

## RESEARCH ARTICLE SUMMARY

PSYCHENCODE2

# Cross-ancestry atlas of gene, isoform, and splicing regulation in the developing human brain

Cindy Wen, Michael Margolis, Ruija Dai, Pan Zhang, Paweł F. Przytycki, Daniel D. Vo, Arjun Bhattacharya, Nana Matoba, Miao Tang, Chuan Jiao, Minsoo Kim, Ellen Tsai, Celine Hoh, Nil Aygün, Rebecca L. Walker, Christos Chatzinakos, Declan Clarke, Henry Pratt, PsychENCODE Consortium, Mette A. Peters, Mark Gerstein, Nikolaos P. Daskalakis, Zhiping Weng, Andrew E. Jaffe, Joel E. Kleinman, Thomas M. Hyde, Daniel R. Weinberger, Nicholas J. Bray, Nenad Sestan, Daniel H. Geschwind, Kathryn Roeder, Alexander Gusev, Bogdan Pasaniuc, Jason L. Stein, Michael I. Love, Katherine S. Pollard, Chunyu Liu\*, Michael J. Gandal\*

**INTRODUCTION:** Genome-wide association studies (GWAS) have identified thousands of loci associated with neurodevelopmental and psychiatric disorders, yet our lack of understanding of the target genes and biological mechanisms underlying these associations remains a major challenge. GWAS signals for many neuropsychiatric disorders, including autism spectrum disorder, schizophrenia, and bipolar disorder, are particularly enriched for gene-regulatory elements active during human brain development. However, the lack of a unified population-scale, ancestrally diverse gene-regulatory atlas of human brain development has been a major

obstacle for the functional assessment of top loci and post-GWAS integrative analyses.

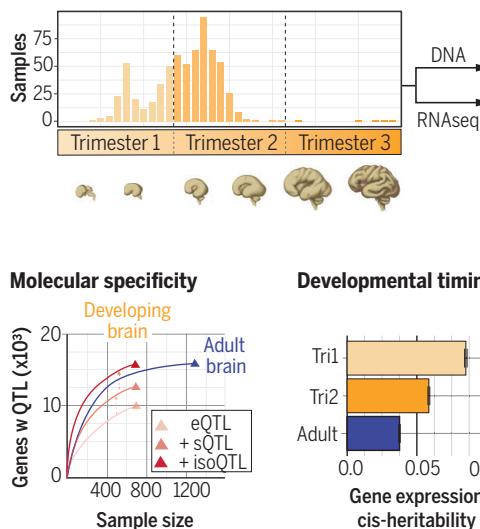
**RATIONALE:** To address this critical gap in knowledge, we have uniformly processed and systematically characterized gene, isoform, and splicing quantitative trait loci (cumulatively referred to as xQTLs) in the developing human brain across 672 unique samples from 4 to 39 postconception weeks spanning European, African-American, and Latino/admixed American ancestries). With this expanded atlas, we sought to specifically localize the timing and molecular features mediating the greatest pro-

portion of neuropsychiatric GWAS heritability, to prioritize candidate risk genes and mechanisms for top loci, and to compare with analogous results using larger adult brain functional genomic reference panels.

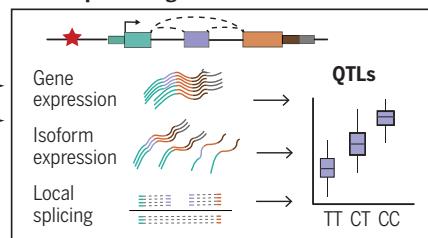
**RESULTS:** In total, we identified 15,752 genes harboring a gene, isoform and/or splicing *cis*-xQTL, including 49 genes associated with four large, recurrent inversions. Highly concordant effect sizes were observed across populations, and our diverse reference panel improved resolution to fine-map underlying candidate causal regulatory variants. Substantially more genes were found to harbor QTLs in the first versus second trimester of brain development, with a notable drop in gene expression and splicing heritability observed from 10 to 18 weeks coinciding with a period of rapidly increasing cellular heterogeneity in the developing brain. Isoform-level regulation, particularly in the second trimester, mediated a greater proportion of heritability across multiple psychiatric GWASs compared with gene expression regulation. Through colocalization and transcriptome-wide association studies, we prioritized biological mechanisms for ~60% of GWAS loci across five neuropsychiatric disorders, with >2-fold more colocalizations observed compared with larger adult brain functional genomic reference panels. We observed convergence between common and rare-variant associations, including a cryptic splicing event in the high-confidence schizophrenia risk gene *SP4*. Finally, we constructed a comprehensive set of developmentally regulated gene and isoform coexpression networks harboring unique cell-type specificity and genetic enrichments. Leveraging this cell-type specificity, we identified >8000 module interaction QTLs, many of which exhibited additional GWAS colocalizations. Overall, neuropsychiatric GWASs and rare variant signals localized more strongly within maturing excitatory- and interneuron-associated modules compared with those enriched for neural progenitor cell types. Results can be visualized at [devbrainhub.gandallab.org](https://devbrainhub.gandallab.org).

Downloaded from https://www.science.org at University Van Amsterdam on April 13, 2025

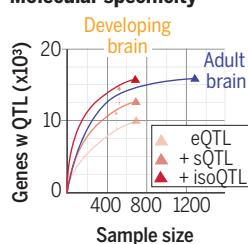
## Developing human neocortex



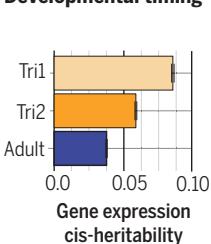
## Transcriptome regulation



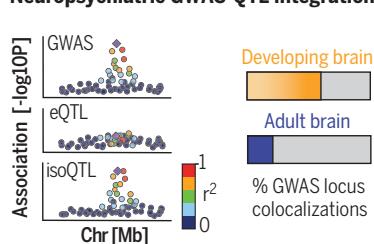
## Molecular specificity



## Developmental timing



## Neuropsychiatric GWAS-QTL integration



**A comprehensive transcriptome regulatory atlas of the developing human neocortex.** RNA-sequencing and single-nucleotide polymorphism genotypes were uniformly integrated within a diverse set of 672 samples of the developing human neocortex. Gene regulation was systematically assessed across the gene, isoform expression, and local splicing levels, yielding 15,752 genes harboring a significant xQTL. Gene regulation was highly dynamic, with a substantial drop observed in gene expression heritability over development. Integrative analyses with neuropsychiatric GWASs uncovered hundreds of candidate risk genes and mechanisms, providing insights into the cellular, molecular, and developmental specificity underlying disease-associated genetic variation.

The list of author affiliations is available in the full article online.

\*Corresponding author. Email: liuch@upstate.edu (C.L.); michael.gandal@pennmedicine.upenn.edu (M.J.G.)

Cite this article as C. Wen et al., Science 384, eadh0829 (2024). DOI: 10.1126/science.adh0829

**S** READ THE FULL ARTICLE AT  
<https://doi.org/10.1126/science.adh0829>

## RESEARCH ARTICLE

## PSYCHENCODE2

# Cross-ancestry atlas of gene, isoform, and splicing regulation in the developing human brain

Cindy Wen<sup>1,2,3</sup>, Michael Margolis<sup>2,3</sup>, Ruija Dai<sup>4</sup>, Pan Zhang<sup>2,3</sup>, Paweł F. Przytycki<sup>5</sup>, Daniel D. Vo<sup>2,3,6,7</sup>, Arjun Bhattacharya<sup>8,9</sup>, Nana Matoba<sup>10,11</sup>, Miao Tang<sup>6,7</sup>, Chuan Jiao<sup>4,12</sup>, Minsoo Kim<sup>2,3</sup>, Ellen Tsai<sup>2,3</sup>, Celine Hoh<sup>2,3</sup>, Nil Aygün<sup>10,11</sup>, Rebecca L. Walker<sup>1,2,3</sup>, Christos Chatzinakos<sup>13,14,15</sup>, Declan Clarke<sup>16</sup>, Henry Pratt<sup>17</sup>, PsychENCODE Consortium†, Mette A. Peters<sup>18</sup>, Mark Gerstein<sup>16,19,20,21</sup>, Nikolaos P. Daskalakis<sup>13,14,15</sup>, Zhiping Weng<sup>17</sup>, Andrew E. Jaffe<sup>22,23,24,25,26,27,28</sup>, Joel E. Kleinman<sup>22,23</sup>, Thomas M. Hyde<sup>22,23,29</sup>, Daniel R. Weinberger<sup>22,23,24,25,29</sup>, Nicholas J. Bray<sup>30</sup>, Nenad Sestan<sup>31,32</sup>, Daniel H. Geschwind<sup>3,33,34</sup>, Kathryn Roeder<sup>35,36</sup>, Alexander Gusev<sup>37,38,39,40</sup>, Bogdan Pasaniuc<sup>1,3,8,34,41</sup>, Jason L. Stein<sup>10,11</sup>, Michael I. Love<sup>10,42</sup>, Katherine S. Pollard<sup>5,43,44</sup>, Chunyu Liu<sup>4,45\*</sup>, Michael J. Gandal<sup>1,2,3,6,7\*</sup>

Neuropsychiatric genome-wide association studies (GWASs), including those for autism spectrum disorder and schizophrenia, show strong enrichment for regulatory elements in the developing brain. However, prioritizing risk genes and mechanisms is challenging without a unified regulatory atlas. Across 672 diverse developing human brains, we identified 15,752 genes harboring gene, isoform, and/or splicing quantitative trait loci, mapping 3739 to cellular contexts. Gene expression heritability drops during development, likely reflecting both increasing cellular heterogeneity and the intrinsic properties of neuronal maturation. Isoform-level regulation, particularly in the second trimester, mediated the largest proportion of GWAS heritability. Through colocalization, we prioritized mechanisms for about 60% of GWAS loci across five disorders, exceeding adult brain findings. Finally, we contextualized results within gene and isoform coexpression networks, revealing the comprehensive landscape of transcriptome regulation in development and disease.

**T**housands of genetic risk loci have been robustly associated with neurodevelopmental and psychiatric disorders by large-scale genome-wide association studies (GWASs) (1, 2). However, because most associated GWAS variants reside within non-coding regions of the human genome, often in large linkage disequilibrium (LD) blocks, the true underlying causal variant(s) and target gene(s) remain largely unknown. Therefore,

the critical defining obstacle of the post-GWAS era is to pinpoint the specific, locus-level molecular impact of GWAS variants at scale (3). Because risk loci are enriched in the regulatory regions of the human genome (4–6), one major approach to addressing this challenge has been to connect risk variants with tissue-specific reference panels of gene expression quantitative trait loci (eQTLs) through statistical colocalization, transcriptome-wide asso-

ciation studies (TWASs), and related approaches (7–9). This has prompted several large-scale efforts (10–16) to generate comprehensive functional genomic compendia connecting population-level allelic variation with gene expression profiles in the human brain. However, although these resources have provided biologically interpretable and meaningful annotations for dozens of neuropsychiatric risk loci, most remain mechanistically unannotated (15, 17).

Gene regulation is highly dependent on the specific underlying developmental stage, tissue, and cellular context (7, 17–20), and increasing evidence implicates the developing human brain in the genetic risk for neuropsychiatric disorders, including autism spectrum disorder (ASD) and schizophrenia (SCZ) (21–23). Therefore, several efforts have begun to characterize the genetic control of gene expression during human brain development, finding that gene regulation is tightly controlled during this process and is enriched for neuropsychiatric risk (11, 12, 24–28). However, because these studies are individually small, the power to pinpoint the developmental timing or elucidate the full extent of gene regulation in the developing human brain has been limited.

The regulation of transcript-isoform structure and diversity, through alternative local splicing and differences in transcriptional start sites (TSS) and termination sites, is also known to be a mechanism critical for human brain development that is implicated in disease pathogenesis (14, 29–32). Indeed, relative to other tissues and species, the landscape of alternative splicing is particularly extensive in the human brain, and its regulation is notably distinct from that of gene expression (16, 33). Although individual studies have begun to

\*Corresponding author. Email: liuch@upstate.edu (C.L.); michael.gandal@pennmedicine.upenn.edu (M.J.G.)

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

<sup>1</sup>Interdepartmental Program in Bioinformatics, University of California, Los Angeles, Los Angeles, CA 90095, USA. <sup>2</sup>Department of Psychiatry, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA 90095, USA. <sup>3</sup>Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA 90095, USA. <sup>4</sup>Department of Psychiatry, SUNY Upstate Medical University, Syracuse, NY 13210, USA. <sup>5</sup>Gladstone Institute of Data Science and Biotechnology, San Francisco, CA 94158, USA. <sup>6</sup>Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA. <sup>7</sup>Lifespan Brain Institute, The Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA. <sup>8</sup>Department of Pathology and Laboratory Medicine, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA 90095, USA. <sup>9</sup>Institute for Quantitative and Computational Biosciences, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA 90095, USA. <sup>10</sup>Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. <sup>11</sup>UNC Neuroscience Center, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. <sup>12</sup>Université Paris Cité, Institute of Psychiatry and Neuroscience of Paris (IPNP), INSERM U1266, Team Krebs, 75014 Paris, France. <sup>13</sup>Department of Psychiatry, Harvard Medical School, Boston, MA 02215, USA. <sup>14</sup>McLean Hospital, Belmont, MA 02478, USA. <sup>15</sup>Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA. <sup>16</sup>Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520, USA. <sup>17</sup>Program in Bioinformatics and Integrative Biology, University of Massachusetts Medical School, Worcester, MA 01650, USA. <sup>18</sup>CNS Data Coordination Group, Sage Bionetworks, Seattle, WA 98109, USA. <sup>19</sup>Program in Computational Biology and Bioinformatics, Yale University, New Haven, CT 06520, USA. <sup>20</sup>Department of Computer Science, Yale University, New Haven, CT 06520, USA. <sup>21</sup>Department of Statistics and Data Science, Yale University, New Haven, CT 06520, USA. <sup>22</sup>Lieber Institute for Brain Development, Baltimore, MD 21205, USA. <sup>23</sup>Department of Psychiatry & Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. <sup>24</sup>Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. <sup>25</sup>Department of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. <sup>26</sup>Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, USA. <sup>27</sup>Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, USA. <sup>28</sup>Neumora Therapeutics, Watertown, MA 02472, USA. <sup>29</sup>Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. <sup>30</sup>MRC Centre for Neuropsychiatric Genetics & Genomics, Division of Psychological Medicine & Clinical Neurosciences, Cardiff University School of Medicine, Cardiff CF24 4HQ, UK. <sup>31</sup>Department of Comparative Medicine, Yale University School of Medicine, New Haven, CT 06520, USA. <sup>32</sup>Department of Neuroscience, Yale University School of Medicine, New Haven, CT 06520, USA. <sup>33</sup>Program in Neurogenetics, Department of Neurology, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA 90095, USA. <sup>34</sup>Institute for Precision Health, University of California, Los Angeles, Los Angeles, CA 90095, USA. <sup>35</sup>Department of Statistics & Data Science, Carnegie Mellon University, Pittsburgh, PA 15213, USA. <sup>36</sup>Computational Biology Department, Carnegie Mellon University, Pittsburgh, PA 15213, USA. <sup>37</sup>Department of Medical Oncology, Division of Population Sciences, Dana-Farber Cancer Institute, Boston, MA 02215, USA. <sup>38</sup>Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA. <sup>39</sup>Harvard Medical School, Boston, MA 02115, USA. <sup>40</sup>Division of Genetics, Brigham and Women's Hospital, Boston, MA 02215, USA. <sup>41</sup>Department of Computational Medicine, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA 90095, USA. <sup>42</sup>Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. <sup>43</sup>Department of Epidemiology & Biostatistics, University of California, San Francisco, San Francisco, CA 94158, USA. <sup>44</sup>Chan Zuckerberg Biohub, San Francisco, CA 94158, USA. <sup>45</sup>Center for Medical Genetics & Human Key Laboratory of Medical Genetics, School of Life Sciences, Central South University, Changsha, Hunan 410008, China.

investigate splicing QTLs (sQTLs) and isoform QTLs (isoQTLs) in the developing brain (24, 25), a lack of uniform data processing has precluded a systematic characterization of these critical mechanisms and their potential relationships with genetic risk factors for neuropsychiatric disorders.

