

NEUROSCIENCE

Consensus molecular environment of schizophrenia risk genes in coexpression networks shifting across age and brain regions

Giulio Pergola^{1,2,3*}, Madhur Parihar¹, Leonardo Sportelli^{1,2}, Rahul Bharadwaj¹, Christopher Borcuk², Eugenia Radulescu¹, Loredana Bellantuono^{2,4}, Giuseppe Blasi^{2,5}, Qiang Chen¹, Joel E. Kleinman^{1,3}, Yanhong Wang¹, Srinidhi Rao Sripathy¹, Brady J. Maher^{1,3,6}, Alfonso Monaco^{4,7}, Fabiana Rossi^{1,2}, Joo Heon Shin¹, Thomas M. Hyde^{1,3,8}, Alessandro Bertolino^{2,5*}, Daniel R. Weinberger^{1,6,9*}

Copyright © 2023 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution License 4.0 (CC BY).

Schizophrenia is a neurodevelopmental brain disorder whose genetic risk is associated with shifting clinical phenomena across the life span. We investigated the convergence of putative schizophrenia risk genes in brain coexpression networks in postmortem human prefrontal cortex (DLPFC), hippocampus, caudate nucleus, and dentate gyrus granule cells, parsed by specific age periods (total $N = 833$). The results support an early prefrontal involvement in the biology underlying schizophrenia and reveal a dynamic interplay of regions in which age parsing explains more variance in schizophrenia risk compared to lumping all age periods together. Across multiple data sources and publications, we identify 28 genes that are the most consistently found partners in modules enriched for schizophrenia risk genes in DLPFC; twenty-three are previously unidentified associations with schizophrenia. In iPSC-derived neurons, the relationship of these genes with schizophrenia risk genes is maintained. The genetic architecture of schizophrenia is embedded in shifting coexpression patterns across brain regions and time, potentially underwriting its shifting clinical presentation.

INTRODUCTION

Schizophrenia (SCZ) risk increases as a function of genetic relatedness with an affected individual (1). The latest rendering of the genome-wide association study (GWAS) approach has identified 287 distinct SCZ-associated loci harboring many more potential risk genes (2). Therefore, an important current challenge is to explain how hundreds or perhaps thousands of spatially distant and biologically diverse genes confer risk for SCZ and which biological pathways they affect. Another outstanding question concerns the molecular environment in which genes mapped to GWAS loci exert their function and which other genes play key roles in such a molecular environment. Most studies so far have emphasized the functional characterization of genes mapped to GWAS loci, which likely act upon a molecular environment including genes not GWAS significant. This perspective is particularly intriguing for a disease such as SCZ that has been modeled as a condition with omnigenic heritability (3).

As the clinical illness presumably is the emergent manifestation of perturbations in cellular systems biology, not in individual genes of small effect, the study of gene coexpression in the brain is a

compelling strategy to identify convergence of SCZ risk in biologically meaningful pathways. Many SCZ risk variants control gene expression (4–9); thus, gene coexpression may be one mechanism governing the convergence of genes mediating SCZ risk into biological pathways of risk (10–12). Interindividual genetic variation associated with risk-enriched coexpressed gene sets has also been linked with intermediate phenotypes and clinical traits of SCZ (13–15).

Although coexpression in the human brain has typically been characterized as a “snapshot” of brains lumped together for a pooled analysis, it stands to reason that brain tissue will present shifting/dynamic coexpression patterns across development and age. A neurodevelopmental perspective of SCZ posits that early developmental dysfunctions in brain regions such as the dorsolateral prefrontal cortex (DLPFC) and the hippocampus may interact with molecular aspects of neurotypical brain development, causing the diagnostic symptoms to emerge later in life (16). Brain circuits whose function involves specific aspects of later maturation, such as the surge in dopaminergic innervation of the striatum and cortex in young adulthood (17), may contribute to this later onset of clinical symptoms. However, premorbid characteristics of SCZ are well described, including delays in developmental milestones, language, and cognitive capacity (18). A provocative yet biologically plausible idea ingrained in the neurodevelopmental perspective is that the same genes, e.g., genes mapped to GWAS loci, may have different effects depending on the molecular environment in which they act at different stages of development.

Notably, the study of gene coexpression related to SCZ risk has focused primarily on postnatal, mostly adult cohorts, thus largely missing a potential neurodevelopmental perspective in which genes mapped to GWAS loci act in concert with partner genes shifting in the course of development. Published reports have

¹Lieber Institute for Brain Development, Johns Hopkins Medical Campus, Baltimore, MD, USA. ²Group of Psychiatric Neuroscience, Department of Translational Biomedicine and Neuroscience, University of Bari Aldo Moro, Bari, Italy. ³Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, MD, USA. ⁴Istituto Nazionale di Fisica Nucleare, Bari, Italy. ⁵Azienda Ospedaliero Universitaria Consorziale Policlinico, Bari, Italy. ⁶Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, MD, USA. ⁷Dipartimento Interateneo di Fisica, Università degli Studi di Bari Aldo Moro, Bari, Italy. ⁸Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD, USA. ⁹Department of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA.

*Corresponding author. Email: daniel.weinberger@libd.org (D.R.W.); alessandro.bertolino@uniba.it (A.B.); giulio.pergola@uniba.it (G.P.)

considered principally the DLPFC rather than a circuit of brain regions also involved in SCZ risk, including the hippocampus (19, 20), and the associative striatum (21, 22). Although risk convergence is likely more dense in specific cell types (11, 23, 24), most work has used bulk homogenate tissue data to date. The availability of cell-specific techniques and published datasets now enables a more granular understanding of genetic risk converging into coexpression gene sets within cells, brain circuits, and multiple time points in the life span.

Prior works on autism and SCZ risk have suggested the utility of parsing coexpression into age period–defined cohorts (25, 26). For example, Werling *et al.* (27) and others reported remarkable differences between prenatal and postnatal gene expression and expression quantitative trait locus (eQTL) patterns (28–30). In this and other studies (31), however, once coexpression modules had been identified, the set of genes composing a module remained fixed. This procedure is ideal for tracking gene expression trajectories but misses how risk gene sets shift over time, potentially converging and diverging at different ages.

Here, we hypothesized that genes mapped to SCZ GWAS loci (2) would show shifting patterns of convergence into gene coexpression pathways at specific time points from early neurodevelopment through aging and that this variation would also have brain region selectivity. We thus defined SCZ gene–enriched modules (named “SCZ risk modules”) based on a stringent consensus between gene lists. We obtained these gene lists by windowing gene proximity to GWAS-significant single-nucleotide polymorphisms (SNPs) detected at $P = 5 \times 10^{-8}$ (“SCZ risk genes”), as detailed in Materials and Methods and by Pergola *et al.* (14). We hypothesized that the placement of a gene within a coexpression gene set is a component of its genetic association with SCZ and that network statistics are also associated with GWAS statistics of a given gene. We tested this hypothesis across ages in a critical circuit implicated in SCZ (DLPFC, hippocampus, and caudate nucleus). By parsing data into age period–specific coexpression networks, we searched for enrichment of coexpression gene sets for SCZ risk, allowing sets to change across age periods. We replicated our main findings using PsychENCODE and other datasets (32). Last, we identified a previously unreported set of consistent molecular partners of SCZ risk in these networks and tested whether coexpression relationships of GWAS genes representing the potential molecular environment surrounding SCZ risk genes hold in induced pluripotent stem cells (iPSCs), providing an *in vitro* platform for mechanistic investigations (33).

RESULTS

Our main dataset consisted of 562 postmortem brains from neurotypical control (NC) European and African American individuals and 186 brains from individuals with a diagnosis of SCZ (Materials and Methods and table S1). We replicated our early developmental main results in DLPFC using three postmortem datasets, including, NC RNA sequencing (RNA-seq) data from (i) 16 individuals aged 6 to 25 years from the UCLA-ASD dataset (32), (ii) 21 European and African American ancestry individuals from prenatal age to 25 years old from the BrainSpan collection (www.brainspan.org/static/download.html), and (iii) 48 European and African Americans subjects ranging from prenatal to 25 years old from the Lieber Institute for Brain Development (LIBD) polyA RNA-seq data (12, 14).

Data from the main and replication datasets were preprocessed separately (see Materials and Methods). Demographic information of all individuals (cases and controls) used across studies are shown in table S1. After preprocessing, we used weighted gene coexpression network analysis (WGCNA) to identify coexpressed gene sets (34). We included all individuals complying with inclusion criteria (Materials and Methods), regardless of age, for each brain region examined in the benchmark networks that we identified. Following WGCNA nomenclature, throughout this paper, we define as “network” the set of all genes considered, which are clustered into “modules,” i.e., a partition of this graph composed of genes highly connected with each other and not overlapping with any other module. While modules are given arbitrary color names, a particular gene set that includes all genes that remain isolated, i.e., not in any highly connected module, is called the “gray” module.

First, we performed a sensitivity analysis to establish the parameters used to identify SCZ risk modules based on enrichment for SCZ risk genes. To this aim, we used previously published networks (5, 10–14, 26, 27, 31, 35, 36). We then tested four hypotheses (each outlined in Fig. 1): (i) Age-parsed networks in neurotypical brains explain more SCZ genetic risk than the same data not age-parsed (benchmark networks). To test age parsing effects, we considered ordinal age periods (perinatal, juvenile, adult, and older adult) on the bulk tissue region samples and identified the shifting patterns of module enrichment across the human life span in those age periods (“Age study”). (ii) The course of SCZ enrichment across age differs between NC and patients with SCZ. Here, we used a sliding window approach to age (“Sliding windows”), which, unlike the age-period study, did not model fixed periods but rather the gradual variation of convergence of SCZ risk genes with age across different brain regions and compared coexpression patterns between the NC and the SCZ samples. (iii) There is a molecular environment to SCZ risk that can be identified by consensus between networks in terms of genes coexpressed with SCZ risk genes. We computed the overlap between our networks’ DLPFC SCZ risk–associated modules and previously published networks to obtain a consensus gene set coexpressed with SCZ risk genes (“Consensus genetic environment”). We thus assessed the reproducibility of coexpression gene sets and identified genes that are typically coexpressed with SCZ risk genes, with the assumption that such genes are important for the biology of SCZ risk in the brain. (iv) We tested whether an *in vitro* biological system reproduces the network features and consensus relationship that we identified in brain in a model of neuronal cells differentiated from human iPSCs (“iPSC study”).

