grants Nos. 82022024, 31970572)
951 science and technology innovation Program of Hunan Province (2021RC4018,
952 2021RC5027)

953 **Author contributions:**

954 Conceptualization: CL, KPW, JP, SHW, BL, XW
955 Methodology: AWS, LC, DZ, FW, MX, MN, DP, YJW, RV, CZ, KG, GG

Investigation: SHW, QL, YJ, CC, CY, AWS

Visualization: QL

Supervision: SHW, CL, CC, BL, KPW

Writing—original draft: SHW, QL, YJ, CL

Writing—review & editing: SHW, CL, CC, QL

Competing interests: Authors declare that they have no competing interests.

Data and materials availability: All raw data from psychENCODE BrainGVEX project are described in https://www.synapse.org/#!Synapse:syn5613798. All code, materials and results used in the analyses are available in https://www.synapse.org/#!Synapse:syn51664324. All data are available in the main text or the supplementary materials.

Figures and Tables

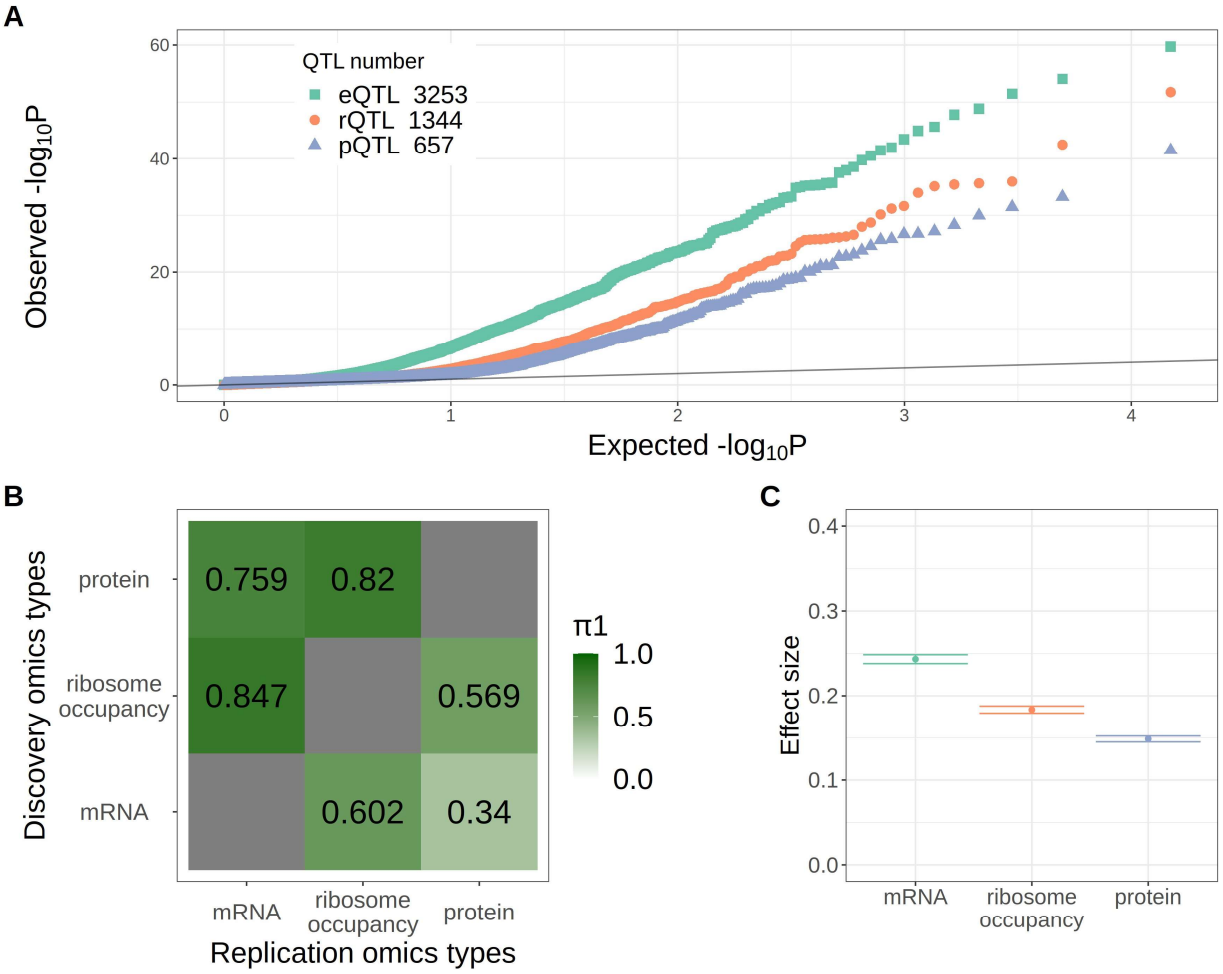


Fig. 1. Genetic regulation of gene expression in the human brain. (A) P-value quantile-quantile plot between the observed (Y-axis) and the expected based on null distribution (X-axis). The black line indicates the expected distribution of p values when there are no real QTL signals. The number of cis-QTLs (i.e. the most significantly associated SNP for each gene) identified at 10% FDR is labeled in the top left inset. **(B)** Replication rate between QTL types. Proportions of QTLs replicated in the other two omics types are listed in the 3X3 matrix. Each row is a discovery omics type and each element of the row correspond to the proportion QTL signals replicated in the omics type specified by the column label. For example, only 34% of the eQTL signals were replicated in the protein data. **(C)** Effect size of CMC eQTL SNPs in BrainGVEX data. Mean and

979 95% confidence interval of absolute per allele effect across all CMC eQTL SNPs that were also

980 analyzed in the BrainGVEX union set is shown.