To address these critical gaps, we present a comprehensive investigation into the genetic regulation of the transcriptome during human brain development by uniformly processing data from 672 distinct samples spanning 4 to 39 postconception weeks (PCW), most of which are within the first (Tri1) and second (Tri2) trimesters. The full integrated dataset is highly diverse, comprising several major continental ancestries including European (EUR; 45%), Latino/admixed American (AMR; 25%), African-American (AFR; 22%), and East Asian/Southeast Asian (EA/SEA; 8%). This well-powered, cross-ancestry resource provides an extensive view of the developmental timing and regulatory landscape of human brain development at gene, isoform, and splicing levels. We observed a substantial drop in the heritability of gene expression and local splicing with development, particularly from 10 to 18 PCW. However, psychiatric GWAS signals were more concentrated within Tri2. We found that isoform-level regulation mediated substantially more psychiatric GWAS heritability than gene regulation. We prioritized candidate risk genes and molecular mechanisms for ~60% of GWAS loci across five neuropsychiatric disorders. Finally, we contextualized these risk mechanisms within a comprehensive set of gene- and isoform-level coexpression networks across human brain development. Results can be visualized at [devbrainhub.gandallab.org](http://devbrainhub.gandallab.org).

### Large-scale functional genomic investigation of human brain development

Here, we present a comprehensive investigation of gene expression, splicing, and isoform regulation during human brain development by integrating data across five individual cohorts (12, 24–27) (fig. S1) comprising 672 distinct donors. After uniform data processing, including imputation into the multi-ancestry TOPMed reference panel (34), and strict quality control (35) (fig. S2), we retained 654 samples with matched genotype and cortical RNA-sequencing (RNA-seq) data spanning 4 to 39 PCW, most falling within the first two trimesters (Tri1,  $N = 216$ ; Tri2,  $N = 433$ ; table S1). Altogether, our large sample size enabled the detection of 10,094 genes with at least one *cis*-eQTL at a false discovery rate (FDR) adjusted  $P < 0.05$  (called “eGenes”) after permutation-based correction for local LD structure [Fig. 1A, fig. S3, and table S1 (35)].

Compared with adult brain eQTL datasets from PsychENCODE ( $N = 1,387$ ) (13, 14) and

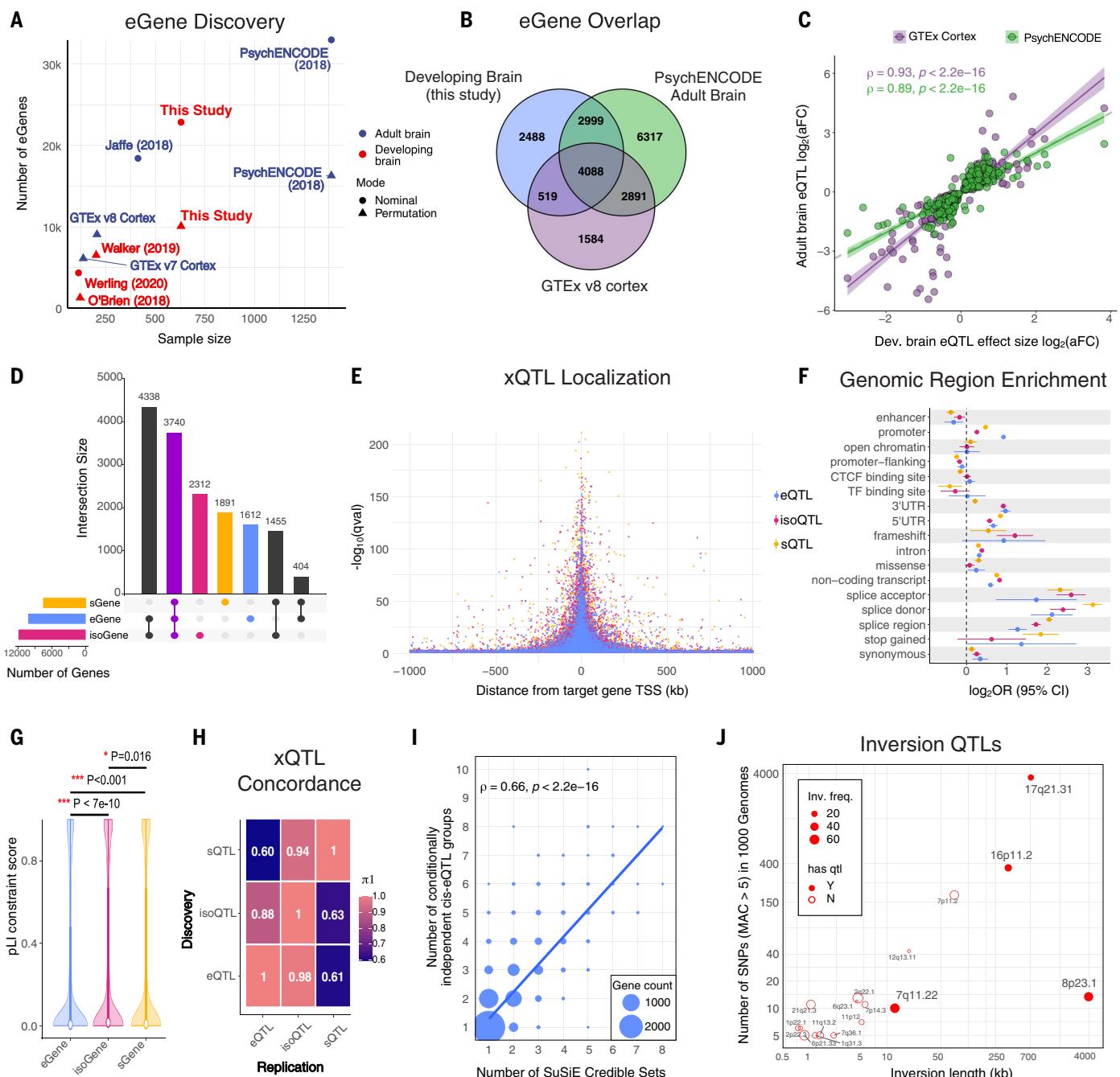
GTEX ( $N = 205$ ) (16), we identified 2488 “fetal-specific” eGenes in this dataset (Fig. 1B and table S1). Among eQTL-eGene pairs conserved between the developing and adult human brain cohorts, we observed strong correlations in effect size as measured by allelic fold change (GTEX: Spearman rho = 0.93,  $P < 2.2 \times 10^{-16}$ , PsychENCODE: Spearman rho = 0.89,  $P < 2.2 \times 10^{-16}$ ; Fig. 1C and fig. S3). Notably, these 2488 fetal-specific eGenes were significantly more intolerant to loss-of-function mutations, as measured by the pLI score (36), compared with eGenes shared between developing and adult time points (Wilcoxon  $P < 2 \times 10^{-16}$ ; fig. S4), suggesting that these context-specific annotations may be more relevant for interpretation of neuropsychiatric GWAS signals, which are under negative selection (37, 38). These fetal-specific eGenes were enriched for spliceosomal pathways and cell type-specific marker genes (39) for a subtype of ventricular radial glia [odds ratio (OR) = 3.3, FDR-corrected  $P = 0.054$ , Fisher’s exact test; fig. S4 and table S1].

Dysregulation of RNA splicing has been strongly implicated in complex disease risk (29), as well as alterations in brain development (24). To systematically investigate RNA splicing regulation in the developing human brain, we used two complementary approaches. First, guided by existing GENCODE reference transcriptome annotations (40), we imputed isoform expression from short-read RNA-seq using Salmon (41). Second, we profiled local alternative splicing events using LeafCutter (42), an annotation-free method that quantifies intron excision ratios. Next, we identified genetic variants that are associated with these isoform- and splicing-level quantitative traits [figs. S5 and S6 (35)], identifying 11,845 and 7490 genes harboring a *cis*-isoQTL or *cis*-sQTL at FDR-corrected  $P < 0.05$  (called “isoGenes” and “sGenes”; table S1), respectively. Among the identified xQTL-containing genes (where xQTL is the sum of eQTL + sQTL + isoQTL), 3740 are shared, whereas 2312 and 1891 only have QTLs detected at the isoform and local splicing levels, respectively (Fig. 1D). As expected, *cis*-xQTLs are strongly clustered near their target genes’ TSS (Fig. 1E) and are correspondingly enriched in functional regions of the genome (Fig. 1F and table S1). eQTLs, isoQTLs, and sQTLs are all significantly differentially enriched in promoter and splicing regions ( $P < 0.05$  with Bonferroni correction; table S1). In addition, we observed that isoGenes and sGenes are more intolerant to loss-of-function mutations than eGenes (Fig. 1G). To investigate the variant-level overlap among *cis*-eQTLs, *cis*-isoQTLs, and *cis*-sQTLs, we calculated the pairwise Storey’s pi<sub>1</sub> statistic (43) that measures the proportion of *cis*-eQTLs, *cis*-isoQTLs, and *cis*-sQTLs in the “Discovery” group (permutation  $q$  value  $< 0.05$ ) that also

exhibit true associations in the “Replication” group (nominal  $P$ s from all association pairs). We found that genetic variants directly associated with RNA splicing (especially isoQTLs) capture orthogonal signals from eQTLs [Fig. 1H (35)]. Finally, to investigate the potential impact of the different library preparations and sequencing depths of the individual substudies, we remapped sQTLs in the Walker (24) and HDBR (27) datasets using our uniform processing pipeline. We observed a high replication rate of our full sQTL results within each individual dataset (Walker pi<sub>1</sub> = 0.93; HDBR pi<sub>1</sub> = 0.89), indicating that the splicing regulation detected in our combined analysis was robust to these potential technical factors.

We next sought to characterize the extent of allelic heterogeneity in the developing human brain, in which multiple potential causal variants regulate gene expression at a given locus. Through stepwise conditional QTL mapping (44) [fig. S3 and table S1 (35)], we identified multiple independent regulatory signals for 3570 eGenes in the developing brain, with some exhibiting up to 10 groups of conditionally independent *cis*-eQTL signals. As expected, the primary *cis*-eQTL (median -0.0340 kb) was closer to the corresponding target gene’s TSS than was the secondary *cis*-eQTL (median 0.4820 kb; Wilcoxon rank sum  $P < 2.2 \times 10^{-16}$ ). The same was also true for tertiary and higher-rank *cis*-eQTLs. As an orthogonal approach, we also performed statistical fine-mapping with SuSiE (45), which estimates credible sets (CSs) of candidate causal variants. The number of SuSiE-estimated CSs was highly concordant with the number of conditionally independent QTL signals for each eGene (Spearman rho 0.66,  $P < 2.2 \times 10^{-16}$ ; Fig. 1I), indicating that fine-mapping is a complementary approach to conditional QTL mapping for detecting independent regulatory signals. Each CS prioritized a median of five single-nucleotide polymorphisms (SNPs), with 2423 eGenes containing exactly one SNP (extended data can be accessed at <https://doi.org/10.7303/syn50897018.5>).

Compared with SNPs and indels, the impact of other classes of genetic variation, such as structural variants, on downstream gene expression remains underexplored. Nevertheless, complex structural variants such as large recurrent inversions have known associations with brain-relevant traits and can affect gene expression extensively (25, 46). To address this, we imputed the genotypes of 17 common [minor allele frequency (MAF)  $> 0.05$ ] inversions into our uniformly processed developing brain data and quantified their effects on gene expression across the transcriptome. Longer inversions were more likely to affect downstream gene expression. We found 49 inversion-associated eQTLs (inv-eQTLs), both



**Fig. 1. The landscape of gene, splicing, and isoform regulation in the developing human brain.** (A) Number of eGenes versus sample size discovered here compared with other human brain studies (12–14, 16, 24–27). (B) Overlap of eGenes between developing brain ( $N = 629$ ), GTEx v8 Brain Cortex ( $N = 205$ ), and PsychENCODE adult brain studies ( $N = 1387$ ). (C) Correlation of eQTL effect size, measured by allelic fold change (aFC), between developing and adult brain datasets. Each dot is a shared eGene-primary eQTL pair between the developing brain and GTEx (247 pairs) or PsychENCODE (253 pairs) datasets. (D) Overlap among eGenes, isoGenes, and sGenes. (E) Distance between the TSS of each target gene of cis-eQTL, cis-isoQTL, and cis-sQTL SNPs. (F) Enrichment of cis-eQTL, cis-isoQTL, and cis-sQTLs within functional regions of the genome.

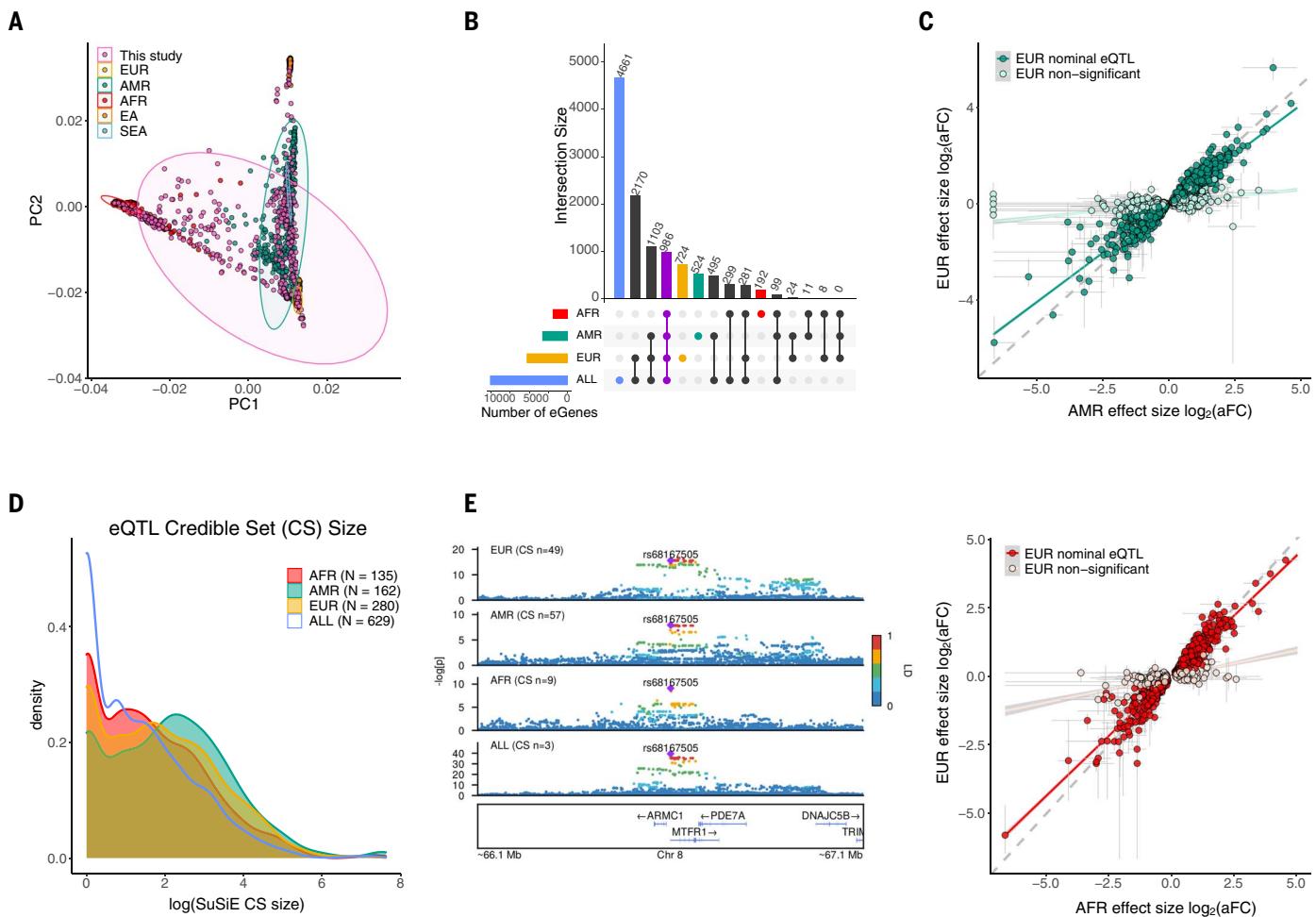
in *cis* and in *trans*, for four of the inversions located at 17q21.31, 16p11.2, 7q11.22, and 8p23.1 [Fig. 1J, fig. S7, and table S1 (35)].