Parameter setting and regional coexpression

We aimed to identify modules enriched for genes in the proximity of genome-wide significant SNPs, hence limiting our search to loci whose association with SCZ was corrected for multiple comparisons. As the loci defined by Trubetsky *et al.* (2) comprised 2194 genes, representing a sizable fraction of the transcriptome, we evaluated several sets of genes at multiple genomic distances from the tagging SNP (35), plus the set of 120 genes prioritized by the authors of the third wave of the Psychiatric Genetics Consortium (PGC3) manuscript (see Materials and Methods). We hypothesized that the distance of coexpressed genes from SNPs varied across different modules, and these data could be informative about the particular SNPs and genes involved. We leveraged data from published networks to establish an appropriate tradeoff between sensitivity and

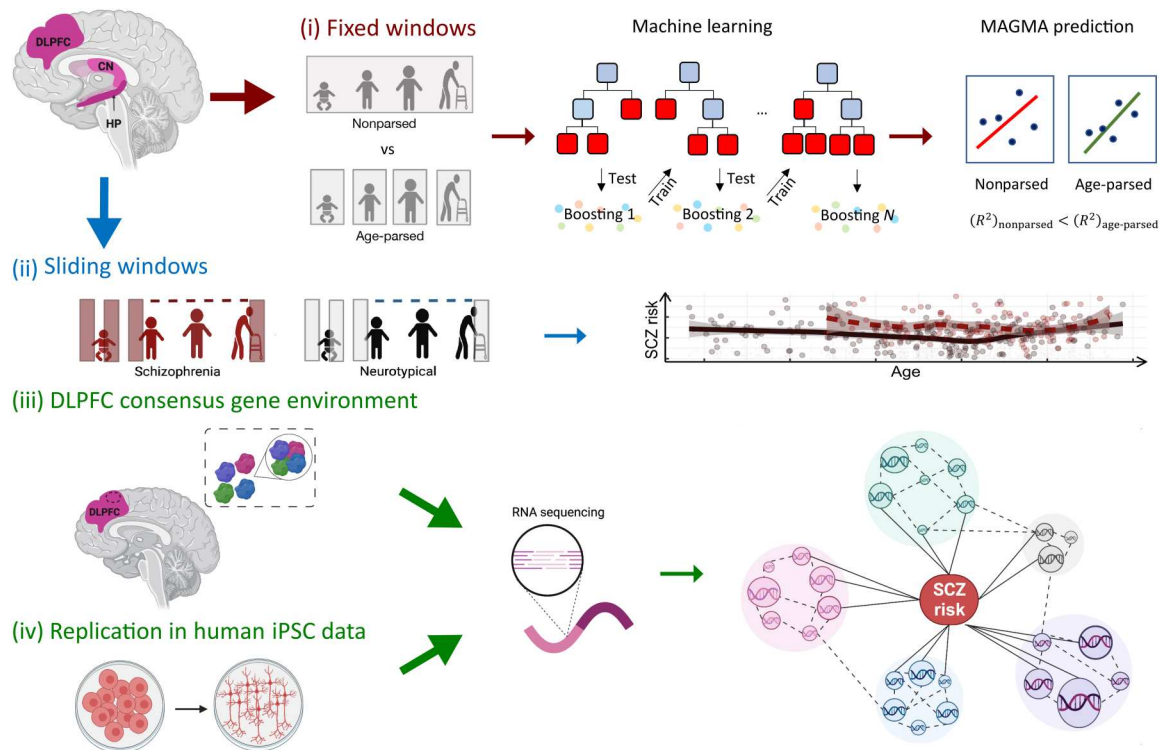


Fig. 1. Graphic summary of the study design. In this study, we evaluate four hypotheses: (i) Age-parsed networks in neurotypical brains explain more SCZ genetic risk than the same data not age-parsed. (ii) The course of SCZ enrichment differs between NCs and patients with SCZ. (iii) There is a molecular environment to SCZ risk that can be identified by consensus between networks in terms of genes coexpressed with SCZ risk genes. (iv) The consensus genetic environment surrounding SCZ risk genes can be reproduced in vitro to perform cell system studies of coexpression networks relevant to SCZ. CN, caudate nucleus; HP, hippocampus; MAGMA, multimer analysis of genomic annotation; SCZ, schizophrenia. The figure was created with BioRender.com and Inkscape.

specificity. Table S2 shows the Bonferroni-corrected significant module count for each network assessed for enrichment for protein-coding genes in the proximity of SCZ-significant SNPs. There was remarkable variability in terms of distance of the significant gene set from the SCZ-significant SNPs (Fig. 2A), as indexed by a coefficient of variation of 1.14. The coefficient of variation was computed as $SD/mean$ of the peak enrichment window genomic distance from the index SNP. The value above 1 indicates that the SD was larger than the mean [$SD = 214$ kilo-base pair (kbp); mean = 186 kbp]. We prioritized modules enriched in at least three genomic distance windows (note S1.1). Figure 2B shows the characterization of these modules in published networks from the DLPFC. By and large, we rediscovered modules already prioritized in the original reports, but the number of modules that we identified with this criterion was generally lower than the original reports, hence reflecting greater stringency. We also identified some modules not previously reported in these articles preceding the PGC3 publication (2). The biological characterization of these modules reflected neuronal biology and enrichment for loss-of-function intolerant genes (Fig. 2B). Following this calibration step, we used the same criterion and prioritized modules enriched in three genomic distance windows in our networks.

Study 1: SCZ risk gene enrichment across brain regions and age

Age parsing improves gene importance prediction models

Most previous studies along with the bulk tissue networks that we identified here across all ages embedded age information in network identification. Here, we generated age period-specific networks. Age periods were defined as a trade-off between sample size and biological meaningfulness: (i) perinatal, fetal life to 6 years; (ii) juvenile, 6 to 25 years (birth to 25 years for caudate nucleus); (iii) adult, 25 to 50 years; and (iv) older adult, over 50 years.

We obtained 13 age period-parsed gene coexpression networks (four age periods \times three brain regions without perinatal caudate and two replication networks). Table S3 summarizes the descriptors of these networks. Given the evidence implicating early neurodevelopmental components of SCZ (26, 37), our central hypothesis in the age study was that age-parsed coexpression explained SCZ genetic risk better than nonparsed coexpression. To assess the explanatory power of parsed networks, we investigated the genetic convergence of the SCZ genetic signal by a continuous measure of that signal across all genes. To this end, we used MAGMA (multi-marker analysis of genomic annotation) in its revised version (38, 39) to compute for each gene a Z score of association with SCZ based on GWAS effect size (2) of SNPs in the GWAS loci. First, we assessed whether splitting the datasets into age periods with a smaller sample size enhanced the accessible biological information. We computed linear models to associate gene importance with