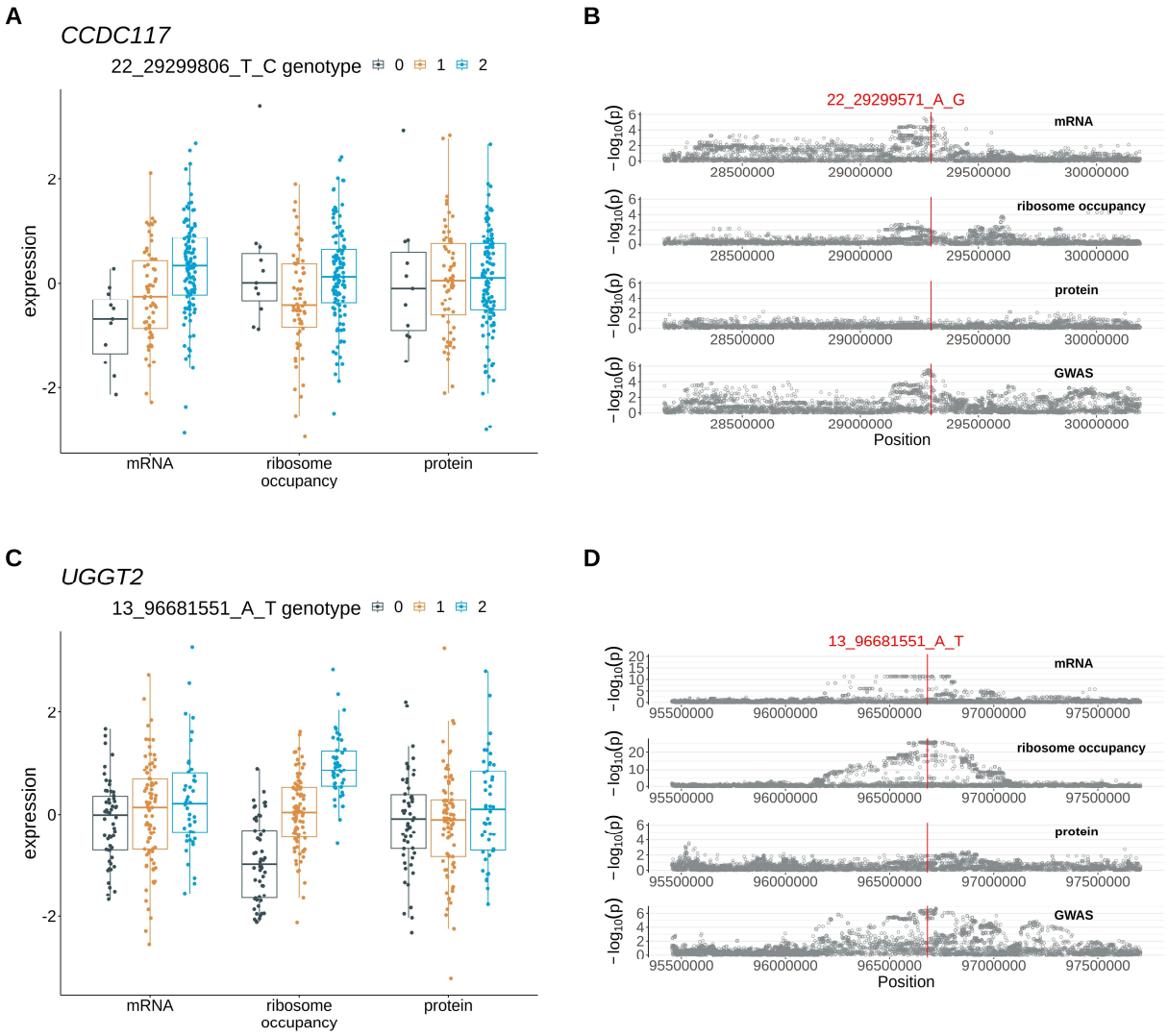


Fig. 2. Signal colocalization between schizophrenia GWAS and omics-specific QTLs. (A, C) Boxplots summarizing normalized gene expression level stratified by QTL genotypes for *CCDC117* esQTLs (A) and *UGGT2* rsQTLs (C). **(B, D)** Manhattan plots showing p value distribution for each QTL type and schizophrenia GWAS for the 1Mb QTL mapping window flanking *CCDC117* (B) and *UGGT2* (D). The red line indicates the position of the lead colocalization SNP between omics-specific QTLs and schizophrenia GWAS.