#### Cross-ancestry gene regulation and fine-mapping

Differences in genetic variation (e.g., allele frequency and LD) across ancestries (47) have

(G) Loss-of-function mutation intolerance, as measured by pLI score, of eGenes, isoGenes, and sGenes. isoGenes and sGenes exhibit significantly less tolerance to loss-of-function mutations than eGenes (Wilcoxon). (H) Storey's pI statistic of the proportion of true associations in the discovery group of QTL (y axis; permutation  $q$  value  $< 0.05$ ) that are also true associations in the replication group of QTL (x axis; all nominal  $P$  values). (I) Number of fine-mapped credible sets discovered. The size of the dots is scaled to the number of genes. (J) Common recurrent inv-eQTLs in the developing brain. Inversions are displayed according to their length and the number of overlapping SNPs. Inversions with significantly associated eGenes have filled circles (FDR-adjusted  $P < 0.05$ ). The size of the circle indicates the population frequency of the inversions.

the potential to increase power for statistical fine-mapping (48, 49). After filtering and imputing the genotype data [fig. S2 (35)], we



**Fig. 2. Cross-ancestry gene regulation and fine-mapping.** (A) Genotype principal components analysis of the developing human brain samples. Sample ancestry was inferred by merging imputed genotypes with 1000 Genomes (47). (B) Comparison of eGenes discovered in the full cross-ancestry dataset ("ALL," N = 629) and in the separate subancestries, EUR (N = 280), AMR (N = 162), and AFR (N = 135).

(C) Correlation of eQTL effect sizes between AMR/AFR (top/bottom) and EUR, as measured by allelic fold change. Each dot is an AMR/AFR eGene-primary eQTL pair and is colored by its nominal significance in EUR. Gray lines denote the lower and upper bounds of aFC. (D) Comparison of fine-mapping credible set sizes between the ancestries. (E) *Cis* associations for the gene *MTFR1* in the cross-ancestry, EUR, AMR, and AFR datasets.

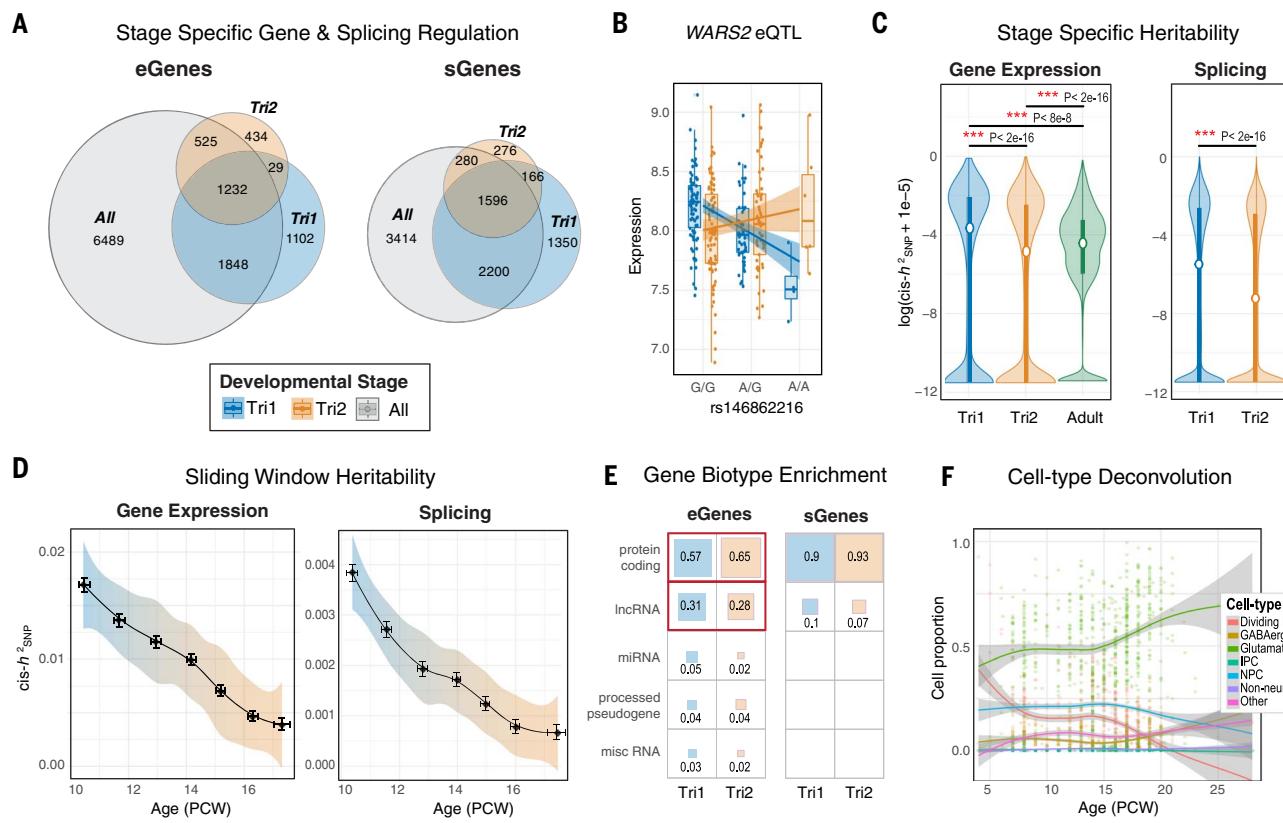
inferred genetic ancestries in the developing brain samples in reference to the 1000 Genomes panel (47). Of the 654 samples, 292 (44.6%) were labeled as EUR, 164 (25.1%) as AMR, 145 (22.2%) as AFR, 29 (4.4%) as SEA, and 24 (3.7%) as EA (Fig. 2A, fig. S1, and table S1). We next independently mapped xQTLs in the three largest ancestry groupings: EUR, AMR, and AFR. A total of 986 eGenes were shared between the multiancestry dataset and the three ancestry groupings (Fig. 2B and table S2). Among the shared eGene-eQTL pairs between ancestry groups, we observed highly consistent effect sizes (AMR-EUR Spearman rho = 0.97,  $P < 2.2 \times 10^{-16}$ ; AFR-EUR Spearman rho = 0.97,  $P < 2.2 \times 10^{-16}$ ; Fig. 2C). We then performed statistical fine-mapping separately in the ancestry groups (45) and showed that AFR, despite having the smallest sample size, had on average the smallest CSs of causal variants (Fig. 2D). Next, we found that the multi-

ancestry dataset substantially reduced the CS size. Taking gene *MTFR1* as an example (Fig. 2E), the gain in fine-mapping resolution was not solely due to the larger sample size, but also reflecting the distinct patterns of LD across groups. *MTFR1* has a complex LD structure in EUR and AMR, reflected in its large CS (EUR N = 49, AMR N = 57), whereas in AFR, with simpler LD, the CS is only nine variants. CS sizes are further reduced in multiancestry fine-mapping (N = 5), and in a *trans*-ancestry fine-mapping framework that leverages functional annotation (51, 52) (N = 3). With these results, we highlight the importance of multiancestry data in refining statistical fine-mapping.

#### Trimester-specific transcriptome regulation

Previous work has implicated the midfetal period of human brain development as a critical window convergently affected by multiple distinct rare genetic risk factors for ASD

and SCZ (21–23). Although common variant-mediated gene regulation is also known to be dependent on developmental context, this has not been systematically evaluated in the developing human brain. To address this critical gap, we conducted trimester-specific analyses to investigate the specificity of genetic regulation during distinct periods of brain development. We called QTLs after first separating samples into similarly powered Tri1 (4 to 13 PCW; EUR, N = 143) and Tri2 (14 to 26 PCW; EUR, N = 145) windows, although there were too few samples in Tri3 (N = 4) to conduct a similar analysis. We identified almost twofold more Tri1 eGenes (Tri1, N = 4211) than Tri2 eGenes (N = 2220), with 1261 eGenes shared between the two trimesters (Fig. 3, A and B; fig. S8; and table S3). Splicing regulation showed a similar pattern with >twofold more sGenes in Tri1 (N = 5312) than in Tri2 (N = 2318; Fig. 3A).



**Fig. 3. Trimester-specific patterns of gene expression and splicing regulation.**

(A) Comparison of eGenes and sGenes identified in Tri1, Tri2, and the full dataset. We identified many more eGenes and sGenes in Tri1 than in Tri2 despite similar sample sizes. (B) Example of a trimester-specific eGene for *WARS2*, where rs146862216 (G>A) is an eQTL in Tri1 ( $\beta = -0.89$ ,  $FDR = 3.88 \times 10^{-13}$ ) but not in Tri2 ( $\beta = -0.03$ ,  $P = 0.71$ ). (C) cis-Heritability of gene expression drops from Tri1 to Tri2 time points, as well as between developing and adult (PsychENCODE) samples (left). A similar drop is observed in splicing heritability between Tri1 and Tri2 (right).

(D) Sliding-window analysis of gene expression (left) and splicing (right) cis-heritability for samples from 10 to 18 weeks. Each dot represents a sliding set of temporally ordered samples ( $N = 150$ ), with mean age ( $\pm SD$ ) on the x axis and median  $cis-h^2_{SNP}$  ( $\pm SD$ ) on the y axis. (E) Comparison of gene biotype enrichment of Tri1-only and Tri2-only eGenes and sGenes. Values associated with each gene type represent the proportion of genes classified within that category. Red boxes highlight significant post hoc  $P$  value. (F) Estimated proportion of seven major cell classes over development assessed by bulk tissue cell-type deconvolution using CIBERSORTx (53).

To further explore this unexpected result, we next calculated the *cis*-window SNP-based heritability ( $cis-h^2_{SNP}$ ) for gene expression, a related measure that is more robust to sample size variation, in Tri1 and Tri2 samples, as well as in adult samples from PsychENCODE ( $N = 1,387$ ). Mirroring eQTL results, we found that *cis* heritability drops significantly from Tri1 to Tri2, with higher  $cis-h^2_{SNP}$  in both Tri1 and Tri2 compared with adult brain samples (all  $P < 2.2 \times 10^{-8}$ , Wilcoxon; Fig. 3C and table S3). To determine whether a particular developmental inflection point could contextualize this drop in heritability, we rank-ordered samples by PCW and performed a sliding window analysis of  $cis-h^2_{SNP}$  across development in equally powered batches ( $N = 150$ ). We again observed a substantial linear drop in  $cis-h^2_{SNP}$  across the entire range from 10 to 18 weeks (Fig. 3D). There were no detectable differences in the number of conditionally independent eQTLs (fig. S8), levels of gene expression, or phenotypic variance across Tri1 and Tri2 that

could have driven these results (fig. S9). To determine whether these observations extended beyond gene expression, we conducted similar analyses using sQTLs, because local splicing event quantifications are highly distinct from expression (29). Again, we observed a significant drop in local splicing heritability from Tri1 to Tri2 ( $P < 2 \times 10^{-16}$ , Wilcoxon), with a linear decrease from 10 to 18 weeks (Fig. 3, C and D; fig. S8; and table S3).

We next sought to investigate potential biological explanations for these observed stage-specific properties of developmental gene regulation. We first assessed whether Tri1 and Tri2 eGenes were differentially enriched among specific gene biotypes, observing a significant shift from noncoding to protein-coding biotypes ( $\chi^2 = 37.1$ ,  $df = 9$ ,  $P = 2.5 \times 10^{-5}$ ; Fig. 3E). However, the heritability drop remained significant when restricted to protein-coding genes, suggesting that coding status was not the primary driver of these findings (fig. S8). The 10 to 20 PCW period is also co-

incident with a shift in cell-type proportion from neural progenitor cells (NPCs) to distinct classes of migrating and maturing excitatory neurons. Thus, we hypothesized that the observed decrease in heritability could reflect genetic regulation in specific cell types as their proportions become more heterogeneous. To assess this, we deconvoluted bulk gene expression using a reference panel of cell type-specific marker genes derived from a meta-analysis of 2.95 million single cells and nuclei from the developing human brain (53, 54). As expected, estimated cell-type proportions were highly dynamic during this window, with a notable increase in excitatory neurons and a concomitant decrease in progenitor populations (Fig. 3F and fig. S8). Consistent with the cellular heterogeneity hypothesis, proportions of these cell types were significantly correlated over time with the drop in  $cis-h^2_{SNP}$  (Dividing\_progenitor:  $R = 0.96$ ,  $FDR = 0.005$ ; glutamatergic:  $R = -0.94$ ,  $FDR = 0.006$ , Pearson; fig. S8). However, heritability was calculated using