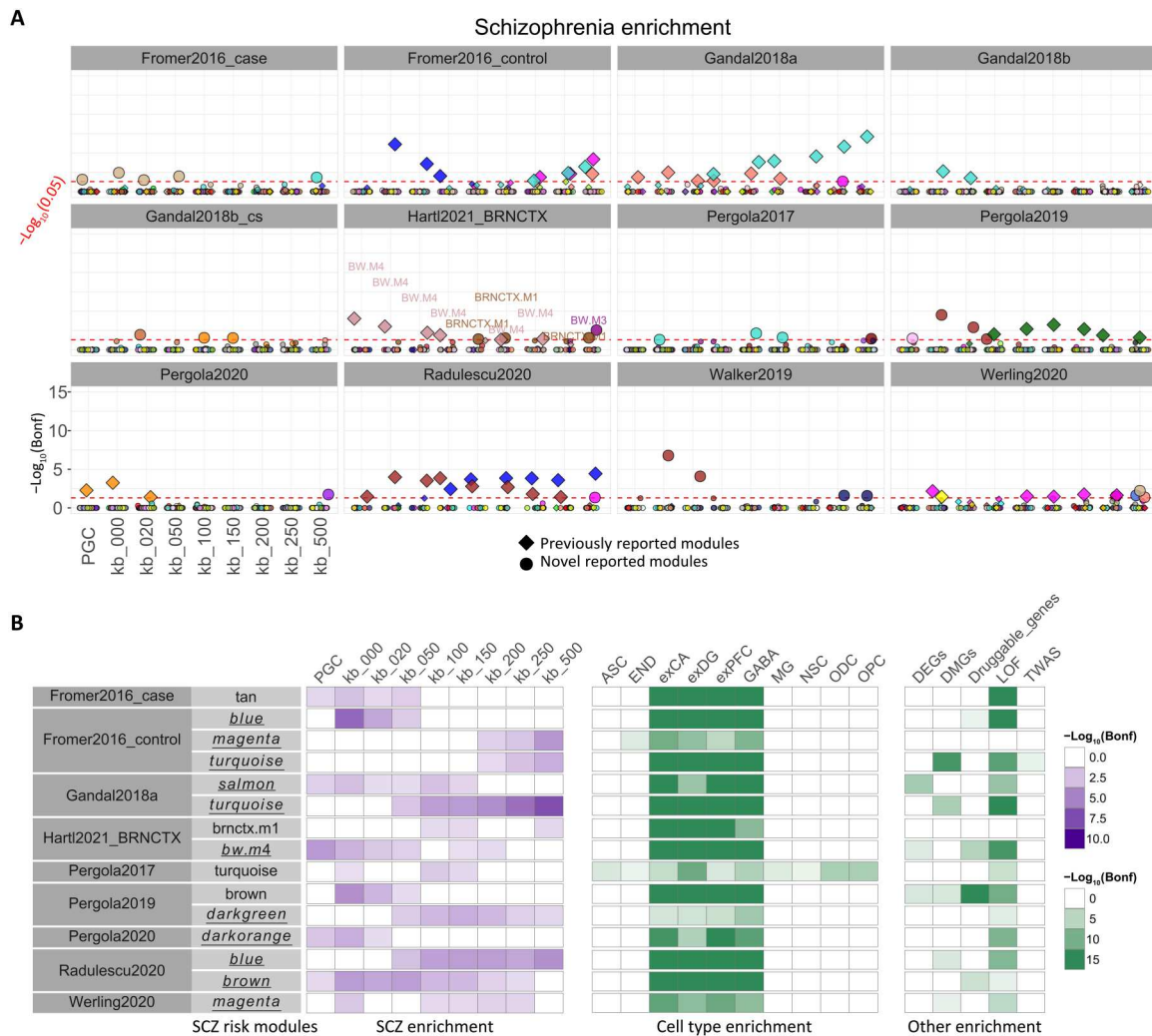


Fig. 2. Coexpression modules enriched for SCZ risk. (A) SCZ risk gene enrichment in previously published networks. SCZ risk was assessed in the coexpression gene sets in 12 previously published networks for nine reference lists of expanding SCZ risk protein-coding genes [bins: PGC (120 prioritized genes) (2), kb_000, kb_020, kb_050, kb_100, kb_200, kb_250, and kb_500]. The y axis represents enrichment value $-\log_{10}(\text{Bonferroni-corrected } P \text{ value})$ for a module in that bin (hypergeometric test; see Materials and Methods). The dotted red line denotes $-\log_{10}(0.05)$ threshold. Significant modules (Bonferroni-corrected P value < 0.05) are magnified. Colors correspond to the name assigned by WGCNA to the coexpression module except in the Hartl2021_BRNCTX network. The gene sets and reference lists were restricted to protein-coding genes. Diamonds (\blacklozenge) represent modules previously highlighted for SCZ risk enrichment in the respective publications. **(B)** Risk module functional characterization. We identified 15 SCZ risk modules across nine published networks as significantly enriched for SCZ risk in at least three of the nine lists of protein-coding SCZ risk genes. Risk modules were characterized for different enrichments (SCZ, cell type, and other; see Materials and Methods). Only significant enrichments (Bonferroni-corrected P value < 0.05) are shown. Module names with underline represent previously reported modules. ASC, astrocytes; DEGs, SCZ differentially expressed genes between patients and controls derived from separate caudate, DLPPFC, and HP studies; DMGs, genes proximal to differentially methylated CpG islands in SCZ; Druggable_genes, SCZ drug target genes; END, endothelial cells; exCA, pyramidal neurons from the hippocampal CA region; exDG, granule neurons from the hippocampal dentate gyrus; exPFC, pyramidal neurons from the prefrontal cortex; GABA, GABAergic interneurons; LOF, loss-of-function intolerant genes; MG, microglia; NSC, neuronal stem cells; ODC, oligodendrocytes; OPC, oligodendrocyte precursor cells; TWAS, SCZ transcriptome-wide association study genes derived from separate caudate, DLPPFC, and HP studies.

coexpression, in which the outcome variables were MAGMA scores obtained for each gene and predictors were module assignment (categorical) and total connectivity of each gene in each network. We compared linear models obtained with the entire cohort (no parsing in three brain regions: DLPFC, hippocampus, and caudate nucleus) versus the 11 age period-parsed networks.

We first evaluated the direct relationship between total gene connectivity and MAGMA score (an example scatter plot can be seen in fig. S1). We found a significant positive correlation in all networks

except the perinatal (all other Bonferroni-corrected $P < 0.001$). This result suggests that genes with higher SCZ MAGMA scores are generally more strongly connected with all other genes, although the effect size was small for all (all Kendall's tau < 0.054).

Both nonparsed and age-parsed gene module assignments were significantly associated with MAGMA scores (complete statistics in note S1.3). In age-parsed networks, we found that perinatal and juvenile networks consistently, across all tissues, showed a stronger association between SCZ importance and module assignment

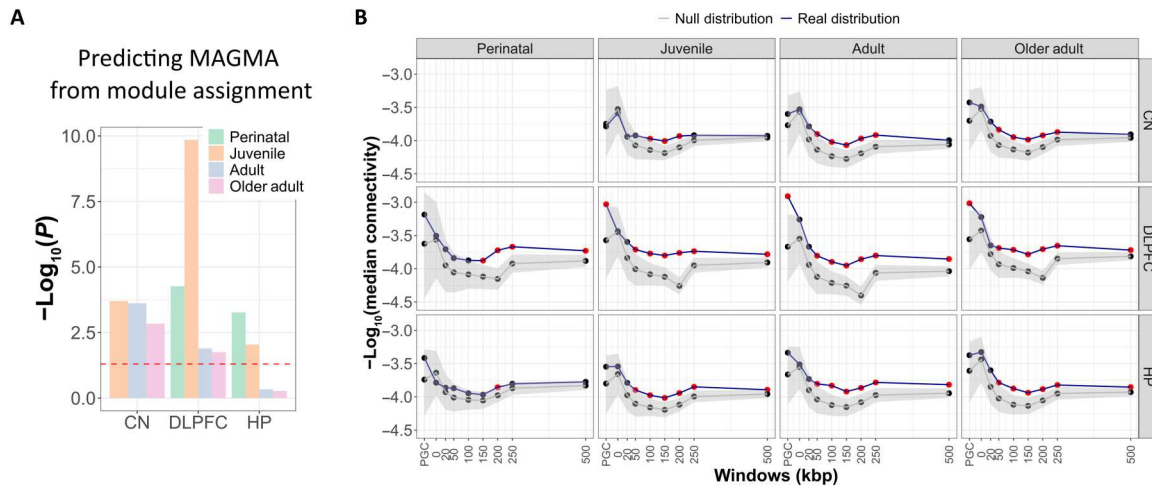


Fig. 3. SCZ risk is better explained in age-parsed networks. (A) A linear model was used to predict MAGMA score from module assignment and total connectivity predictors considering all 11 age-parsed networks. The graph shows the $-\log_{10}(P)$ for each age-parsed network to predict MAGMA scores from module assignment within this linear model. (B) Within-set gene connectivity (coexpression) for different SCZ gene sets, across tissue and age. On the y axis is the median of the gene-wise connectivity for each set, with each SCZ gene set represented on the x axis. Connectivity values were compared to a null distribution of gene sets of equal size, length, GC content and average expression. The null 95% confidence interval (CI) is in gray shade; the mean value is the gray line. Red dots represent SCZ gene sets above the null 95% CI.

than adult networks (Fig. 3A). This finding suggests that genetic risk for SCZ is more aggregated in networks during earlier development than in adulthood, hence pointing to a functional role of SCZ risk genes in early development. Thus, evaluating gene assignments and trajectories during these early stages is relevant for understanding the biology of SCZ risk. In addition, perinatal hippocampus total gene connectivity was negatively associated with MAGMA, which was not the case in DLPFC networks. This result means that genes most associated with SCZ were overall less connected with all other genes (and thus less central in the network) in hippocampal in contrast to DLPFC networks, independently of their module assignment.

Computing age-parsed networks explained significantly more variance in MAGMA scores for SCZ than using the same samples without age parsing to compute a single network per tissue (Vuong test for non-nested models, $z = 8.4$, $P < 2.2 \times 10^{-16}$). To validate this result, we implemented a regression algorithm to predict MAGMA scores to compare the prediction performance of our machine learning workflow (Fig. 4A; see Materials and Methods and method S2.2) between age-parsed and nonparsed datasets (see Fig. 4). Both datasets contained 21,751 genes and the same intrinsic attributes but different expression and coexpression properties.

The R^2 distributions obtained for the Full and Age-parsed datasets are shown in Fig. 4B. A Wilcoxon test indicated a significant improvement ($P = 0.015$) in the performance obtained with the Age-parsed dataset. Among the most relevant features in predicting the MAGMA score were intrinsic attributes such as gene width and GC content, the median expression variables, and the module membership metrics (KME) associated with the modules of the coexpression network (Fig. 4, C and D). Explanatory KME variables for modules that are not significantly enriched for SCZ are related to predicting the entire range of MAGMA score values, also including low-score genes (some module membership variables have a significant negative correlation with the MAGMA score). In the age-

parsed prediction models, the features selected in each and every run included the median gene expression in perinatal DLPFC/hippocampus and juvenile DLPFC/hippocampus module membership but not in adult age stages.

Together, these analyses suggest that (i) genes with high MAGMA score for SCZ tend to be coexpressed, as module membership contributes to SCZ importance prediction into an unseen testing set, and (ii) parsing genes into age periods, data being equal, reveals stronger associations between networks and a continuous measure of gene association with SCZ beyond GWAS-significant loci.

SCZ top loci genes converge into coexpression modules in multiple brain regions, though especially in early-stage DLPFC

The publications we used to perform the sensitivity analysis for module characterization focused on cortical samples. We then compared these published results with our networks from DLPFC, hippocampus, and caudate nucleus with and without parsing for age. We interrogated the properties of SCZ risk genes across age-parsed networks, asking whether sets of putative SCZ risk genes (defined from PGC3 and at varying genomic windows from PGC3 SNPs) were highly coexpressed with each other regardless of module membership. We found that SCZ risk genes are more coexpressed with each other as compared to a null distribution of gene sets of equal set size, gene length, GC content, and average expression. Results vary across brain regions and age, with peak statistics around 200 kbp from GWAS-significant SNPs (Fig. 3B). The perinatal results were not principally driven by postnatal samples, as shown by significant departure of connectivity estimates from the null distribution (windows: 150 to 500 kbp; $P[\text{empirical}] < 0.001$).

In most networks, SCZ risk genes were depleted, i.e., underrepresented in hypergeometric tests, in the gray module, more extensively than control gene sets negative for SCZ. Because gray modules in WGCNA are composed of nonclustered genes, these results suggest that SCZ risk genes tend to be coexpressed in modules