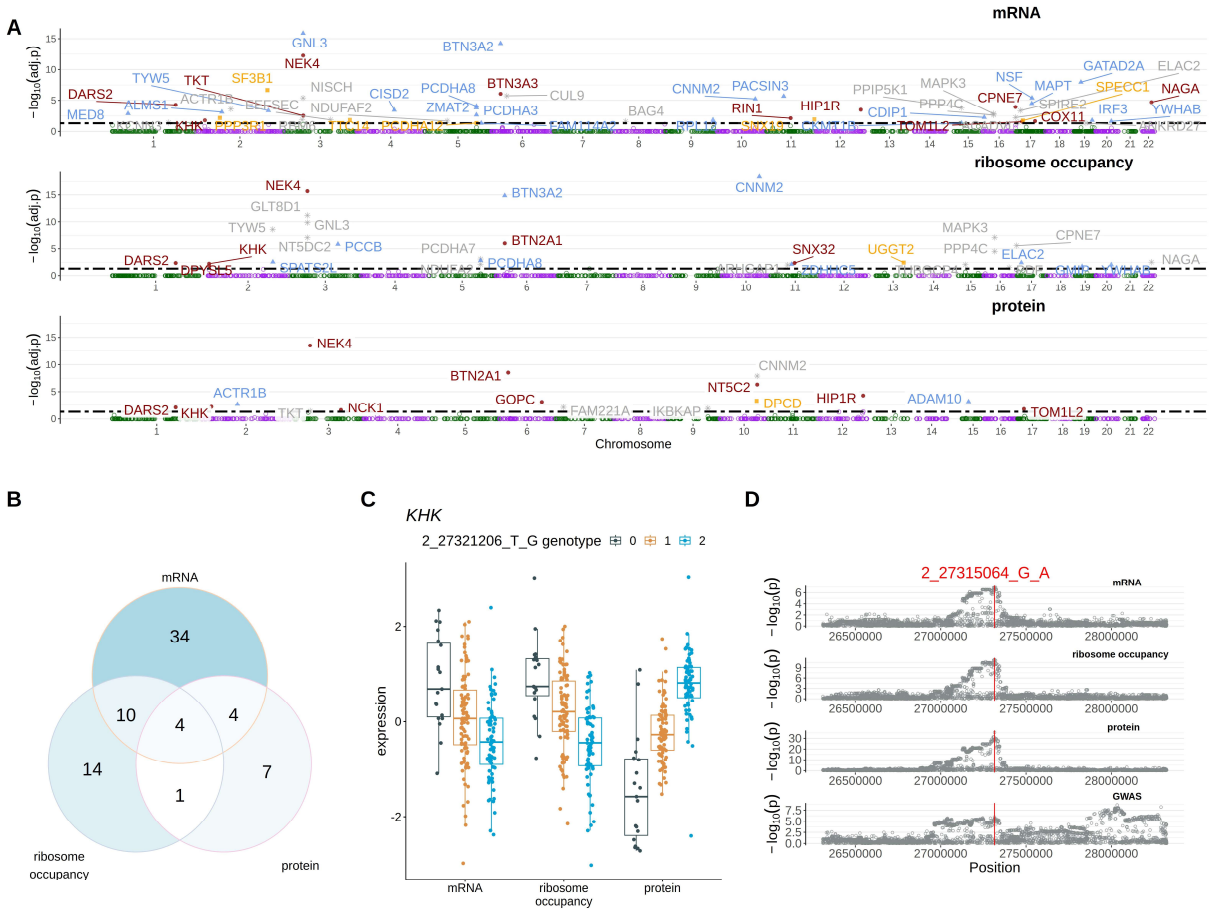