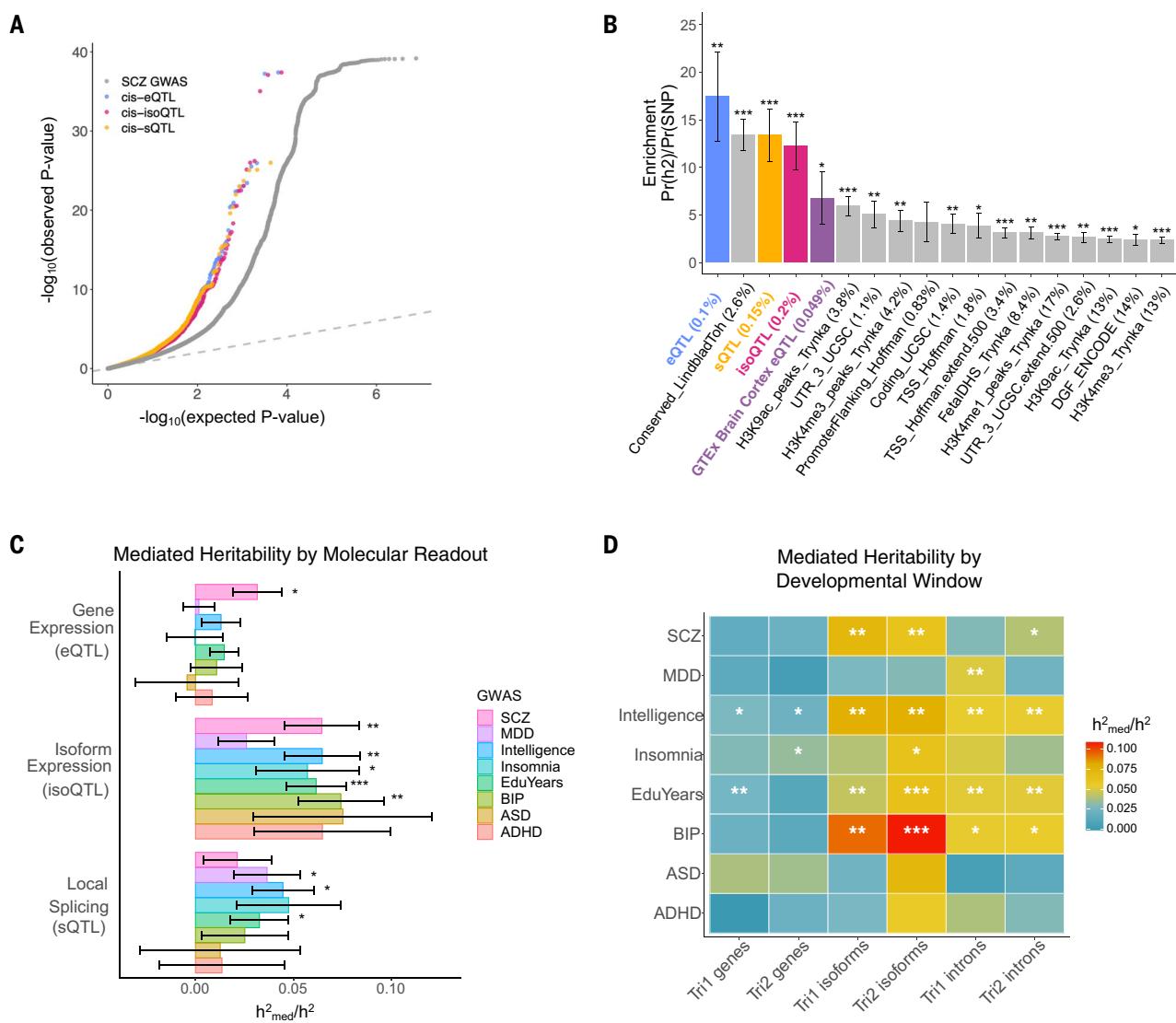
gene expression data corrected for hidden covariates (see the materials and methods and fig. S8) known to control for cell-type heterogeneity across samples (7, 55), which we verified (fig. S8), suggesting an indirect relationship. As an alternative hypothesis, the drop in heritability could reflect intrinsic properties of neurons as they mature. We tested this by comparing gene expression *cis-h<sup>2</sup>SNP* in a matched cohort of cultured primary human NPCs (phNPCs) and their differentiated neuronal progeny (56). Mirroring our findings, we observed significantly greater gene expression *cis-h<sup>2</sup>SNP* in phNPCs compared with their differentiated neuron counterparts ( $P = 0.0009$ ,

Wilcoxon rank sum test; fig. S8). This is consistent with a reported >twofold increase in eGenes (56) and noncoding regulatory elements (57) in phNPCs compared with neurons. Altogether, these results underscore the temporally dynamic, context-specific nature of gene regulation in the human brain, likely reflecting the complex interplay between changing cellular proportions and intrinsic properties of neuronal maturation.

### Temporal and molecular specificity of neuropsychiatric GWAS

To prioritize the potential underlying biological context(s) through which genetic risk

is conferred, we next sought to integrate our expanded set of functional genomic annotations with neuropsychiatric GWAS results. Using the well-powered SCZ GWAS as an example, we observed substantial enrichment of test statistics when subsetting to developing brain *cis*-eQTLs, *cis*-isoQTLs, and *cis*-sQTLs compared with background variants (Fig. 4A). We next used stratified LD-score regression [S-LDSC (6, 58)] to estimate the degree to which *h<sup>2</sup>SNP* was enriched among these xQTLs. We generated continuous annotations for genetic variants based on fine-mapping posterior inclusion probabilities (59), and found SCZ GWAS signals to be highly enriched among developing



**Fig. 4. Integrative analysis of xQTLs with neuropsychiatric GWAS results.**

(A) Quantile-quantile plot of SCZ GWAS  $P$  values, subsetted by top *cis*-eQTLs, *cis*-isoQTLs, and *cis*-sQTLs compared with all background GWAS SNPs. (B) S-LDSC enrichment of SCZ GWAS heritability within developing brain xQTLs (GTEx v8) compared with background functional annotations. The proportional genomic coverage of SNPs within each annotation is shown in parentheses. (C) Estimated proportion ( $\pm$ SE) of GWAS  $h^2_{\text{SNP}}$  mediated by the *cis* genetic component of gene, isoform, and intron (splicing) regulation. IsoQTLs mediate the greatest degree of heritability for multiple neuropsychiatric traits in the developing brain compared with eQTLs and sQTLs. (D) Estimated proportion of GWAS  $h^2_{\text{SNP}}$  mediated by the *cis* genetic component of trimester-stratified gene, isoform, and intron (splicing) regulation. For (B), (C), and (D), \*\*\*FDR < 0.001, \*\*FDR < 0.01, and \*FDR < 0.05.

brain regulatory elements, more so than in the adult brain [Fig. 4B and table S4 (35)]. When jointly modeling developing brain eQTLs, isoQTLs, and sQTLs, we observed that sQTLs and isoQTLs captured greater enrichment than eQTLs [fig. S10 and table S4 (35)]. These observations were consistent when extending across multiple neuropsychiatric GWAS, as well as within trimester-specific annotations (fig. S10 and table S4). Overall, these results highlight the importance of the developmental context, as well as splicing and isoform regulation, for interpreting potential mechanisms underlying psychiatric GWAS signals.

Although the S-LDSC results clearly demonstrated greater enrichment of  $h^2_{\text{SNP}}$  among prenatal (relative to postnatal) gene-regulatory variants, and splicing and isoform regulation (relative to total gene expression), these analyses are based on small genomic windows centered around top xQTLs. It can be difficult to compare results across annotations that differ substantially in genomic coverage, and it has been hypothesized that the top xQTLs may not overlap strongly with complex traits under negative selection (37, 38). To address these issues, we next leveraged the mediated expression score regression (MESC) framework (38) to estimate the proportion of heritability that is mediated by the *cis*-genetic component of assayed genes, isoforms, and introns ( $h^2_{\text{med}}/h^2_g$ ). Whereas S-LDSC is restricted to the subset of significant xQTLs, MESC estimates enrichments genome wide, including assayed features with low expression heritability (38). We chose five GWASs of brain-relevant neuropsychiatric traits: SCZ, ASD, bipolar disorder (BIP), attention deficit hyperactivity disorder (ADHD), and major depression disorder (MDD) (1, 2, 60–62). Across these traits, we consistently observed greater heritability mediated by isoform regulation (isoform  $h^2_{\text{med}}/h^2_g$  6.45 ± 1.92%) than that of gene expression (3.15 ± 1.25%) or splicing (2.13 ± 1.74%), although the overall extent of mediation remained small [Fig. 4C and table S4 (35)]. These findings remained consistent after excluding high pLI genes ( $\text{pLI} > 0.9$ ), indicating that these results are not simply a reflection of our observation that genes intolerant to loss-of-function mutations are more likely to harbor isoQTLs than eQTLs (fig. S10). Extending these analyses across trimester-specific annotations, although less well powered, pointed to isoform regulation in Tri2 as being particularly important, especially for BIP (paired *t* test  $P = 0.04$  on isoform  $h^2_{\text{med}}/h^2_g$  across all traits; Fig. 4D and table S4). Our findings also remained consistent when restricting to protein-coding genes and their isoforms, indicating that the results were not biased by the shift of eGene biotype between trimesters (fig. S10). Altogether, these results highlight the importance of isoform expression regulation in the develop-

ing brain as a critical mediator of psychiatric GWAS heritability and Tri2 as the time frame during which genetic risks converge in the developing human brain.

### Colocalization and isoform-level TWAs

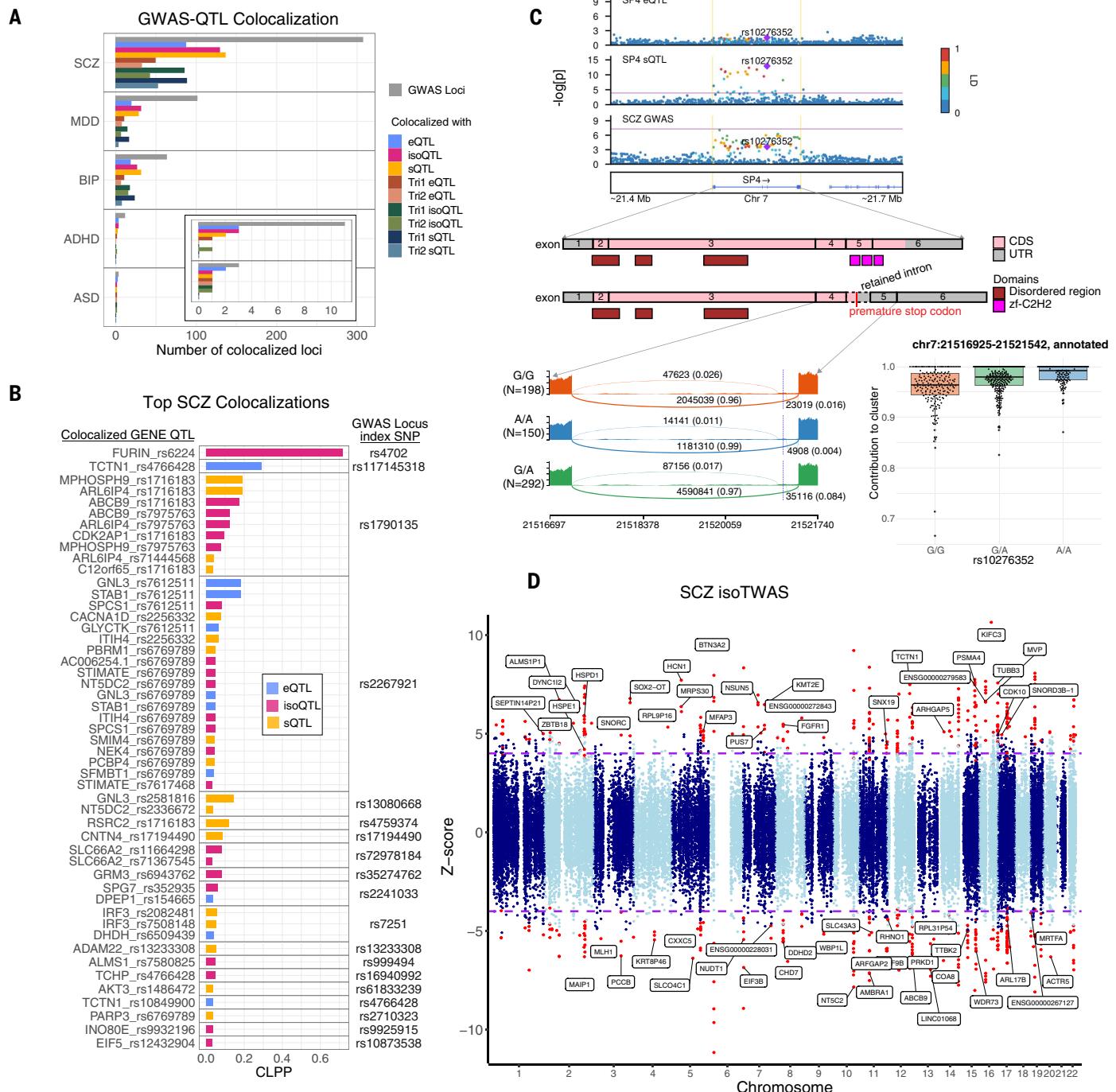
To move from broad GWAS enrichments to locus-specific risk genes and molecular mechanisms, we next conducted a systematic colocalization analysis using two methods eCAVIAR (9), which performs joint probabilistic fine-mapping and estimates the colocalization posterior probability (CLPP) that a given variant is “causal” in both the GWAS and xQTL datasets, and COLOC, which estimates the posterior probability (PP4) that GWAS and xQTL association signals at a given locus are consistent with a shared causal variant (63, 64). Across the combined 485 psychiatric GWAS loci, we prioritized 293 (60%) above the established cutoffs of  $\text{PP4} > 0.7$  or  $\text{CLPP} > 0.01$  [Fig. 5, A and B; fig. S11; and table S5 (35)]. For context, restricting our results to eQTLs in SCZ and BIP, we identified 57 loci with prioritized eQTLs using eCAVIAR, whereas a recent, much larger adult brain eQTL study with an effective sample size of 3154 identified 20 significant colocalizations ( $\text{CLPP} > 0.01$ ) for the two disorders (15).

Consistent with heritability enrichment patterns, we observed many more GWAS loci harboring colocalized isoQTLs and sQTLs compared with standard eQTLs (Fig. 5, A and B; fig. S11; and table S5). To confirm the consistency with previous studies, we identified the well-established neuropsychiatric risk gene *FURIN* with an isoQTL colocalization observed across both SCZ and BIP (Fig. 5B and fig. S11) (10). The colocalized variant rs6224 (located at intron 13) is in moderately strong LD ( $R^2 = 0.7$ ) with rs4702 located at the 3'-untranslated region (3'-UTR) of *FURIN* that has been CRISPR validated as an eQTL (65, 66). Further, we prioritized a candidate common variant mechanism for SCZ at the *SP4* locus, a transcription factor and high-confidence rare variant gene implicated in SCZ risk (1, 67, 68). Although *SP4* was known to be the target gene in this GWAS locus, variant-to-gene mapping was limited by the lack of an observable *SP4* eQTL signal in adult or fetal brain. Here, we identified a cryptic splicing event in *SP4* colocalizing with both SCZ and BIP GWASs (Fig. 5C and fig. S11). The risk variant rs10276352 (G>A) was identified as an sQTL associated with increased inclusion of a 181-base pair cryptic, unannotated exon (chr7:21521120-21521300). The inclusion of this cryptic exon was predicted to introduce a frameshift and premature stop codon between canonical exons 4 and 5, likely resulting in nonsense-mediated decay. The resulting truncated protein, if any, would be missing the zinc finger domain that is critical for *SP4*'s DNA-binding activity (Fig. 5C).

Other candidate disease mechanisms prioritized by these analyses included splicing dysregulation in BIP of *SCN2A* (sQTL: rs17183814, CLPP = 0.82; fig. S11 and table S5), a voltage-gated sodium channel subunit with rare variant associations in ASD and epilepsy (69, 70); downregulation of the metabotropic glutamate receptor *GRM3* in SCZ (PP4 = 0.99; eQTL/isoQTL: rs6943762, CLPP = 0.06/0.08; Fig. 5B and table S5); and isoform-level dysregulation in SCZ of *KMT2E* (PP4 = 0.99; isoQTL: rs2428162, CLPP = 0.02; table S5), a lysine methyltransferase chromatin modifier and high-confidence ASD risk gene. In MDD, several top colocalized genes are associated with clathrin-coated vesicles (*DENND1B*, *SH3GL2*, and *RAB27B*; fig. S11).