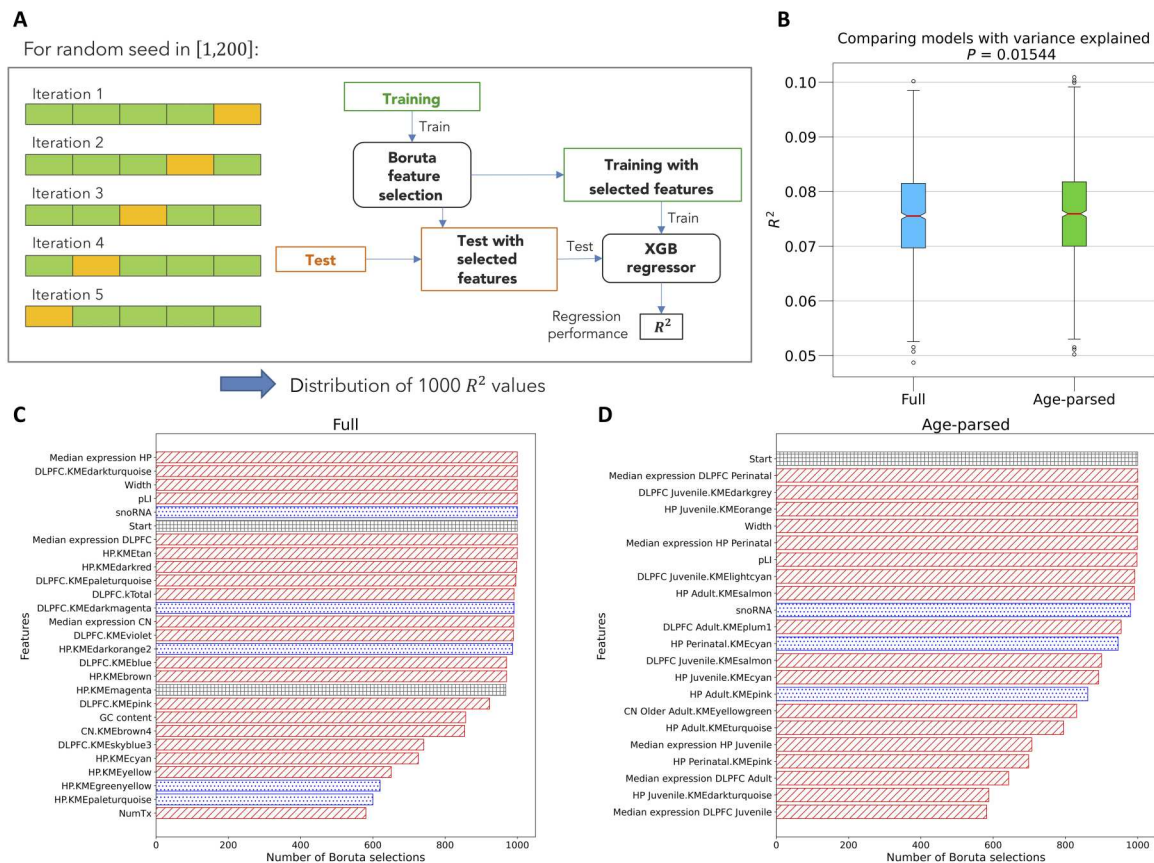


Fig. 4. The proposed machine learning model is more effective at predicting MAGMA score when trained on age-parsed rather than age-aggregated network features. (A) Workflow of the analysis: Using a random number generator, the 21,751 genes are randomly assigned to five different partitions, and for each partition, the algorithm performs a round of a fivefold cross-validation, stratified with respect to the chromosomes to which genes belong; then, it applies the Boruta feature selection algorithm to the training set and trains an XGBoost regressor on the same set, using only the selected features. XGBoost predicts the MAGMA score of genes in the test set; One thousand values of R^2 are collected during the analysis, corresponding to different random seeds and different iterations of fivefold cross validation. (B) Distributions of the 1000 R^2 values obtained in the MAGMA predictions based on the Full dataset (light blue), containing variables related to expression and connectivity in the three age-aggregated networks, and on the Age-parsed dataset (green), containing analogous variables for the 11 age-parsed networks. The Wilcoxon test reveals a statistically significant ($P = 0.01544$) improvement of performance in the Age-parsed dataset with respect to the Full dataset. (C) Features of the Full dataset selected by the Boruta algorithm in more than 500 runs. The bars are filled with red lines, blue dots, or gray squares, depending on whether the Pearson correlation of the corresponding feature with the MAGMA score is, respectively, positive and statistically significant ($P < 0.05$), negative and statistically significant ($P < 0.05$), or not significant. (D) Features of the Age-parsed dataset selected by the Boruta algorithm in more than 500 runs. The meaning of the patterns used to fill the bars is the same as for (C).

with other genes, rather than not (note S1.4 and fig. S2). This finding is consistent with the correlation between MAGMA scores and connectivity shown above. The underrepresentation in gray, however, may relate at least in part to the fact that SCZ risk genes tend to be more expressed in brain tissue and highly expressed genes are less likely to be in the gray module (note S1.1 and fig. S3). Eight of 11 age-parsed networks included at least one SCZ risk-enriched module (Fig. 5A). The genomic window in which most of these modules were significant included genes localized up to 200 kbp from the SCZ index SNP. SCZ risk modules were generally selective for SCZ genes and were not enriched for other disorders (fig. S4), although hippocampal modules were an exception, representing enrichment for multiple pathologies. DLPFC enrichments were not driven by many genes localized in few loci but were characterized by the convergence of many loci, while the opposite case was seen in the hippocampus (note S1.4). Considering that all DLPFC networks identified coexpression convergence in at least one

module, DLPFC aggregation of SCZ risk genes appears to endure across multiple age periods in contrast to caudate or hippocampus, which are relatively age restricted. In summary, the coexpression of genes in the proximity of GWAS-significant SNPs associated with SCZ is found most convincingly in the DLPFC.

SCZ risk modules had predominantly neuronal cell specificity (Fig. 5B). Accordingly, a study of dentate gyrus tissue enriched for granule cells via laser capture microdissection (23) showed a greater degree of SCZ risk gene convergence on coexpression patterns compared to bulk hippocampus RNA-seq and insight into the context-dependent functions of these genes otherwise unavailable (note S1.2, methods S2.1, and fig. S5). In addition, several of the age-parsed modules were enriched in genes proximal to CpG islands differentially methylated in SCZ, which was not the case in nonparsed networks, suggesting a different gene aggregation and consistent with the expected epigenetic variation associated with age. Note how this property contrasts with the sparse other

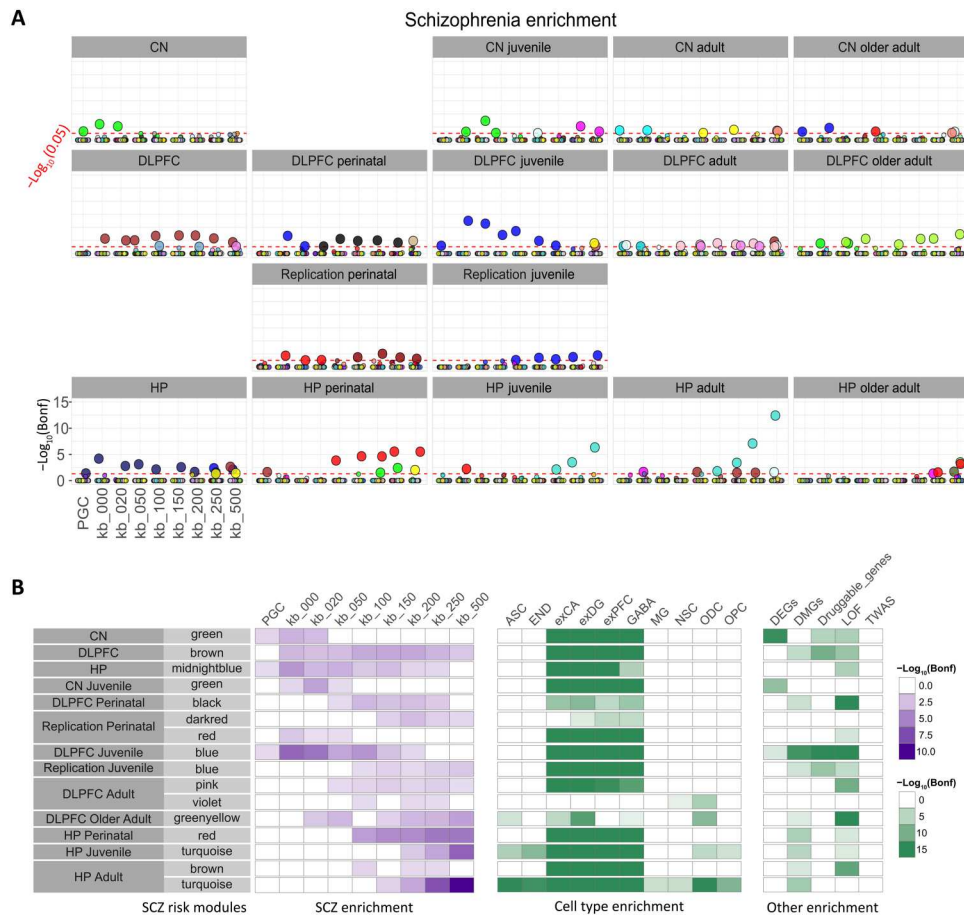


Fig. 5. SCZ risk modules in networks parsed by age. (A) SCZ risk genes enrichment in the generated networks. Only networks with modules significantly enriched for SCZ top protein-coding risk genes are shown in (B). See Fig. 2 caption for details on the tests computed and abbreviations in (A) and (B).

enrichments found in previously published enriched modules, which were not parsed by age and mainly relied on adult brain samples (Fig. 2B).

Regulomic analyses identified several transcription factors whose targets were overrepresented across multiple networks (Fig. 6A), with the most evidence favoring *KLF15*, *MAZ*, and *SP2*. Other transcription factors of the *SP* family with a targetome profile similar to *SP2* were also included, such as *SP4*, a gene prioritized by the PGC3 GWAS (2) and enriched for rare damaging coding variants in SCZ at exome-wide significance (40). SCZ risk modules were similar across regions in terms of top gene ontologies (e.g., involving synapses; Fig. 6B).

Risk convergence trajectories point to specific and shared gene functional profiles in the DLPFC and hippocampus across age

The premise of this work was that SCZ risk gene coexpression might change across ages. We measured the overlap between modules in sequential age periods in terms of Jaccard index (JI) at 200 kbp (as reported above, the extension window with the greatest number of significant SCZ risk modules). In the DLPFC, most perinatally coexpressed SCZ risk genes (module black) continue to be coexpressed in juvenile samples (module blue), along with many of their coexpression partners (Fig. 7A), a pathway strongly enriched for neurodevelopment (note S1.5 for additional details and

statistics). In the context of a prominent overlap between SCZ risk modules across age periods, the coexpression of a cohesive subset of SCZ risk genes remained high (67% still coexpressed and 33% shifted), while most of their module partners shifted (52% partner genes). As the perinatal window included individuals aged from fetal life to 6 years old, we performed a sensitivity analysis to assess whether this difference in continuity of coexpression after birth was driven by postnatal samples (note S1.5). We found that the prenatal module red including only fetal samples recapitulated the perinatal black enriched module (preservation Z score = 41; see fig. S6); also in this case, the proportion of SCZ risk genes coexpressed in the juvenile age stage was higher (46% prenatal genes still coexpressed in the juvenile samples and 54% shifted) than the proportion of non-SCZ risk genes (68% partner genes shifted). We further reproduced this finding in our replication datasets, where we identified another perinatal and another juvenile DLPFC network, respectively. In the replication perinatal DLPFC, module red included 21 SCZ risk genes, of which 12 (57%) were also coexpressed in the replication juvenile blue module, whereas 43% shifted module membership. The 66 genes overlapping between perinatal and juvenile SCZ risk modules were strongly enriched for modulation of chemical synaptic transmission (17 genes, $P_{\text{adjust}} = 2.46 \times 10^{-7}$) and regulation of synaptic plasticity (9 genes, $P_{\text{adjust}} = 6.64 \times 10^{-6}$). Once again, while most SCZ risk genes were still