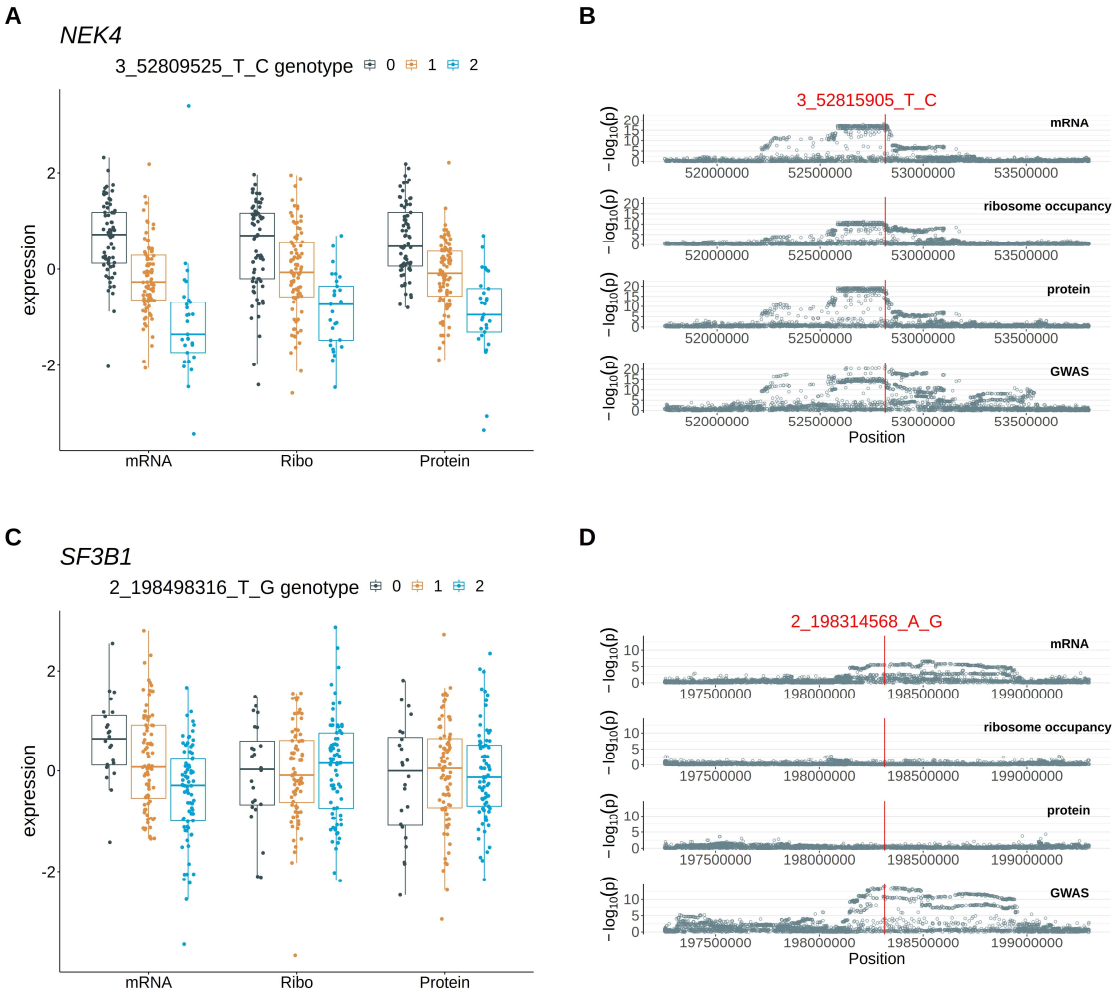