We next conducted an isoTWAS (71) to identify genes and isoforms with *cis*-regulated expression associated with SCZ risk. isoTWAS identified 536 isoforms across 271 distinct genes with significant isoTWAS associations [Bonferroni-corrected  $P < 0.05$  across genes, familywise error rate-corrected  $P < 0.05$  across transcripts of the same gene, permutation  $P < 0.05$  (35)]. To leverage local LD and to account for SNP weight correlations, we performed fine-mapping on isoTWAS associations that passed permutation testing (Fig. 5D and table S5). This analysis resulted in 129 putatively causal isoforms from 107 distinct genes that fell into 90% credible sets, with 15 of these genes having a pLI score > 0.9. Of these, 69 isoforms from 57 distinct genes were within 500 kb of independent GWAS-significant loci [41 distinct GWAS loci with lead GWAS SNP  $< 5 \times 10^{-8}$ , defined by LD clumping with GWAS  $P$  used for ranking and a  $R^2$  threshold of 0.2 (72)]. Comparing isoforms prioritized through isoTWAS and/or isoQTL colocalization revealed 14 isoforms that were prioritized with both methods, including *SLC9C2*, *KMT2E*, and *ABCB9*. The relatively low overlap could be caused by the low power of probabilistic colocalization methods (73). With its increased power, isoTWAS also captured additional notable associations, including genetically mediated up-regulation of *HCN1-20I*, the dominant isoform of the hyperpolarization-activated cation channel *HCN1*. An additional advantage of isoTWAS over colocalization analyses was the ability to map isoform-level associations outside of GWAS-significant loci, additionally identifying 60 putatively causal isoforms across 50 distinct genes that were found outside a GWAS-significant locus.

When examining results from gene prioritization analyses, we observed several instances in which a single variant was associated with multiple distinct QTLs and overlapped with disease GWAS signal. For example, rs6769789 was identified as an xQTL for multiple genes on chromosome 3 that colocalized with SCZ GWAS (Fig. 5B). Because of pleiotropy, it can



**Fig. 5. Neuropsychiatric risk gene prioritization through colocalization and isoTWASs.** (A) Total number of GWAS loci exhibiting significant colocalization ( $CLPP > 0.01$  or  $PP4 > 0.7$ ) with specific developing brain xQTL annotations. (B) Colocalization between SCZ GWAS and developing brain xQTLs ranked by CLPP and grouped by GWAS locus, as indicated by the index SNP to the right. Only top colocalization results for protein-coding genes at genome-wide significant loci are shown. (C) Top: locus plots of SCZ GWAS with SP4 eQTLs and sQTLs. A significant colocalization ( $CLPP = 0.01$ ) is observed for SCZ GWAS with a cryptic splicing event in SP4. SP4 does not have a detectable eQTL in the

developing brain. Middle: Gene structure of *SP4* with and without cryptic exon inclusion, likely resulting in nonsense-mediated decay. Bottom left: sashimi plot showing the density of exon and junction read mapping for intron cluster chr7:21516925-21521542 stratified by the colocaled sQTL. Bottom right: sQTL rs10276352 (G>A) increases the contribution to cluster of annotated intron chr:21516925-21521542. The SCZ risk allele increases cryptic exon inclusion. (D) Developing brain isoTWAS associations with SCZ GWASs. Each dot represents an isoform, and adjusted  $P < 0.05$  isoforms are shown in red. Genes of fine-mapped isoforms near a GWAS locus are labeled.

be particularly difficult to distinguish whether one (of many potential) molecular effects is the true mediator of the SNP trait effect. In

attempt to address this, we applied MRLocus (74), which leverages a Mendelian randomization framework for loci with allelic heterozy-

ogeneity using conditionally independent QTLs for these mediators to further support (or refute) observed gene-to-trait effects. Among

14 TWAS-identified genes with more than three conditionally independent eQTLs, the MRLocus framework provided additional support for seven genes with FDR < 0.2 [table S5 (35)]. In the pleiotropic rs6769789 locus described above, for example, there was an association between eQTL and GWAS effect sizes among conditionally independent eQTLs for *NT5DC2* (FDR = 0.125), providing additional evidence in support of this gene-trait association (fig. S12).

### Network-level contextualization of developmental gene regulation

To identify developmentally relevant transcriptional networks, investigate them for common and rare genetic risk, and connect that risk with key biological and cellular processes, we performed robust weighted gene coexpression network analysis (rWGCNA), an unsupervised method that clusters genes into modules based on shared patterns of expression (22, 75–77). We constructed networks using gene- and isoform-level quantifications across all samples ( $N = 642$ ), as well as within samples filtered by trimester (Tri1, Tri2) or chromosomal sex (XX, XY). In total, we identified 124 gene and isoform coexpression modules exhibiting enrichment for all major developmental cell types (39) and recapitulating early biochemical processes and developmental pathways [Fig. 6A (35)]. This expanded on previous work (12, 18, 19, 21, 22, 24) by incorporating an order of magnitude more samples, generating isoform-level modules for >120,000 transcripts, and contextualizing findings within trimester- and sex-specific network contexts.

Overlaying enrichment for common and rare variation within the hierarchical structure defined by coexpression pinpointed groups of modules that harbor most of neurodevelopmental disease risk [Fig. 6B, fig. S13, and table S6 (35)]. Broadly, we found that isoform modules were more likely to capture cell-type marker enrichment (two-sided Fisher's exact test on FDR < 0.1 hits,  $P = 8 \times 10^{-5}$ , OR = 2.37), demonstrating the importance of incorporating splicing and isoform expression in a network context. Neuropsychiatric GWAS signals localized within modules enriched for deep layer excitatory, maturing excitatory, and interneuron cell populations while exhibiting depletion for neural progenitor modules. Rare variation was more likely to be enriched in deep layer excitatory, maturing excitatory, and oligodendrocyte precursor modules, suggesting that most disease-associated variation perturbs the maturation of neuronal cell types (fig. S14).

At the gene level, a group of modules (M1, M2, and M3) enriched for chromatin remodeling and histone modification pathways, with hub genes including *EP300*, *EP400*, *ARID1A*, *KMT2E*, and *POGZ*, converged with maturing excitatory and inhibitory neuron marker genes

and showed strong enrichment for rare variation associated with ASD and developmental delay (DD) (Fig. 6C). These modules highly overlapped with the DD- and ASD-risk yellow module described in Walker *et al.* (24), building upon the previously described excitatory neuron enrichment for these disorders. Our analysis partitioned the hub genes of the yellow module into a new group of deep-excitatory synaptic projection modules (M83, M84, and M85; fig. S15), which exhibited strong enrichment for cross-disorder common variation and rare variation risk for ASD, epilepsy, and SCZ. The strongest genetic risk for SCZ was observed in two groups of interrelated gene modules, M93 and M94 (group 1) and M86, M87, and M88 (group 2) (fig. S15), which are enriched for synaptic gene ontologies and excitatory neuron marker genes. A closely related isoform module, M82, captured this SCZ signal and harbored hub transcripts with known risk, like the canonical transcripts of *TRIO* and *SYNGAP1*, as well as multiple transcripts of *ANK3* (fig. S15). Using module preservation analysis on our trimester-specific networks, we identified M17 as a Tri1-specific module that shows distinct enrichment for severe developmental disorder risk genes and ontologies for neurogenesis and DNA binding, with hub genes such as *GREB1L*, *LRP2*, and *CASZ1* (fig. S15).

At the isoform level, M59 was notable for its association with ADHD common variant risk and enrichment for mitochondrial and bioenergetic pathways. Hub genes for this module included multiple proteasome subunit genes (the PSMD family), and cell-type analysis revealed association with excitatory, microglial, and pericyte populations (Fig. 6D). Furthermore, a group of neuronally enriched modules (M65, M66, and M67; fig. S15) displayed robust BIP, education attainment (EA), and SCZ GWAS enrichments while harboring high-confidence NDD risk genes as hubs including *MEF2C* and *SATB2*, similar to the prenatally enriched M37 module described in Li *et al.* (18). Distinct transcripts of the known ASD risk gene *SOX5* were identified as hubs for two isoform modules, M120 and M122, but M120 alone exhibit a strong and specific enrichment for ASD risk genes (fig. S15).

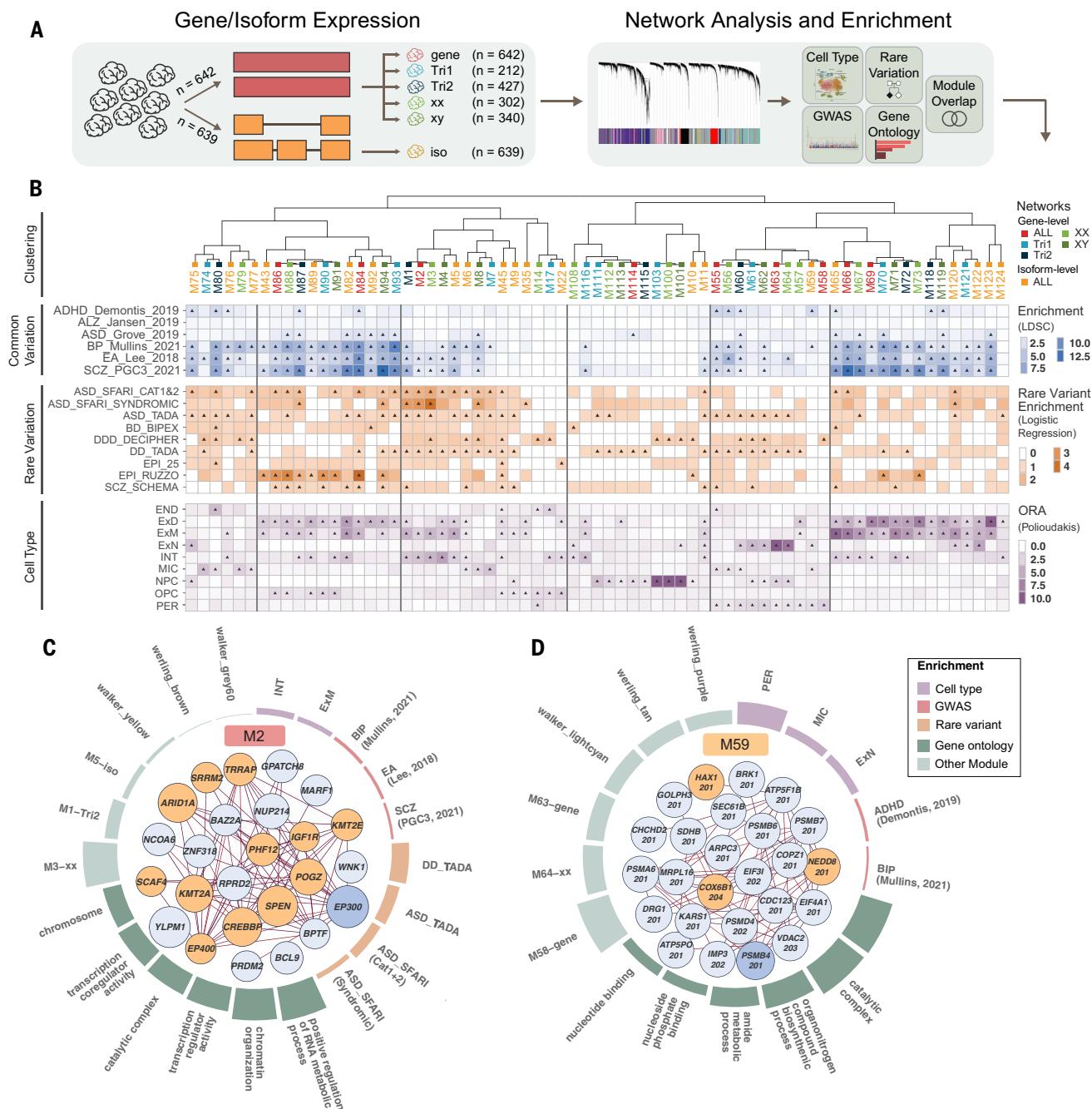
### Cell-type specificity through module-interaction QTLs

Given that gene regulation often occurs within cell type-specific contexts, we leveraged two orthogonal approaches to interrogate the cellular specificity of our xQTL results in the developing human brain. First, we applied the network-based framework implemented in CellWalker to integrate developing brain single-cell chromatin accessibility with eQTL results (78) using the largest existing atlas from midgestation telencephalon (79). Chro-

matin accessibility complements RNA-seq by localizing eQTLs or active regulatory elements. CellWalker mapped 21.7% of eQTL-containing genes to a specific cellular context in the developing human brain, corresponding to 3739 of the bulk-derived eQTLs [Fig. 7A and table S7 (35)]. Second, we leveraged the fact that coexpression modules capture cell-type-specific processes (Fig. 6A) to identify module-interacting eQTL (ieQTLs). To identify such ieQTLs, in which SNP-gene associations are modulated by the levels of a given module, we tested for an interaction effect between genotype and module eigengenes expression [Fig. 7B (35)]. Across all gene and isoform modules, 8008 ieQTLs were identified following permutation testing [FDR-adjusted  $P < 0.05$  (35)], of which 3960 were specific to a single module (Fig. 7C). To validate the cell-type specificity of these module ieQTLs, we examined their overlap with eQTLs identified separately within cultured neurons and progenitors (56). Using the Storey's *p*1 statistic (43), we observed substantial concordance between module-interacting SNP-gene pairs that are true associations in external neuron and progenitor eQTL datasets (Fig. 7C). Forty modules had a *p*1 difference >0.2 between neuron and progenitor eQTLs, suggesting cell type-specific genetic regulation (Fig. 7C). Finally, integrating disease GWAS signals, we identified significant colocalizations (CLPP > 0.01) between ieQTLs and 22 SCZ GWAS loci, 13 of which were not found with bulk *cis*-eQTLs (table S5). For example, we identified a SCZ colocalization with an ieQTL of *BRINP2* and the Tri1 gene module M93 (rs17659437, CLPP = 0.01653) (Fig. 7D). M93 is enriched for deep layer excitatory neuron markers, cross-disorder psychiatric GWAS signal, and synapse-related pathways (Fig. 7E), with eigengene expression increasing across development (Fig. 7F). Correspondingly this module shows greater *p*1 concordance with neuronal compared with progenitor eQTLs (Fig. 7C). *BRINP2* is involved in BMP and retinoic acid signaling pathways, both of which play critical roles in brain development and neuropsychiatric disease. In sum, these analyses demonstrate how bulk QTLs can be annotated to specific cell types, to uncover context-specific gene-regulatory variation relevant to neuropsychiatric disease mechanisms in the developing human brain.