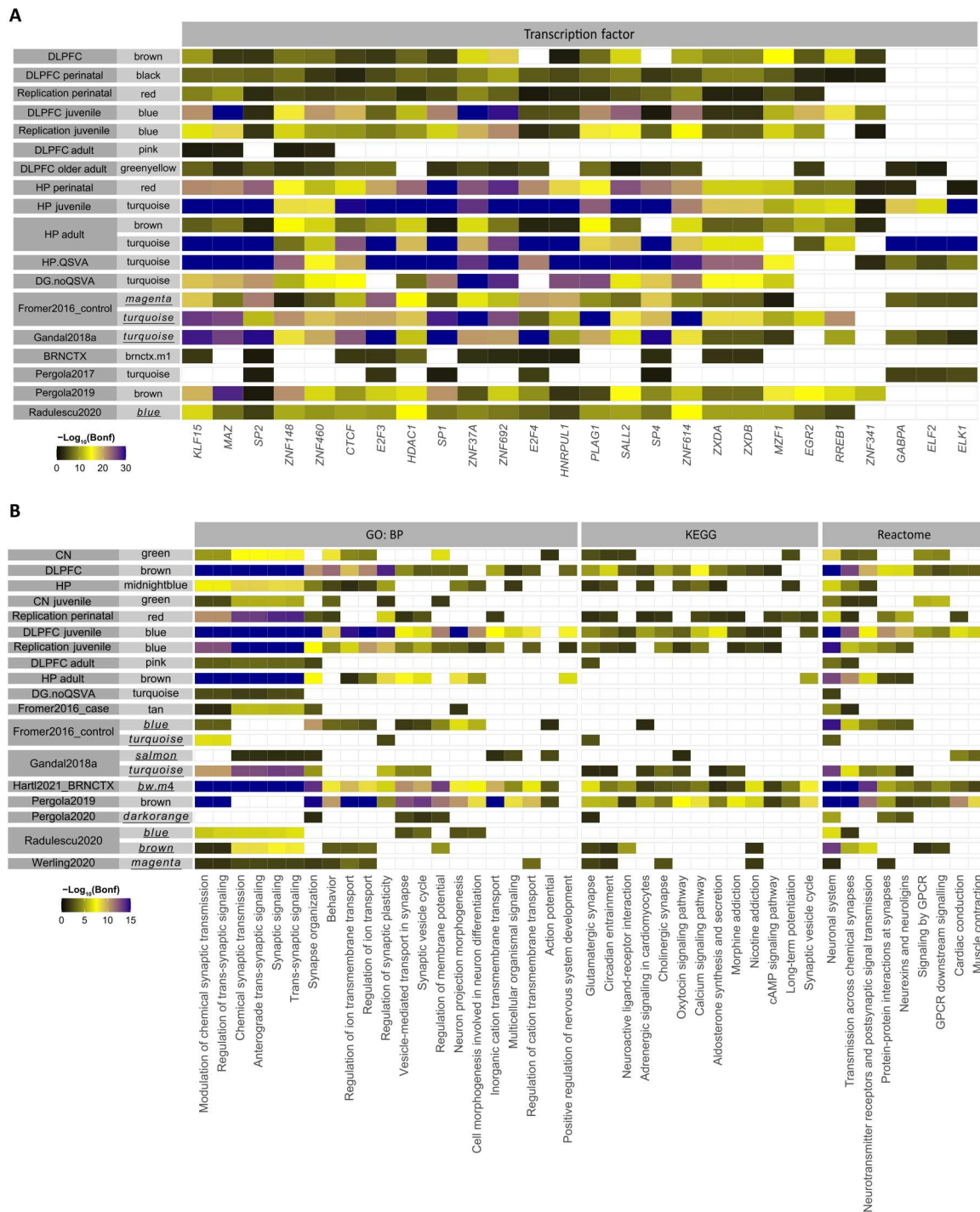


Fig. 6. Regulome enrichment and functional characterization of schizophrenia risk modules. To obtain this figure, transcription factor target enrichments were first restricted to corrected $P < 0.05$. Then, for modules significant for SCZ in at least three extension windows (shown in Figs. 2 and 5 and fig. S5), a union of top three most significant transcription factors for each module is considered for plotting. A white block indicates that the targetome of the transcription factor is not significantly overrepresented in that module. **(A)** Transcription factor targetome enrichment. **(B)** Gene ontology: biological processes (GO: BP), Kyoto Encyclopedia of Genes and Genomes (KEGG), and Reactome pathways. DG, dentate gyrus granule cell layer. For both (A) and (B), processes/transcription factors significant in less than four risk modules and risk modules with less than two significant processes/transcription factors are not shown. Underlined module names represent modules previously reported as associated with SCZ.

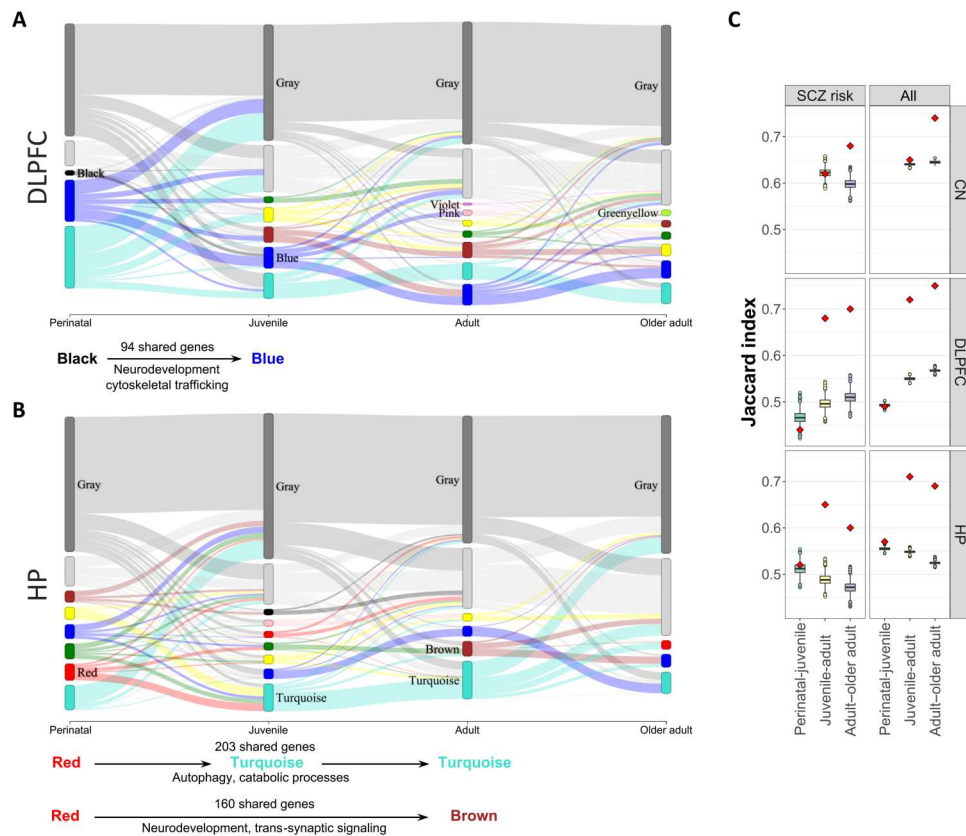


Fig. 7. Coexpression transitions of SCZ risk genes at various age periods. Module membership of SCZ risk genes (200-kbp list; see Materials and Methods) for age-parsed networks arranged ordinally. Modules for a network are represented as nodes. The size of node corresponds to the number of SCZ genes in the module. The link width represents the number of shared SCZ genes between modules from two age-parsed networks. Only the risk modules (from Fig. 5) are labelled. The bottom text highlights select conserved “streams” between risk modules across age. **(A)** DLPFC age-parsed networks. **(B)** HP age-parsed networks. In **(A)** and **(B)**, a general discontinuity from perinatal to postnatal ages is evident, but some risk gene sets continue to be expressed in the perinatal-to-juvenile transition. For clarity of illustration, modules with less than 30 SCZ risk genes for the 200-kbp list that are not enriched for SCZ are combined into a hypothetical “lightgray module.” **(C)** Gray composition age transitions in JI across brain regions, where box plots depict the null distribution, and the red diamond represents the actual JI measured. The left panels considered only SCZ risk genes, whereas the right panels considered all genes. Notice a transition with low JI between the gray modules of the perinatal and juvenile period in DLPFC. The gray module represents the unclustered genes not belonging to any modules in the network.

coexpressed, their non-SCZ partners shifted (64%). Accordingly, table S4 shows that all SCZ risk genes most consistently coexpressed into SCZ risk modules, even in adult age, were already coexpressed in perinatal age.

In the hippocampus, the SCZ risk module in the perinatal age period split into two separate developmental coexpression pathways: One mirrored early DLPFC age period coexpression related to neurodevelopment, and another involved autophagy functions. In the DLPFC, autophagy was associated with SCZ risk genes only in the perinatal age period, suggesting that the hippocampus continues coexpressing autophagy-related genes with SCZ risk genes in adulthood (Fig. 7B). The pattern of continuity and shift in coexpression that we observed in perinatal and juvenile DLPFC was not apparent in the hippocampus, neither in the main dataset nor in the replication (note S1.5). Caudate age-parsed networks were not considered because we only found an enriched module in the juvenile network (Fig. 5).

Sankey plots (Fig. 5, A and B) further show that gray module composition changes between perinatal to juvenile periods (juvenile to adult in the caudate nucleus, Sankey plot not shown) and later

transitions. These changes were quantified by computing a JI across age adjacent gray modules. Across early life transitions, the JI measured between gray modules did not significantly differ from the null distribution (gray modules were no more similar than two randomly selected gene sets), indicating that gray module composition changes between early life periods. Instead, the JI was significantly higher across later life transitions than the null distribution, indicating that gray composition remains stable in later life periods (Fig. 7C). For additional results, along with Sankey plots for all genes, not just SCZ risk genes, in the networks, see note S1.5. Interactive versions of full Sankey plots are available online at <https://doi.org/10.5281/zenodo.5676480>.

Study 2: Differences between neurotypical and patient-derived networks
Lifespan trajectories of coexpression show early SCZ risk convergence and differences between patients with SCZ and NCs

Given that preassigned age periods do not capture the gradual changes intrinsic to neurodevelopment and results may be affected

by the sampling density of subjects along the age dimension, we identified networks in sliding windows of 40 subjects with steps of one subject and compared NC with patients with SCZ. The same samples analyzed in the fixed age period study were used here, with the addition of cohorts of patients with SCZ. We thus generated 316 networks from the DLPFC (221 from NC and 95 from patients with SCZ), 306 networks from the hippocampus (234 from NC and 72 from patients with SCZ), 329 networks from the caudate (219 from NC and 110 from patients with SCZ), and 77 networks from dentate gyrus (46 from NC and 31 from patients with SCZ). Figure 8A shows early SCZ enrichment in neurotypical DLPFC that is superseded later, first by the hippocampus and then by the caudate. The most notable finding is that networks derived from patients with SCZ persistently show greater

enrichment for SCZ in caudate nucleus. In contrast, the dentate gyrus data, enriched for neuronal cell type, mainly showed higher enrichment in controls at the age range overlapping with patient-derived networks, reaching a higher peak of enrichment than most other brain regions across age. Although different age trends may reflect age-related changes in cell composition in patients with SCZ or the effect of drugs, stress, chronicity, etc., it is noteworthy that estimated neuronal proportion based on cell decomposition did not explain the difference in the caudate observed between patients and controls (Fig. 8B).