Fig. 3. Schizophrenia risk genes identified from each of the three omics types RNA-Seq (mRNA), ribo-seq (ribosome occupancy), and proteomics (protein) using S-PrediXcan. (A) Manhattan plots showing significance level (i.e. -log10 FWER from S-PrediXcan) of gene-schizophrenia association across the genome for genes that pass the 5% FWER significance cutoff. The black horizontal dotted line indicated the significance cutoff. Risk genes are color-coded according to the MR test results. Grey asterisks mark the risk genes that failed the two-sample MR tests; dark red solid circle marks the risk genes that pass both one-sample MR tests (passing both one-sample MR tests suggest that transcriptionally regulated protein level differences between individuals drives the disease risk); blue triangle marks the risk genes that pass one of the two one-sample MR tests; orange rectangle marks the risk genes that failed both

1002 one-sample MR tests. **(B)** Venn diagram illustrating the number and corresponding percentage of
1003 overlapping risk genes between omics types. **(C)** Boxplots summarizing normalized gene
1004 expression level stratified by eQTL genotypes for *KHK*. **(D)** Manhattan plots showing p value
1005 distribution for each QTL type and schizophrenia GWAS for the 1Mb QTL mapping window
1006 flanking *KHK*. The red line indicates the position of the lead colocalization SNP between eQTLs
1007 and schizophrenia GWAS.
1008

1009



1010

1011 **Fig. 4. Signal colocalization between schizophrenia GWAS and eQTLs of example risk**

1012 **genes**. **(A, C)** Boxplots summarizing normalized gene expression level stratified by QTL

1013 genotypes for *NEK4* eQTLs (A) and *SF3B1* eQTLs (C). **(B, D)** Manhattan plots showing p value

1014 distribution for each QTL type and schizophrenia GWAS for the 1Mb QTL mapping window

1015 flanking *NEK4* (B) and *SF3B1* (D). The red line indicates the position of the colocalization lead

1016 SNP between eQTLs and schizophrenia GWAS.

1017