### Discussion

Here, we present a broad view of the landscape of gene, isoform, and splicing regulation in the developing human brain across >650 distinct donors. We further provide xQTL maps specific to the first and second trimesters of human brain development, as well as across three genetic ancestries, leveraging the resulting allelic diversity through cross-ancestry fine-mapping to narrow in on underlying candidate



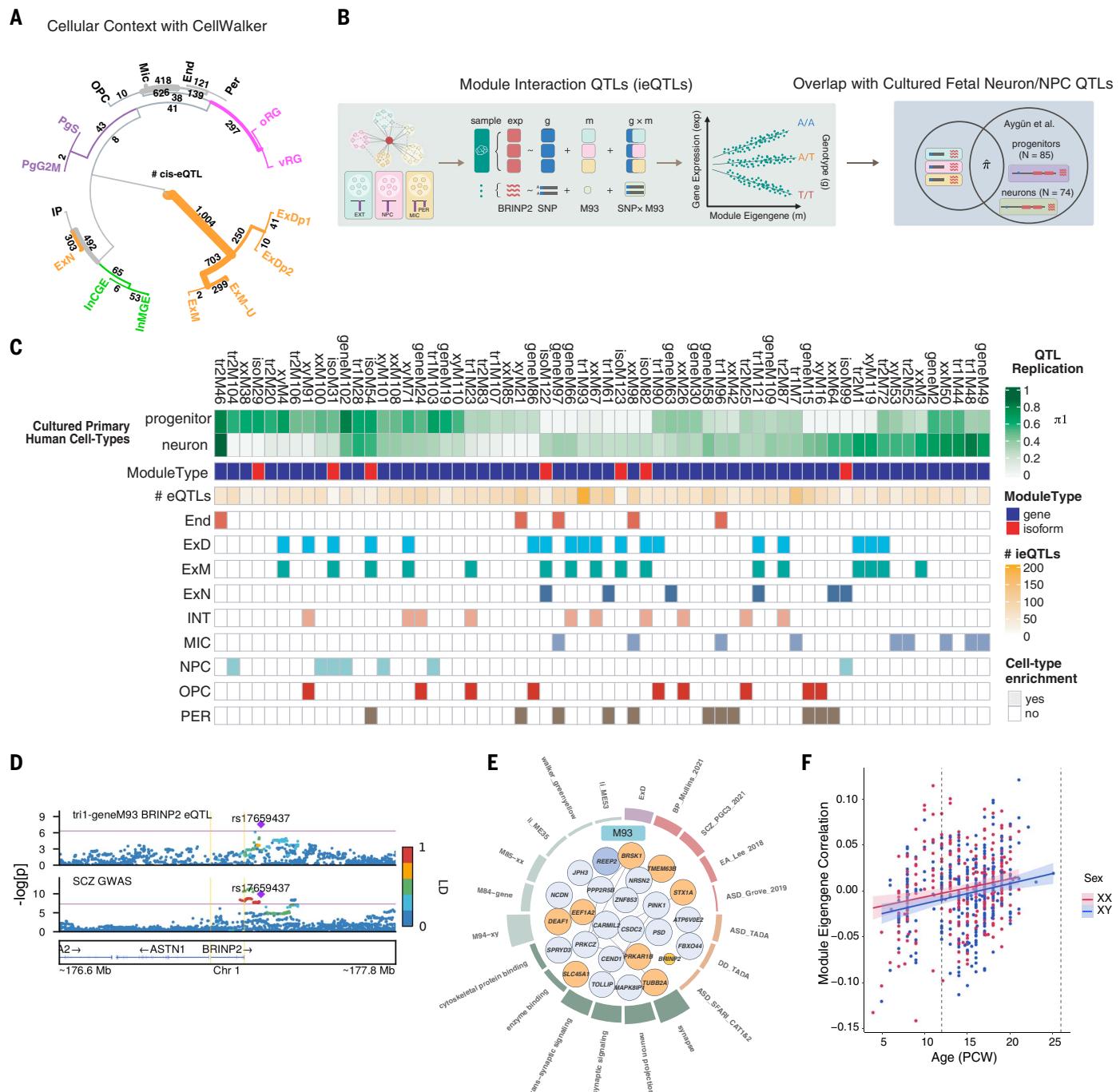
**Fig. 6. Systems-level integration of risk variation with developmental gene and isoform coexpression.** (A) Workflow for construction of gene- and isoform-level coexpression networks, followed by cell-type, pathway, and disease gene enrichment analyses. Separate gene coexpression networks were built to capture trimester- and sex-specific effects. (B) Top, hierarchical clustering of modules from gene, isoform, trimester, and sex-stratified networks through biweight midcorrelation of eigengenes. Middle, heatmaps depicting module-level enrichment for neuropsychiatric GWAS signal ( $-\log_{10}P_{\text{enrich}}$  from S-LDSC and MAGMA) and ORs for rare variation and cell-type enrichment (truncated at 10). Triangles indicate FDR-

corrected  $P < 0.1$  significance. (C) Annotations for M2, a development-wide disease-associated chromatin regulation module. Center: top module ("hub") genes with circle size reflecting module membership (kME) and orange shading indicating genes with associated high-confidence neuropsychiatric disorder-associated rare variants. Thin edges represent topologic overlap, and solid edges indicate protein-protein interactions from the STRING database. Surrounding: circular bar plot highlighting module enrichment for cell types (purple), common (red) and rare (orange) variation, GO terms (dark green), and module overlap (light green). (D) Annotations for M59, an ADHD-associated mitochondrial/proteasome isoform module.

causal genetic associations. With this work, we prioritized twofold more risk genes and candidate molecular mechanisms for neurodevelopmental and psychiatric disorder GWAS loci

compared with much larger adult brain reference panels, highlighting the importance of developmental context when interpreting genetic risk variation. Finally, we construct gene and

isoform-level coexpression networks to place this polygenic risk within an unsupervised, systems-level developmental and cell type-specific context.



**Fig. 7. Module-interacting eQTLs and context-specific GWAS colocalization.**

(A) Hierarchical clustering of *cis*-eQTL enrichments among specific developing brain cell types as mapped by CellWalker. Outermost numbers denote results from single-cell-type label analysis. Inner numbers denote results from a broader, multilevel label analysis. (B) Schematic of module interaction ieQTL mapping and validation in cultured neurons and progenitors. (C) Results from module ieQTL mapping. From top to bottom,  $\pi_1$  statistics depicting ieQTL

overlap with eQTLs from cultured neurons or NPCs, molecular feature (gene or isoform), number of ieQTLs, and cell-type enrichment. A total of 62 modules with  $\pi_1 > 0.2$  in either neurons or NPCs are shown. (D) Colocalization between SCZ GWAS and *BRINP2* ieQTL rs17659437 in M93 ( $CLPP = 0.02$ ). (E) Annotation for M93, a SCZ/BIP enriched deep-layer neuronal, synaptic module. (F) Trajectory of M93 eigengene expression across brain development colored by biological sex.

Our results have several notable implications for future investigations of the functional genomic underpinnings of brain-relevant complex traits and disorders. First, several findings highlight the substantial impact of develop-

mental stage on gene regulation, including a marked drop from 10 to 18 PCW in *cis*-heritability of gene expression and splicing. Because QTL detection and *cis* heritability are interrelated, this explains why we identified a

similar number of eGenes on par with larger adult brain studies. Furthermore, although disease genes tend to be under negative selection, which is inversely related to gene expression heritability (37, 38), our data suggest

that this relationship is likely further moderated by developmental stage. Indeed, the developing brain at Tri2 harbors greater enrichment for neuropsychiatric GWAS signal than at postnatal time points, consistent with results from rare variant enrichment analyses (22). There are several potential biological explanations for the observed temporal drop in heritability. Dynamic changes in neurogenesis and neuronal cell fate specification during this period lead to an increasing diversity of underlying cell types. Therefore, this decline could reflect distinct gene-regulatory mechanisms specific to certain cell types, especially progenitor and excitatory neuronal lineages, as their proportions become more variable. Conversely, we also found evidence that heritability is intrinsically higher in phNPCs compared with their differentiated neuronal counterparts, consistent with a reported larger number of chromatin accessibility peaks in phNPCs (57). Both biological explanations likely contribute, and future population-level developmental brain atlases with single-cell resolution and multi-omic readouts will be needed to further address this question.

Second, we observed prominent differences in the ability of distinct molecular features to capture neuropsychiatric risk mechanisms. Specifically, isoform-level regulation mediated substantially more neuropsychiatric disease heritability than eQTLs across several distinct GWAS. Further, isoQTLs and sQTLs colocalized with many more GWAS loci than standard eQTLs across both developing or adult brain datasets. Given that isoform quantification requires accurate imputation from short-read RNA-seq, guided by genomic reference annotations that are notably incomplete with respect to the human brain (80), we expect these differences to only become further accentuated as our understanding of the landscape of alternative splicing and isoform complexity grows through the continued adoption of emerging long-read RNA-seq technologies.

Third, with the ever-increasing number of available functional genomic reference panels, including from this study, it is becoming increasingly clear that there will soon be multiple potential, prioritized functional genomic mechanisms at a given GWAS locus. This can occur when a GWAS SNP tags a haplotype containing an underlying structural variant (such as an inversion), with multiple pleiotropic effects on *cis* gene regulation. Mendelian randomization-based approaches can help to address this challenge, but will require larger reference panels capable of capturing a greater extent of conditionally independent QTL signals across the transcriptome. Further, despite the promise shown by single-cell QTL studies, they will only exacerbate this issue while lacking the ability to profile splicing and isoform-level regulation, which we have

shown to be the most high yield for mechanistic inference.

Given that isoform-level coexpression networks in particular recapitulate cell type-specific biological processes and genetic regulation, future work should continue to build larger bulk-tissue reference panels of the developing human brain guided by more complete, matching isoform-level annotations from long-read sequencing. Finally, greater visibility is needed, especially in Tri3 of brain development, which is critically underrepresented in our sample.

## Methods summary

Here, we integrated and uniformly processed genotype and RNA-seq data from five developing human brain datasets (12, 24–27), for a total of 654 unique samples passing strict quality control. The genotype data were filtered by MAF, missingness, and Hardy-Weinberg Equilibrium (HWE), followed by imputation into the multiancestry TOPMed reference panel using the Michigan Imputation Server (34, 81). After imputation, SNP variants were combined across cohorts and merged with 1000 Genomes (47) to infer ancestry. RNA-seq reads were aligned to GRCh37 with GENCODE v29lift37 annotations (40, 82) using STAR-2.7.3a (83). Gene and isoform quantifications were calculated using Salmon v1.0.0 (41). QC metrics were calculated using PicardTools. Genes and isoforms with an expression level of >0.1 transcripts per million in >25% of the subjects were included in the analysis. Local splicing was quantified using Leafcutter v0.2.7 (42). Introns and intron clusters that passed extensive quality controls, including >50 reads per cluster, >500 kb intron length, >0.001 of reads in a cluster support an intron, and >5 reads per intron, were included in the analysis.

*cis*-eQTL, *cis*-isoQTL, and *cis*-sQTL mapping was conducted using FastQTL (84) for variants within a *cis* window of 1 Mb. QTL mapping covariates included sex, age, top five genotype PCs, and Hidden Covariates with Prior (HCP) factors. Genes with adjusted beta-approximated permutation  $P < 0.05$  were defined as eGenes, isoGenes, and sGenes. Population-specific xQTLs were mapped in EUR, AMR, and AFR samples. Trimester-specific xQTLs were mapped in EUR Tri1 and Tri2 samples. xQTLs were fine-mapped using SuSIE (45), and cross-ancestry eQTLs were further fine-mapped using PAINTOR (51, 52) while leveraging functional annotations. Gene expression and splicing heritability were estimated using a restricted maximum likelihood algorithm implemented in Julia (85). Cell-type proportions were estimated for seven major cell classes with CIBERSORTx (53) using a single-cell reference panel of meta-analyzed cell-type markers (“metamarkers”) from 37 primary fetal brain scRNA-seq datasets

comprising ~2.95 million individual nuclei and cells (54).

Stratified LD-score regression (6) was applied to investigate the enrichment of neuropsychiatric GWAS heritability among developing brain QTLs, which were quantitatively annotated on the basis of their probability of being causal using maxCPP (59). The MESC framework (38) was used to estimate the proportion of GWAS heritability mediated by the *cis* genetic component of genes, isoforms, and introns, which were further stratified into trimesters. The isoTWAS (71) framework was applied to identify isoforms with *cis*-regulated expression associated with neuropsychiatric GWASs. Statistical colocalization between xQTLs and GWASs was performed using eCAVIAR v2.2 (9) and COLOC, and variants with CLPP > 0.01 or molecular features with PP4 > 0.7 were defined as colocalized.

Robust weighted gene correlation network analysis (rWGCNA) was performed to identify coexpression modules using gene and isoform quantifications, as well as trimester- and sex-specific contexts. ieQTLs were identified by testing the interaction effect between genotype and module eigengenes. The enrichment of ieQTLs in neurons and progenitors (56) was examined using Storey’s *p*il statistic (43). CellWalker (78) leveraged single-cell RNA-seq and ATAC-seq data to map bulk eQTLs to specific cell types.

See the supplementary materials for the full materials and methods (35).

## REFERENCES AND NOTES

1. V. Trubetskoy *et al.*, Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature* **604**, 502–508 (2022). doi: [10.1038/s41586-022-04434-5](https://doi.org/10.1038/s41586-022-04434-5); pmid: [35396580](https://pubmed.ncbi.nlm.nih.gov/35396580/)
2. J. Grove *et al.*, Identification of common genetic risk variants for autism spectrum disorder. *Nat. Genet.* **51**, 431–444 (2019). doi: [10.1038/s41588-019-0344-8](https://doi.org/10.1038/s41588-019-0344-8); pmid: [30804558](https://pubmed.ncbi.nlm.nih.gov/30804558/)
3. M. J. Gandal, V. Leppa, H. Won, N. N. Parikhshak, D. H. Geschwind, The road to precision psychiatry: Translating genetics into disease mechanisms. *Nat. Neurosci.* **19**, 1397–1407 (2016). doi: [10.1038/nn.4409](https://doi.org/10.1038/nn.4409); pmid: [27786179](https://pubmed.ncbi.nlm.nih.gov/27786179/)
4. L. D. Ward, M. Kellis, Evidence of abundant purifying selection in humans for recently acquired regulatory functions. *Science* **337**, 1675–1678 (2012). doi: [10.1126/science.1225057](https://doi.org/10.1126/science.1225057); pmid: [22956687](https://pubmed.ncbi.nlm.nih.gov/22956687/)
5. M. T. Maurano *et al.*, Systematic localization of common disease-associated variation in regulatory DNA. *Science* **337**, 1190–1195 (2012). doi: [10.1126/science.1222794](https://doi.org/10.1126/science.1222794); pmid: [22955828](https://pubmed.ncbi.nlm.nih.gov/22955828/)
6. H. K. Finucane *et al.*, Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015). doi: [10.1038/ng.3404](https://doi.org/10.1038/ng.3404); pmid: [26414678](https://pubmed.ncbi.nlm.nih.gov/26414678/)
7. L. M. Hernandez *et al.*, Transcriptomic Insight Into the Polygenic Mechanisms Underlying Psychiatric Disorders. *Biol. Psychiatry* **89**, 54–64 (2021). doi: [10.1016/j.biopsych.2020.06.005](https://doi.org/10.1016/j.biopsych.2020.06.005); pmid: [32792264](https://pubmed.ncbi.nlm.nih.gov/32792264/)
8. A. Gusen *et al.*, Integrative approaches for large-scale transcriptome-wide association studies. *Nat. Genet.* **48**, 245–252 (2016). doi: [10.1038/ng.3506](https://doi.org/10.1038/ng.3506); pmid: [26854917](https://pubmed.ncbi.nlm.nih.gov/26854917/)
9. F. Hormozdiari *et al.*, Colocalization of GWAS and eQTL Signals Detects Target Genes. *Am. J. Hum. Genet.* **99**, 1245–1260 (2016). doi: [10.1016/j.ajhg.2016.10.003](https://doi.org/10.1016/j.ajhg.2016.10.003); pmid: [27866706](https://pubmed.ncbi.nlm.nih.gov/27866706/)