To further characterize these networks biologically, we correlated MAGMA gene scores representing the association with SCZ with gene ontology enrichments across modules within each sliding window network (Fig. 8, C and D). There is a strong association

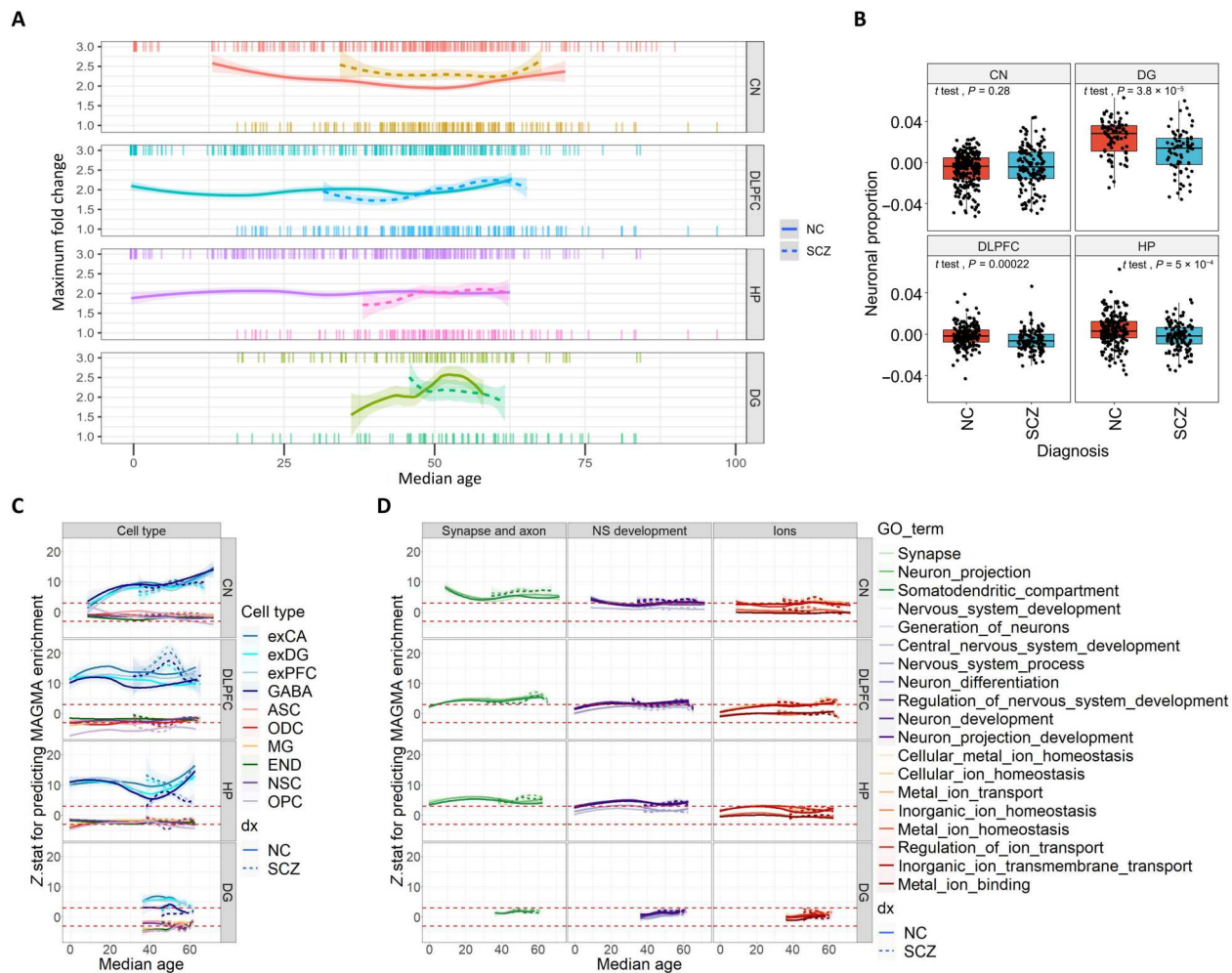


Fig. 8. Across module association of cell type and GO with MAGMA enrichment in sliding window networks. (A) Overrepresentation of SCZ top loci genes in age-varying sliding windows. Multiple coexpression networks were identified with samples first chronologically arranged and then selected via a sliding window of width 40 samples, shifting one sample at a time. The x axis represents the median age of individuals in the window. The y axis represents the maximum SCZ fold enrichment (200-kbp list) among all modules in that window. Only the smoothed line connecting maximum fold enrichments is shown. The shaded area represents the 95% CI of smoothed line. The top rug and bottom rug show the individual NC and SCZ sample ages, respectively. (B) Neuronal proportion in age-matched samples. Estimated neuronal proportion was significantly greater in neurotypical individuals than patients with SCZ in DLPFC, HP, and DG (*t* test; $*P < 0.05$). (C) We derived *P* values for cell type and MAGMA enrichment for all modules in all sliding window networks. A robust linear model was constructed for each network to evaluate the association between cell type enrichment and MAGMA enrichment. (D) The gene ratios of GO term lists were also associated with MAGMA enrichment scores across modules. For (C) and (D), the y axis shows the Z-stat for cell type or GO term enrichment predicting MAGMA enrichment; the x axis shows the median age of the sliding window network. The red dashed lines indicate a nominal *P* = 0.05. SCZ, patients with schizophrenia; NS, nervous system; NSC, neural stem cells.

with neuronal cell types in both cases and NC at all ages, except in young caudate NC networks (Fig. 8C). This association suggests a large overlap of neuronal and SCZ enrichment across modules. Gene ontology term associations also show consistency between NC and case networks. The most significant gene ontology terms are those related to synapse and axon, which are above nominal significance across all ages in the caudate, DLPFC, and hippocampus (Fig. 8D, green lines). The only differences across diagnoses are found in the caudate with synapse- and axon-related terms, suggesting a greater overlap in modules between enrichment for SCZ risk and for synaptic function in patients. This result also suggests that the SCZ risk genes responsible for the greater convergence found in networks derived from patients (Fig. 8A) are preferentially synaptic genes. In the hippocampus, modules derived from patients with SCZ are relatively less enriched for GABAergic neurons compared to excitatory neuronal types (Fig. 8C). This brain region also shows the strongest association to nervous system development (Fig. 8D, violet lines), found at the same level of hippocampus synapse and axon terms, which is interesting given the potential role of adult neurogenesis in this region.

Study 3: Consensus genetic environment Identification of reliable coexpression partners of SCZ risk genes in perinatal and later postnatal prefrontal cortex

Last, we performed a consensus analysis to test the hypothesis that a consistent molecular environment surrounds SCZ risk genes, i.e., a potential locus of molecular convergence involved in how diverse small effect genes might translate toward a common phenotype. Across multiple analyses, SCZ risk genes were most robustly coexpressed in neurotypical DLPFC modules, as outlined above. Considering that most previous studies of gene coexpression networks investigated the DLPFC, we adopted a consensus approach to identify the genes most often included in coexpression modules overrepresenting top SCZ loci genes in the current and these prior studies.

Table 1 shows genes coexpressed in SCZ risk modules in at least two-thirds of the networks that had at least one module enriched with SCZ risk genes in (i) cortical fetal samples [the perinatal age period-parsed network and the networks published in (26, 27, 31)] and (ii) postnatal networks [juvenile, adult, older adult DLPFC networks presented here, plus those reported by others (5, 10–14, 35)] (Fig. 9A). These “consensus” GWAS partner genes include four from the perinatal consensus (whereas 50 genes were found using the SCZ-negative gene list) and 24 genes in the SCZ list in the postnatal consensus (with none above the threshold in the negative list), all protein-coding genes. This finding suggests that including only four networks, as in the perinatal study, may be insufficient to disentangle biological signal from noise in a consensus analysis. The perinatal findings should thus be taken as exploratory.

In this full list of consensus genes, SCZ risk genes defined on the basis of PGC3 loci were a minority (18%). Only three genes of the 28 appeared among the 120 PGC3 prioritized genes (*DLGAP2*, *ATP2A2*, and *OPCML*). Therefore, genes not located in GWAS-significant loci are often coexpressed with PGC3 loci genes, more often than most PGC3 genes themselves. The lack of overlap between the consensus gene lists and GWAS loci genes aligns with the finding that SCZ risk genes shift coexpression partners over time. Although these genes were found coexpressed with SCZ risk genes only postnatally, most of them are expressed at higher levels in perinatal and juvenile than in adult life (Fig. 9E). We used H-MAGMA (28, 39),

which considers contact points in chromatin organization to map variants to genes within MAGMA, to characterize the fetal and adult risk score of each gene (Table 1). Gene ontology analyses on this list of 28 perinatal/postnatal genes point to the functions “chemical synaptic transmission” (seven hits, 5.6-fold enrichment, and $q_{\text{FDR}} = 0.02$), “cation transport” (eight hits, 4.7-fold enrichment, and $q_{\text{FDR}} = 0.02$), and “cAMP biosynthetic process” (two hits, 125.5-fold enrichment, and $q_{\text{FDR}} = 0.02$). The top associations of consensus genes with gene ontologies are represented in Fig. 9B.

We noticed that most of the 28 consensus genes have high probability of being loss-of-function intolerant. Loss-of-function intolerance is related to the depletion of disruptive variants. The most frequent coexpression partners of SCZ risk genes are more loss-of-function intolerant than background genes: *pLI* for SCZ risk genes averaged 0.49, whereas for consensus non-GWAS genes, it averaged 0.61; both estimates surpass those of all other protein-coding genes in the genome expressed across at least four perinatal or six postnatal networks ($N = 9,552$; background median $pLI = 0.22$; one-sample Wilcoxon signed-ranked test, $t_{30} = 8.6$, one-sided $P = 0.006$). High *pLI* is thought to characterize potential SCZ core genes (11), and frequent coexpression partners of SCZ risk genes share this feature. No transcription factors or microRNA (miRNA) targets were significantly overrepresented in the consensus sets.