10. M. Fromer *et al.*, Gene expression elucidates functional impact of polygenic risk for schizophrenia. *Nat. Neurosci.* **19**, 1442–1453 (2016). doi: [10.1038/nn.4399](https://doi.org/10.1038/nn.4399); pmid: [27668389](https://pubmed.ncbi.nlm.nih.gov/27668389/)
11. A. E. Jaffe *et al.*, Developmental and genetic regulation of the human cortex transcriptome illuminate schizophrenia pathogenesis. *Nat. Neurosci.* **21**, 1117–1125 (2018). doi: [10.1038/s41593-018-0197-y](https://doi.org/10.1038/s41593-018-0197-y); pmid: [30050107](https://pubmed.ncbi.nlm.nih.gov/30050107/)
12. D. M. Werling *et al.*, Whole-genome and RNA sequencing reveal variation and transcriptomic coordination in the developing human prefrontal cortex. *Cell Rep.* **31**, 107489 (2020). doi: [10.1016/j.celrep.2020.03.053](https://doi.org/10.1016/j.celrep.2020.03.053); pmid: [32268104](https://pubmed.ncbi.nlm.nih.gov/32268104/)
13. D. Wang *et al.*, Comprehensive functional genomic resource and integrative model for the human brain. *Science* **362**, eaat8464 (2018). doi: [10.1126/science.aat8464](https://doi.org/10.1126/science.aat8464); pmid: [30545857](https://pubmed.ncbi.nlm.nih.gov/30545857/)
14. M. J. Gandal *et al.*, Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. *Science* **362**, eaat8127 (2018). doi: [10.1126/science.aat8127](https://doi.org/10.1126/science.aat8127); pmid: [30545856](https://pubmed.ncbi.nlm.nih.gov/30545856/)
15. B. Zeng *et al.*, Multi-ancestry eQTL meta-analysis of human brain identifies candidate causal variants for brain-related traits. *Nat. Genet.* **54**, 161–169 (2022). doi: [10.1038/s41588-021-00987-9](https://doi.org/10.1038/s41588-021-00987-9); pmid: [35088635](https://pubmed.ncbi.nlm.nih.gov/35088635/)
16. F. Aguet *et al.*, The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318–1330 (2020). doi: [10.1126/science.aaz1776](https://doi.org/10.1126/science.aaz1776); pmid: [32913098](https://pubmed.ncbi.nlm.nih.gov/32913098/)
17. B. D. Umans, A. Battle, Y. Gilad, Where Are the Disease-Associated eQTLs? *Trends Genet.* **37**, 109–124 (2021). doi: [10.1016/j.tig.2020.08.009](https://doi.org/10.1016/j.tig.2020.08.009); pmid: [32912663](https://pubmed.ncbi.nlm.nih.gov/32912663/)
18. M. Li *et al.*, Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science* **362**, eaat7615 (2018). doi: [10.1126/science.aat7615](https://doi.org/10.1126/science.aat7615); pmid: [30545854](https://pubmed.ncbi.nlm.nih.gov/30545854/)
19. H. J. Kang *et al.*, Spatio-temporal transcriptome of the human brain. *Nature* **478**, 483–489 (2011). doi: [10.1038/nature10523](https://doi.org/10.1038/nature10523); pmid: [22031440](https://pubmed.ncbi.nlm.nih.gov/22031440/)
20. C. Colantuoni *et al.*, Temporal dynamics and genetic control of transcription in the human prefrontal cortex. *Nature* **478**, 519–523 (2011). doi: [10.1038/nature10524](https://doi.org/10.1038/nature10524); pmid: [22031444](https://pubmed.ncbi.nlm.nih.gov/22031444/)
21. A. J. Willsey *et al.*, Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. *Cell* **155**, 997–1007 (2013). doi: [10.1016/j.cell.2013.10.020](https://doi.org/10.1016/j.cell.2013.10.020); pmid: [24267886](https://pubmed.ncbi.nlm.nih.gov/24267886/)
22. N. N. Parikshak *et al.*, Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. *Cell* **155**, 1008–1021 (2013). doi: [10.1016/j.cell.2013.10.031](https://doi.org/10.1016/j.cell.2013.10.031); pmid: [24267887](https://pubmed.ncbi.nlm.nih.gov/24267887/)
23. S. Gulsuner *et al.*, Spatial and temporal mapping of de novo mutations in schizophrenia to a fetal prefrontal cortical network. *Cell* **154**, 518–529 (2013). doi: [10.1016/j.cell.2013.06.049](https://doi.org/10.1016/j.cell.2013.06.049); pmid: [23911319](https://pubmed.ncbi.nlm.nih.gov/23911319/)
24. R. L. Walker *et al.*, Genetic control of expression and splicing in developing human brain informs disease mechanisms. *Cell* **179**, 750–771.e22 (2019). doi: [10.1016/j.cell.2019.09.021](https://doi.org/10.1016/j.cell.2019.09.021); pmid: [31626773](https://pubmed.ncbi.nlm.nih.gov/31626773/)
25. H. E. O’Brien *et al.*, Expression quantitative trait loci in the developing human brain and their enrichment in neuropsychiatric disorders. *Genome Biol.* **19**, 194 (2018). doi: [10.1186/s13098-018-1567-1](https://doi.org/10.1186/s13098-018-1567-1); pmid: [30419947](https://pubmed.ncbi.nlm.nih.gov/30419947/)
26. A. E. Jaffe *et al.*, Developmental regulation of human cortex transcription and its clinical relevance at single base resolution. *Nat. Neurosci.* **18**, 154–161 (2015). doi: [10.1038/nn.3898](https://doi.org/10.1038/nn.3898); pmid: [25501035](https://pubmed.ncbi.nlm.nih.gov/25501035/)
27. S. J. Lindsay *et al.*, HDBR expression: A unique resource for global and individual gene expression studies during early human brain development. *Front. Neuroanat.* **10**, 86 (2016). doi: [10.3389/fnana.2016.00086](https://doi.org/10.3389/fnana.2016.00086); pmid: [27833533](https://pubmed.ncbi.nlm.nih.gov/27833533/)
28. R. Birnbaum, A. E. Jaffe, T. M. Hyde, J. E. Kleinman, D. R. Weinberger, Prenatal expression patterns of genes associated with neuropsychiatric disorders. *Am. J. Psychiatry* **171**, 758–767 (2014). doi: [10.1176/appi.ajp.2014.13111452](https://doi.org/10.1176/appi.ajp.2014.13111452); pmid: [24874100](https://pubmed.ncbi.nlm.nih.gov/24874100/)
29. Y. I. Li *et al.*, RNA splicing is a primary link between genetic variation and disease. *Science* **352**, 600–604 (2016). doi: [10.1126/science.aad9417](https://doi.org/10.1126/science.aad9417); pmid: [27126046](https://pubmed.ncbi.nlm.nih.gov/27126046/)
30. B. Raj, B. J. Blencowe, Alternative splicing in the mammalian nervous system: Recent insights into mechanisms and functional roles. *Neuron* **87**, 14–27 (2015). doi: [10.1016/j.neuron.2015.05.004](https://doi.org/10.1016/j.neuron.2015.05.004); pmid: [26139367](https://pubmed.ncbi.nlm.nih.gov/26139367/)
31. X. Zhang *et al.*, Cell-type-specific alternative splicing governs cell fate in the developing cerebral cortex. *Cell* **166**, 1147–1162.e15 (2016). doi: [10.1016/j.cell.2016.07.025](https://doi.org/10.1016/j.cell.2016.07.025); pmid: [27565344](https://pubmed.ncbi.nlm.nih.gov/27565344/)
32. S. M. Weyn-Vanhentenryck *et al.*, Precise temporal regulation of alternative splicing during neural development. *Nat. Commun.* **9**, 2189 (2018). doi: [10.1038/s41467-018-04559-0](https://doi.org/10.1038/s41467-018-04559-0); pmid: [29875359](https://pubmed.ncbi.nlm.nih.gov/29875359/)
33. D. Garrido-Martín, B. Borsari, M. Calvo, F. Reverter, R. Guijó, Identification and analysis of splicing quantitative trait loci across multiple tissues in the human genome. *Nat. Commun.* **12**, 727 (2021). doi: [10.1038/s41467-020-20578-2](https://doi.org/10.1038/s41467-020-20578-2); pmid: [33526779](https://pubmed.ncbi.nlm.nih.gov/33526779/)
34. D. Taliun *et al.*, Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. *Nature* **590**, 290–299 (2021). doi: [10.1038/s41586-021-03205-y](https://doi.org/10.1038/s41586-021-03205-y); pmid: [33568819](https://pubmed.ncbi.nlm.nih.gov/33568819/)
35. The full materials and methods are available as supplementary materials.
36. M. Lek *et al.*, Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016). doi: [10.1038/nature19057](https://doi.org/10.1038/nature19057); pmid: [27535533](https://pubmed.ncbi.nlm.nih.gov/27535533/)
37. H. Mostafavi, J. P. Spence, S. Naqvi, J. K. Pritchard, Limited overlap of eQTLs and GWAS hits due to systematic differences in discovery. *bioRxiv* (2022), p. 2022.05.07.491045. doi: [10.1101/2022.05.07.491045](https://doi.org/10.1101/2022.05.07.491045)
38. D. W. Yao, L. J. O’Connor, A. L. Price, A. Gushev, Quantifying genetic effects on disease mediated by assayed gene expression levels. *Nat. Genet.* **52**, 626–633 (2020). doi: [10.1038/s41588-020-0625-2](https://doi.org/10.1038/s41588-020-0625-2); pmid: [32424349](https://pubmed.ncbi.nlm.nih.gov/32424349/)
39. D. Polioudakis *et al.*, A single-cell transcriptomic atlas of human neocortical development during mid-gestation. *Neuron* **103**, 785–801.e8 (2019). doi: [10.1016/j.neuron.2019.06.011](https://doi.org/10.1016/j.neuron.2019.06.011); pmid: [31303374](https://pubmed.ncbi.nlm.nih.gov/31303374/)
40. A. Frankish *et al.*, GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res.* **47** (D1), D766–D773 (2019). doi: [10.1093/nar/gky955](https://doi.org/10.1093/nar/gky955); pmid: [30357393](https://pubmed.ncbi.nlm.nih.gov/30357393/)
41. R. Patro, G. Duggal, M. I. Love, R. A. Irizarry, C. Kingsford, Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017). doi: [10.1038/nmeth.4197](https://doi.org/10.1038/nmeth.4197); pmid: [28263959](https://pubmed.ncbi.nlm.nih.gov/28263959/)
42. Y. I. Li *et al.*, Annotation-free quantification of RNA splicing using LeafCutter. *Nat. Genet.* **50**, 151–158 (2018). doi: [10.1038/s41588-017-0004-9](https://doi.org/10.1038/s41588-017-0004-9); pmid: [29229983](https://pubmed.ncbi.nlm.nih.gov/29229983/)
43. J. D. Storey, R. Tibshirani, Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 9440–9445 (2003). doi: [10.1073/pnas.1530509100](https://doi.org/10.1073/pnas.1530509100); pmid: [12883005](https://pubmed.ncbi.nlm.nih.gov/12883005/)
44. O. Delaneau *et al.*, A complete tool set for molecular QTL discovery and analysis. *Nat. Commun.* **8**, 15452 (2017). doi: [10.1038/ncomms15452](https://doi.org/10.1038/ncomms15452); pmid: [28516912](https://pubmed.ncbi.nlm.nih.gov/28516912/)
45. G. Wang, A. Sarkar, P. Carbonetto, M. Stephens, A simple new approach to variable selection in regression, with application to genetic fine mapping. *J. R. Stat. Soc. Series B Stat. Methodol.* **82**, 1273–1300 (2020). doi: [10.1111/rssb.12388](https://doi.org/10.1111/rssb.12388); pmid: [37220626](https://pubmed.ncbi.nlm.nih.gov/37220626/)
46. M. A. Iakram *et al.*, Common variants at 6q22 and 17q21 are associated with intracranial volume. *Nat. Genet.* **44**, 539–544 (2012). doi: [10.1038/ng.2504](https://doi.org/10.1038/ng.2504)
47. A. Auton *et al.*, A global reference for human genetic variation. *Nature* **526**, 68–74 (2015). doi: [10.1038/nature15393](https://doi.org/10.1038/nature15393); pmid: [26432245](https://pubmed.ncbi.nlm.nih.gov/26432245/)
48. N. Zaitlin, B. Pașaniciuc, T. Gur, E. Ziv, E. Halperin, Leveraging genetic variability across populations for the identification of causal variants. *Am. J. Hum. Genet.* **86**, 23–33 (2010). doi: [10.1016/j.ajhg.2009.11.016](https://doi.org/10.1016/j.ajhg.2009.11.016); pmid: [20085711](https://pubmed.ncbi.nlm.nih.gov/20085711/)
49. A. Mahajan *et al.*, Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nat. Genet.* **46**, 234–244 (2014). doi: [10.1038/ng.2897](https://doi.org/10.1038/ng.2897); pmid: [24509480](https://pubmed.ncbi.nlm.nih.gov/24509480/)
50. P. Mohammadi, S. E. Castel, A. B. Brown, T. Lappalainen, Quantifying the regulatory effect size of *cis*-acting genetic variation using allelic fold change. *Genome Res.* **27**, 1872–1884 (2017). doi: [10.1101/gr.216747.116](https://doi.org/10.1101/gr.216747.116); pmid: [29021289](https://pubmed.ncbi.nlm.nih.gov/29021289/)
51. G. Kichaev *et al.*, Integrating functional data to prioritize causal variants in statistical fine-mapping studies. *PLOS Genet.* **10**, e1004722 (2014). doi: [10.1371/journal.pgen.1004722](https://doi.org/10.1371/journal.pgen.1004722); pmid: [25357204](https://pubmed.ncbi.nlm.nih.gov/25357204/)
52. G. Kichaev, B. Pasaniuc, Leveraging functional-annotation data in trans-ethnic fine-mapping studies. *Am. J. Hum. Genet.* **97**, 260–271 (2015). doi: [10.1016/j.ajhg.2015.06.007](https://doi.org/10.1016/j.ajhg.2015.06.007); pmid: [26189819](https://pubmed.ncbi.nlm.nih.gov/26189819/)
53. A. M. Newman *et al.*, Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat. Biotechnol.* **37**, 773–782 (2019). doi: [10.1038/s41587-019-0114-2](https://doi.org/10.1038/s41587-019-0114-2); pmid: [31061481](https://pubmed.ncbi.nlm.nih.gov/31061481/)
54. J. Werner, J. Gillis, Preservation of co-expression defines the primary tissue fidelity of human neural organoids. *bioRxiv* (2023), p. 2023.03.31.535112. doi: [10.1101/2023.03.31.535112](https://doi.org/10.1101/2023.03.31.535112)
55. S. Kim-Hellmuth *et al.*, Cell type-specific genetic regulation of gene expression across human tissues. *Science* **369**, eaaz8528 (2020). doi: [10.1126/science.aaz8528](https://doi.org/10.1126/science.aaz8528); pmid: [32913075](https://pubmed.ncbi.nlm.nih.gov/32913075/)
56. N. Aygün *et al.*, Brain-trait-associated variants impact cell-type-specific gene regulation during neurogenesis. *Am. J. Hum. Genet.* **108**, 1647–1668 (2021). doi: [10.1016/j.ajhg.2021.07.011](https://doi.org/10.1016/j.ajhg.2021.07.011); pmid: [34416157](https://pubmed.ncbi.nlm.nih.gov/34416157/)
57. D. Liang *et al.*, Cell-type-specific effects of genetic variation on chromatin accessibility during human neuronal differentiation. *Nat. Neurosci.* **24**, 941–953 (2021). doi: [10.1038/s41593-021-00858-w](https://doi.org/10.1038/s41593-021-00858-w); pmid: [34017130](https://pubmed.ncbi.nlm.nih.gov/34017130/)
58. B. K. Bulik-Sullivan *et al.*, LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015). doi: [10.1038/ng.3211](https://doi.org/10.1038/ng.3211); pmid: [25642630](https://pubmed.ncbi.nlm.nih.gov/25642630/)
59. F. Hormozdiari *et al.*, Leveraging molecular quantitative trait loci to understand the genetic architecture of diseases and complex traits. *Nat. Genet.* **50**, 1041–1047 (2018). doi: [10.1038/s41588-018-0148-2](https://doi.org/10.1038/s41588-018-0148-2); pmid: [29942083](https://pubmed.ncbi.nlm.nih.gov/29942083/)
60. D. Demontis *et al.*, Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat. Genet.* **51**, 63–75 (2019). doi: [10.1038/s41588-018-0269-7](https://doi.org/10.1038/s41588-018-0269-7); pmid: [30478444](https://pubmed.ncbi.nlm.nih.gov/30478444/)
61. D. M. Howard *et al.*, Genome-wide association study of depression phenotypes in UK Biobank identifies variants in excitatory synaptic pathways. *Nat. Commun.* **9**, 1470 (2018). doi: [10.1038/s41467-018-03819-3](https://doi.org/10.1038/s41467-018-03819-3); pmid: [29662059](https://pubmed.ncbi.nlm.nih.gov/29662059/)
62. N. Mullins *et al.*, Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. *Nat. Genet.* **53**, 817–829 (2021). doi: [10.1038/s41588-021-00857-4](https://doi.org/10.1038/s41588-021-00857-4); pmid: [34002096](https://pubmed.ncbi.nlm.nih.gov/34002096/)
63. C. Giambartolomei *et al.*, Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLOS Genet.* **10**, e1004383 (2014). doi: [10.1371/journal.pgen.1004383](https://doi.org/10.1371/journal.pgen.1004383); pmid: [24830394](https://pubmed.ncbi.nlm.nih.gov/24830394/)
64. D. J. Schaid, W. Chen, N. B. Larson, From genome-wide associations to candidate causal variants by statistical fine-mapping. *Nat. Rev. Genet.* **19**, 491–504 (2018). doi: [10.1038/s41576-018-0016-z](https://doi.org/10.1038/s41576-018-0016-z); pmid: [29844615](https://pubmed.ncbi.nlm.nih.gov/29844615/)
65. Y. Hou *et al.*, Schizophrenia-associated rs4702 G allele-specific downregulation of FURIN expression by miR-338-3p reduces BDNF production. *Schizophr. Res.* **199**, 176–180 (2018). doi: [10.1016/j.schres.2018.02.040](https://doi.org/10.1016/j.schres.2018.02.040); pmid: [29499696](https://pubmed.ncbi.nlm.nih.gov/29499696/)
66. N. Schrode *et al.*, Synergistic effects of common schizophrenia risk variants. *Nat. Genet.* **51**, 1475–1485 (2019). doi: [10.1038/s41588-019-0497-5](https://doi.org/10.1038/s41588-019-0497-5); pmid: [3148722](https://pubmed.ncbi.nlm.nih.gov/3148722/)
67. J. M. Fu *et al.*, Rare coding variation provides insight into the genetic architecture and phenotypic context of autism. *Nat. Genet.* **54**, 1320–1331 (2022). doi: [10.1038/s41588-022-01104-0](https://doi.org/10.1038/s41588-022-01104-0); pmid: [35982160](https://pubmed.ncbi.nlm.nih.gov/35982160/)
68. T. Singh *et al.*, Rare coding variants in ten genes confer substantial risk for schizophrenia. *Nature* **604**, 509–516 (2022). doi: [10.1038/s41586-022-04556-w](https://doi.org/10.1038/s41586-022-04556-w); pmid: [35396579](https://pubmed.ncbi.nlm.nih.gov/35396579/)
69. M. Wolff *et al.*, Genetic and phenotypic heterogeneity suggest therapeutic implications in SCN2A-related disorders. *Brain* **140**, 1316–1336 (2017). doi: [10.1093/brain/aww054](https://doi.org/10.1093/brain/aww054); pmid: [28379373](https://pubmed.ncbi.nlm.nih.gov/28379373/)
70. P. P. Zandi *et al.*, Amygdala and anterior cingulate transcriptomes from individuals with bipolar disorder reveal downregulated neuroimmune and synaptic pathways. *Nat. Neurosci.* **25**, 381–389 (2022). doi: [10.1038/s41593-022-01024-6](https://doi.org/10.1038/s41593-022-01024-6); pmid: [35260864](https://pubmed.ncbi.nlm.nih.gov/35260864/)
71. A. Bhattacharya *et al.*, Isoform-level transcriptome-wide association uncovers extensive novel genetic risk mechanisms for neuropsychiatric disorders in the human brain. *Nat. Genet.* **55**, 2117–2128 (2023). doi: [10.1038/s41588-023-01560-2](https://doi.org/10.1038/s41588-023-01560-2); pmid: [38036788](https://pubmed.ncbi.nlm.nih.gov/38036788/)
72. F. Privé, H. Aschard, A. Ziyatdinov, M. G. B. Blum, Efficient analysis of large-scale genome-wide data with two R packages: Bigstatsr and bigsnpr. *Bioinformatics* **34**, 2781–2787 (2018). doi: [10.1093/bioinformatics/bty185](https://doi.org/10.1093/bioinformatics/bty185); pmid: [29617937](https://pubmed.ncbi.nlm.nih.gov/29617937/)
73. A. Hukku *et al.*, Probabilistic colocalization of genetic variants from complex and molecular traits: Promise and limitations. *Am. J. Hum. Genet.* **108**, 25–35 (2021). doi: [10.1016/j.ajhg.2020.11.012](https://doi.org/10.1016/j.ajhg.2020.11.012); pmid: [33084433](https://pubmed.ncbi.nlm.nih.gov/33084433/)
74. A. Zhu *et al.*, MRLocus: Identifying causal genes mediating a trait through Bayesian estimation of allelic heterogeneity. *PLOS Genet.* **17**, e1009455 (2021). doi: [10.1371/journal.pgen.1009455](https://doi.org/10.1371/journal.pgen.1009455); pmid: [33872308](https://pubmed.ncbi.nlm.nih.gov/33872308/)
75. P. Langfelder, S. Horvath, WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008). doi: [10.1186/1471-2105-9-559](https://doi.org/10.1186/1471-2105-9-559); pmid: [19114008](https://pubmed.ncbi.nlm.nih.gov/19114008/)