We derived empirical P values for each consensus gene by permuting module membership 10,000 times and identifying a null distribution of consensus genes (see Table 1 and Materials and Methods). Each consensus gene had a higher probability of being coexpressed with SCZ risk genes across both perinatal and postnatal networks than expected by chance. Moreover, obtaining ≥ 4 (perinatal) or ≥ 24 (postnatal) consensus genes as output of this analysis by chance is unlikely for randomly selected genes (perinatal empirical $P = 0.03$; postnatal empirical $P = 1 \times 10^{-4}$).

We looked at the expression of consensus genes in DLPFC 10 \times single-cell sequencing data taken from (41) (https://github.com/LieberInstitute/10xPilot_snRNAseq-human), which has preannotated cell type and neuronal subtype clusters. Expression levels of consensus genes were consistently higher in neuronal cell types versus others (fig. S7). We then evaluated the environment around consensus genes for each network. For each network, we generated gene sets from the union of all module partners of consensus genes (even if that module was not SCZ enriched). Across all networks, consensus partners were very strongly enriched for SCZ risk genes, whereas the union of modules not including consensus genes was never enriched for SCZ risk genes despite this meta-set including more genes (meta_set_with_consensus median = 3086, SD = 1467; meta_set_without_consensus median = 6653, SD = 3470; fig. S8), indicating that consensus SCZ risk pathways may be dispersed across multiple SCZ risk modules and may not come out as statistically significant when compared to each other. The consensus partner set showed enrichments related to synapse and neuron (fig. S9A). Kyoto Encyclopedia of Genes and Genomes (KEGG) annotations revealed additional enrichments in circadian entrainment, oxytocin signaling, and addiction (fig. S9B). The flow of consensus genes across age mirrored closely the flow of SCZ risk genes, with several overlapping modules (Fig. 9D).

Study 4: Consensus genes in iPSC-derived neurons

Last, we explored whether this consensus set of SCZ coexpression partners is coexpressed with SCZ risk genes in a cell model of early

Table 1. Consensus coexpression gene sets. Perinatal consensus considered four studies including fetal samples. Postnatal consensus considered 11 networks only including postnatal samples. Hits, the number of enriched modules each gene appeared in; total, the number of networks each gene was annotated in. Chr, chromosome; NA, not assessed; SCZ, lists of genomic extension windows including the gene among the top SCZ loci genes (distances in kbp); pLI, probability of loss-of-function intolerance (loss of function both with homozygous and heterozygous deleterious mutations); DEGs, SCZ differentially expressed genes (+/-), over/underexpressed in patients with SCZ relative to NCs); DMGs, differentially methylated genes; P value, empirical P value obtained from permutation of consensus genes. References in the table: [1] (5), [2] (91), [3] (94), [4] (90), [5] (93), and [6] (89).

Ensembl ID	Symbol	Chr	SCZ	pLI	MAGMA	H- MAGMA adult	H- MAGMA fetal	DEGs	DMGs	Hits/ total (ratio)	P value
Perinatal consensus											
ENSG00000101204	CHRNA4	20		0.021	2.13	1.49	4.08			3/4 (0.75)	1 × 10 ⁻⁴
ENSG00000119522	DENND1A	9		0.90	1.41	1.11	1.45			3/4 (0.75)	1 × 10 ⁻⁴
ENSG00000162104	ADCY9	16	500	0.97	2.15	0.96	1.44	[1] [-]	[2,3]	3/4 (0.75)	2 × 10 ⁻⁴
ENSG00000173175	ADCY5	3		0.99	-0.2	-0.26	-0.32		[2]	3/4 (0.75)	2 × 10 ⁻⁴
ENSG00000196588	MKL1	22	0-500	0.54	2.59	4.76	4.58			8/10 (0.80)	2 × 10 ⁻⁴
ENSG00000198010	DLGAP2	8	PGC	0.95	4.88	2.67	4.5		[3,4]	7/9 (0.78)	9 × 10 ⁻⁴
ENSG00000215012	RTL10	22		5.8 × 10 ⁻⁶	0.88	1.11	1		[2]	6/8 (0.75)	5 × 10 ⁻⁴
ENSG00000002746	HECW1	7		0.99	0.4	0.36	1.69		[3,5]	8/11 (0.73)	1 × 10 ⁻⁴
ENSG00000105649	RAB3A	19		0.87	0.57	-0.1	-0.31		[3]	8/11 (0.73)	1 × 10 ⁻⁴
ENSG00000174437	ATP2A2	12	PGC, 0-500	0.99	6.02	6.05	6.22			8/11 (0.73)	2 × 10 ⁻⁴
ENSG00000175931	UBE2O	17		0.99	2.83	0.99	0.79		[3]	8/11 (0.73)	2 × 10 ⁻⁴
ENSG00000184156	KCNQ3	8		0.98	4.64	3.22	1.98		[3]	8/11 (0.73)	1 × 10 ⁻⁴
ENSG00000198929	NOSTAP	1		0.12	0.28	1.01	0.67			8/11 (0.73)	2 × 10 ⁻⁴
ENSG00000084731	KIF3C	2		0.91	1.31	3.24	1.9			7/10 (0.70)	4 × 10 ⁻⁴
ENSG00000130226	DPP6	7		0.97	2.64	1.94	3.38		[3,4]	7/10 (0.70)	3 × 10 ⁻⁴
ENSG00000139767	SRRM4	12		0.41	2.72	2.91	1.63		[3,4]	7/10 (0.70)	2 × 10 ⁻⁴
ENSG00000171435	KSR2	12		0.99	2.62	4.23	4.31		[3]	7/10 (0.70)	2 × 10 ⁻⁴
ENSG00000178235	SLITRK1	13		0.99	0.24	0.92	2.54			7/10 (0.70)	4 × 10 ⁻⁴
ENSG00000183715	OPCML	11	PGC, 0-500	0.58	6.3	2.78	0.03		[2,3,4]	7/10 (0.70)	3 × 10 ⁻⁴
Postnatal consensus											
ENSG00000241973	PI4KA	22		0.0008	1.76	1.53	2.32		[1] [-]	7/10 (0.70)	2 × 10 ⁻⁴
ENSG00000110427	KIAA1549L	11		0.0037	0.8	1.4	-0.92			6/9 (0.67)	0.0012
ENSG00000133958	UNC79	14		0.99	3.52	3.63	3.08			6/9 (0.67)	9 × 10 ⁻⁴
ENSG00000139182	CLSTN3	12		0.99	2.35	0.7	1.87			6/9 (0.67)	0.0014
ENSG00000144406	UNC80	2		0.14	-0.65	-0.14	0.05			6/9 (0.67)	6 × 10 ⁻⁴
ENSG00000145934	TENIM2	5		0.99	3.41	1.63	1.5			6/9 (0.67)	8 × 10 ⁻⁴
ENSG00000147724	FAM135B	8		0.99	2.66	4.81	4.72			6/9 (0.67)	4 × 10 ⁻⁴
ENSG00000155093	PTPRN2	7		0.017	2.41	0.64	0.17		[2,3,4,6]	6/9 (0.67)	0.0011
ENSG00000242732	RGAG4	X		0.015	NA	NA	NA			6/9 (0.67)	0.001