76. M. J. Gandal *et al.*, Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. *Science* **359**, 693–697 (2018). doi: [10.1126/science.aad6469](https://doi.org/10.1126/science.aad6469); pmid: [29439242](#)
77. N. N. Parikhshak, M. J. Gandal, D. H. Geschwind, Systems biology and gene networks in neurodevelopmental and neurodegenerative disorders. *Nat. Rev. Genet.* **16**, 441–458 (2015). doi: [10.1038/nrg3934](https://doi.org/10.1038/nrg3934); pmid: [26149713](#)
78. P. F. Przytycki, K. S. Pollard, CellWalker integrates single-cell and bulk data to resolve regulatory elements across cell types in complex tissues. *Genome Biol.* **22**, 61 (2021). doi: [10.1186/s13059-021-02279-1](https://doi.org/10.1186/s13059-021-02279-1); pmid: [33583425](#)
79. R. S. Ziffrin *et al.*, Single-cell epigenomics reveals mechanisms of human cortical development. *Nature* **598**, 205–213 (2021). doi: [10.1038/s41586-021-03209-8](https://doi.org/10.1038/s41586-021-03209-8); pmid: [34616060](#)
80. D. Zhang *et al.*, Incomplete annotation has a disproportionate impact on our understanding of Mendelian and complex neurogenetic disorders. *Sci. Adv.* **6**, eaay8299 (2020). doi: [10.1126/sciadv.aay8299](https://doi.org/10.1126/sciadv.aay8299); pmid: [32917675](#)
81. S. Das *et al.*, Next-generation genotype imputation service and methods. *Nat. Genet.* **48**, 1284–1287 (2016). doi: [10.1038/ng.3656](https://doi.org/10.1038/ng.3656); pmid: [27571263](#)
82. A. Frankish *et al.*, GENCODE 2021. *Nucleic Acids Res.* **49** (D1), D916–D923 (2021). doi: [10.1093/nar/gkaa1087](https://doi.org/10.1093/nar/gkaa1087); pmid: [33270111](#)
83. A. Dobin *et al.*, STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013). doi: [10.1093/bioinformatics/bts635](https://doi.org/10.1093/bioinformatics/bts635); pmid: [23104886](#)
84. H. Ongen, A. Buil, A. A. Brown, E. T. Dermitzakis, O. Delaneau, Fast and efficient QTL mapper for thousands of molecular phenotypes. *Bioinformatics* **32**, 1479–1485 (2016). doi: [10.1093/bioinformatics/btv722](https://doi.org/10.1093/bioinformatics/btv722); pmid: [26708335](#)
85. M. Kim *et al.*, Multivariate variance components analysis uncovers genetic architecture of brain isoform expression and novel psychiatric disease mechanisms. medRxiv 2022.10.18.22281204 [Preprint] (2022); <https://doi.org/10.1101/2022.10.18.22281204>.
86. C. Wen, M. Margolis, M. Gandal, “gandallab/devBrain\_xQTL,” Zenodo (2023); <http://dx.doi.org/10.5281/ZENODO.8336503>.

## ACKNOWLEDGMENTS

**Funding:** This work was supported by the Simons Foundation (SFARI Bridge to Independence award to M.J.G.) and the National Institute of Mental Health (NIMH) of the National Institutes of Health (NIH) (grants R01-MH121521 and R01-MH123922 to M.J.G.). Data were generated as part of the PsychENCODE Consortium, which is supported by the following NIMH grants: U01DA048279, U01MH103339, U01MH103340, U01MH103346, U01MH103365, U01MH103392, U01MH116438, U01MH116441, U01MH116442, U01MH116448, U01MH116489, U01MH116492, U01MH122590, U01MH122591, U01MH122592, U01MH122849, U01MH122678, U01MH122681, U01MH116487, U01MH122509, R01MH094714, R01MH105472, R01MH105898, R01MH109677, R01MH109715, R01MH110905, R01MH110920, R01MH110921, R01MH110926, R01MH110927, R01MH110928, R01MH111721, R01MH117291, R01MH117292, R01MH117293, R21MH102791, R21MH103877, R21MH105853, R21MH105881, R21MH109956, R56MH14899, R56MH14901, R56MH14911, R01MH125516, R01MH126459, R01MH129301, R01MH126393, R01MH121521, R01MH116529, R01MH129817, R01MH117406, and P50MH106934 awarded to: A. Abzyov, N. Ahituv, S. Akbarian, K. Brennand, A. Chess, G. Cooper, G. Crawford, S. Dracheva, P. Farnham, M.J.G., M.G., D.H.G., F. Goes, J. F. Hallmayer, V. Haroutunian, T.M.H., A.E.J., P. Jin, M. Kellis, J.E.K., J. A. Knowles, A. Kriegstein, C.L., C. E. Mason, K. Martinowich, E. Mukamel, R. Myers, C. Nemerooff, M.A.P., D. Pinto, K. S. Pollard, K. Ressler, P. Roussos, S. Sanders, N.S., P. Sklar, M. P. Snyder, M. State, J. L. Stein, P. Sullivan, A. E. Urban, F. Vaccarino, S. Warren, D. R. Weinberger, S. Weissman, Z.W., K. White, A. J. Willsey, H. Won, and P. Zandi. Additional data were generated from samples provided by the Joint MRC/Wellcome Trust (grant 099175/Z/12/Z) and the Human Developmental Biology Resource ([www.hdbi.org](https://www.hdbi.org)) grants MC/PC/13047, MR/L010674/1, and MR/L010674/2).

**Author contributions:** This study was conceived and designed by M.J.G. and C.W. with input from C.L. and K.S.P. Data was provided by D.H.G., N.S., N.J.B., D.R.W., A.E.J., T.M.H., and R.L.W. Analyses were performed by C.W., M.M., R.D., P.Z., P.F.P., D.D.V., A.B., N.M., C.J., M.K., E.T., C.H., N.A., and M.I.L. supervised by M.J.G. with contributions from C.C., D.C., H.P., M.G., N.P.D., Z.W., K.R., A.G., B.P., M.A.P., J.L.S., M.I.L., K.S.P., and C.L. C.W. and M.J.G. wrote the manuscript with major contributions from M.M.

and A.B. and critical input from all authors. **Competing interests:** This article was prepared while M.A.P. was employed at Sage Bionetworks. M.J.G. and D.H.G. receive grant funding from Mitsubishi Tanabe Pharma America. Z.W. cofounded Rgenta Therapeutics and serves as a scientific adviser for the company and a member of its board. The remaining authors declare no competing interests. The opinions expressed in this article are the authors' own and do not reflect the view of the National Institute on Aging, the National Institutes of Health, the Department of Health and Human Services, or the US Government. **Data and materials availability:** Uniformly processed gene, isoform, and splicing quantifications, harmonized metadata, full xQTL summary statistics, and other extended data are available at <https://doi.org/10.7303/syn50897018.5>. Raw genotype and RNA-seq data from the original studies can be accessed through the dbGaP with accession number phs001900 (24), the European Genome-phenome Archive under study accession EGAS00001003214 (25), the PsychENCODE Knowledge Portal at <https://doi.org/10.7303/syn2157948> (12), the HDBI portal at <https://www.hdbi.org/expression> (27), and BioProject PRJNA245228 at <https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA245228> (26). All analysis code and scripts are available on GitHub at [https://github.com/gandallab/devBrain\\_xQTL](https://github.com/gandallab/devBrain_xQTL) and Zenodo (86). An interactive web portal is available at [devbrainhub.gandallab.org](https://devbrainhub.gandallab.org). **License information:** Copyright © 2024 the authors; some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. <https://www.science.org/about/science-licenses-journal-article-reuse>

## SUPPLEMENTARY MATERIALS

[science.org/doi/10.1126/science.adh0829](https://science.org/doi/10.1126/science.adh0829)

PsychENCODE Authors and Affiliations

Materials and Methods

Figs. S1 to S15

Tables S1 to S7

References (87–121)

MDAR Reproducibility Checklist

Submitted 14 February 2023; accepted 7 March 2024  
10.1126/science.adh0829