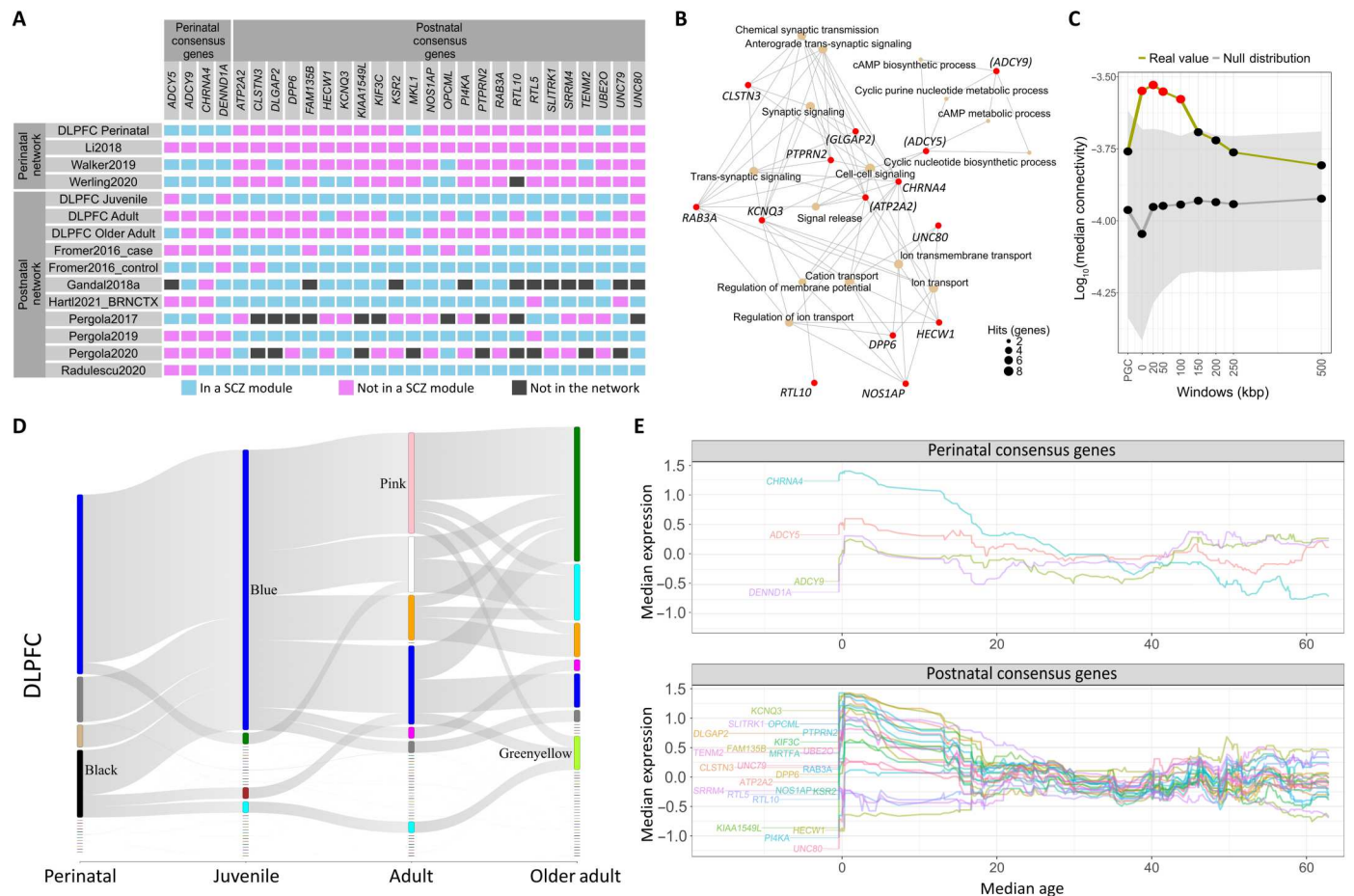


Fig. 9. Consensus genes and their relationship with PGC3 prioritized genes. (A) Overlap of four perinatal and 24 postnatal consensus genes in the prefrontal networks according to their membership to a SCZ risk module across at least two-thirds of 11 networks. Cyan boxes represent genes in SCZ risk modules. Magenta boxes represent membership to modules not significantly enriched for SCZ. Dark gray boxes represent genes not annotated in that specific network. The figure illustrates how coexpression partners of SCZ risk genes change from perinatal to postnatal stages. (B) Gene-Concept Network (CNet) plot of the top 15 most significant ($P_{\text{adjust}} < 0.05$) associated GO biological processes to postnatal ($n = 24$) and perinatal ($n = 4$) consensus genes. Nodes for processes are colored beige. Node size represents the number of hits. Nodes for genes are colored red. Genes also present in the GWAS are marked with round brackets. Only genes that are involved with the top 15 GO biological processes are shown (14 of 28 genes). (C) Connectivity between SCZ risk genes defined at various windows of genomic extension and 22 non-GWAS postnatal consensus genes in an iPSC dataset. Plots show $-\log_{10}(\text{median gene-wise connectivity})$ (y axis) of consensus genes to each SCZ gene set (x axis) from an iPSC-derived gene coexpression network. A null distribution of $-\log_{10}(\text{median gene-wise connectivity})$ was computed from randomly permuted non-SCZ and nonconsensus genes, after matching for gene set size, GC content, gene length, and average expression (95% CI for the mean in shaded gray). (D) Sankey plot of consensus gene membership across age stages in the DLPFC. Thin nodes not connected by links represent modules containing 200-kbp SCZ risk genes but not consensus genes. (E) Median rank-normalized expression across sliding window samples by age for the four prenatal (top) and 24 postnatal (bottom) consensus genes in the NC DLPFC.

brain development. We analyzed RNA-seq from neurons derived from human iPSCs designed to have a wide separation of individual genetic risk for patients with SCZ compared with control samples (33). We found that 22 of the 23 non-GWAS consensus genes reported in Table 1 were also expressed in iPSC-derived neurons. We verified whether sets of putative SCZ risk genes (defined from PGC3 and at varying genomic windows from PGC3 SNPs) were highly connected with consensus genes (not including those already labeled as SCZ risk genes). We found that, even in these stem cell-derived neurons, the preferential connectivity pattern of consensus genes with putative SCZ risk genes is preserved and remains significantly higher than background genes matched for length, GC content, and average expression (Fig. 9C).

DISCUSSION

We investigated the convergence of genetic risk for SCZ into coexpression networks at different periods of the neurotypical life span across different brain regions to address how SCZ risk genes change their coexpression partners across age and to identify a molecular environment that might be particularly relevant to enabling the pathogenic effects of genetic risk. The evaluation of GWAS top loci and whole-genome risk yielded consistent results regarding a profile of convergence of SCZ risk on early development coexpression patterns in the DLPFC, followed by the juvenile hippocampus and caudate nucleus convergence. The key findings of this report are as follows:

1) Parsing by age periods explained more variance in a score of gene importance for SCZ compared to lumping all age periods together in a single network.

2) SCZ risk genes tend to be more coexpressed than random genes matched for potential confounders, and cluster into specific modules, especially in early-stage DLPFC. The finding that perinatal and juvenile networks showed greater maximum overrepresentation relative to adult networks suggests that SCZ risk genes cluster best into coexpression networks in younger neurotypical brain tissue.

3) SCZ risk genes tend to continue to be clustered together from perinatal to juvenile life in DLPFC, but most of their coexpression partners change. This finding was replicated when using exclusively prenatal samples and was further replicated in external datasets, suggesting that these SCZ risk genes continue to exert combined effects while part of their molecular environment changes across development, potentially underlying the age-associated shifts in clinical symptomatology.

4) Networks derived from the caudate nucleus of patients with SCZ show greater convergence of SCZ risk genes than networks from age-matched NC. This effect is not explained by cell type abundance inferred via deconvolution.

5) Despite age changes in coexpression, a set of 28 genes is reliably found in modules enriched for SCZ risk genes in the DLPFC across multiple publications, yet 82% of these genes are not in GWAS-significant SCZ loci. These non-GWAS enrichment-related genes are at least as intolerant to loss-of-function mutations as GWAS genes and significantly more so than the background and may represent a consistent molecular environment that mediates SCZ risk.

6) Most of these consensus genes are highly expressed in perinatal life and are also expressed and correlated with SCZ risk genes in iPSC-derived human neurons. This evidence suggests an *in vitro* platform for experimental investigation of their role in enabling SCZ biological risk.

Together, these findings suggest that mRNA coexpression shows a convergence of SCZ genetic risk situated in the context of age period and brain regions specificity, revealing potentially important biological signal beyond what has been identified with GWAS.

SCZ gene coexpression changes across the neurotypical life span

The aim to identify when and where in the brain SCZ risk genes converge is based on the putative role of neurodevelopment in SCZ. These results emphasize early convergence of SCZ risk and potentially shed light on its shifting clinical presentation (42–44). Recent evidence supports this view. SCZ candidate risk genes are expressed earlier in humans than nonhuman primates, especially during prenatal development (45). Early life gene expression and chromatin accessibility are associated with SCZ genetic risk (26, 46, 47); SCZ genetic risk shows higher penetrance in the presence of perinatal environmental risk (48); combined GWAS/chromatin contact analyses support early life genetic risk convergence for psychiatric disorders (28). Most recently, Cameron *et al.* (24) have shown that genes with high PGC3 associations, as estimated via MAGMA, are preferentially expressed in developing neurons of the frontal cortex during the second trimester of gestation.

Our data suggest that assessing coexpression changes across the life span identifies brain region specificity, which is not equally

accessible with “nonparsed” coexpression approaches. For example, SCZ risk modules enriched for the targetomes of *SP4*, a transcription factor prioritized in genetic associations with SCZ (2, 40), are found in the perinatal and juvenile DLPFC but not in any nonparsed network (Fig. 6A). Consistent with the privileged role of the DLPFC in genetic risk for SCZ, we found that in this region and not in the hippocampus, SCZ risk genes in perinatal life were still coexpressed in juvenile individuals, while changing their coexpression environment partners. SCZ risk genes were peripheral in hippocampal perinatal networks, and the enriched modules overrepresented risk for other disorders and not just SCZ, effects not observed in early-stage DLPFC. Note that our benchmark networks identified using the same samples but without parsing did incorporate age information. As in prior studies, age could be traced back to the first principal component of module gene expression (27, 31). However, the aggregation of genes in prior studies and in our benchmark networks was likely driven by gene expression changes during development, without the opportunity to identify how coexpression partners change during the life span. In particular, the neurotypical function of gene modules associated with SCZ appears more fundamental during perinatal and early developmental periods. As most of the evidence currently available about SCZ genetic risk in coexpression networks derives from adults, the developmental perspective that we provide here has likely been missed in previous work. The network resource developed in this work is freely available and can be used for further studies.

Several of our findings are relevant to a system neuroscience approach to SCZ. The privileged position of the DLPFC in the functional genomics of SCZ agrees with many prior postmortem DLPFC and imaging studies of living individuals (49–51). As previous studies established for gene expression (26, 27, 29, 30), also for coexpression, there are differences between prenatal and postnatal DLPFC. Our results support this idea as well, but to a lesser degree, for hippocampus. However, we also found that RNA-seq from the dentate gyrus of the hippocampus showed greater coexpression of SCZ risk genes than bulk tissue and peak coexpression in NC in the sliding window analysis, supporting the relevance of SCZ risk gene convergence for neurons and suggesting cell type beyond brain region specificity. A convergence of SCZ risk genes on neuronal and especially synaptic biology is the same conclusion reached in the PGC3 report (2). To the extent that coexpression recapitulates cell types, cellular districts, and functional pathways, the genes most often coexpressed with SCZ risk genes should share their main functions and, in fact, do so. The additional enrichment analyses also highlight spatiotemporal aspects of SCZ risk gene coexpression. For instance, differentially methylated CpG islands (DMGs) are coexpressed principally in perinatal or juvenile networks; and differentially expressed genes (DEGs) are coexpressed in the caudate nucleus. The caudate nucleus SCZ enrichment time course aligns with dopaminergic maturation (37, 44). In the sliding window analysis, patients with SCZ generally showed greater enrichment for SCZ risk genes in caudate nucleus networks, the brain region with maximum dopaminergic innervation. This finding may be associated with D_2 receptor-targeting antipsychotic treatment in patients with SCZ or with the impact of variants in genes relevant to dopaminergic transmission (52). Notably, our results based on a connectivity match between networks across brain regions and age failed to support previous findings of less connected

coexpression networks in the DLPFC of patients with SCZ (5). We hypothesize that the networks available with the present work may afford greater explanatory power in the translation of postmortem brain insights into neuroscience and clinical predictions in living patients with the latest available approaches (13–15, 53–60).

The developmental milieu of SCZ risk genes

Arguably, our most unexpected and novel result is the identification of a set of mutation intolerant consensus genes that are especially likely to be coexpressed with SCZ risk genes. This set of genes is coexpressed with SCZ genes in the large majority of the networks that we examined, regardless of often emphasized differences across datasets, preprocessing pipelines, and parameter setting. The method that we used to derive this consensus is very conservative, as confirmed by the statistical significance of finding very large consensus sets. The gene-wise *P* values that we derived control for bias related to module size of the consensus genes and are highly significant for all genes.

We successfully predicted on the basis of the postmortem data that these genes would be more coexpressed than matched background genes with the set of 120 PGC prioritized genes in iPSC-derived human neurons. There is no obvious clustering of these genes by chromosome, hence discounting the possibility of locus-related artifacts. By merging modules containing consensus genes, we showed strong enrichment for SCZ risk genes, which was not found when merging modules without consensus genes. Gene ontologies for the consensus gene set and their modules once again emphasized neurons, synapses, and ion transport. It is tempting to speculate that these genes mediate at least in part the read-out of SCZ risk at the level of biological systems and that they are involved in translating SCZ genetic risk into SCZ biological risk.

There are multiple nonexclusive explanations as to why genetic variants relevant to the biology of a heritable disease might not show GWAS-associated signal, which may be related to their high loss-of-function intolerance. For example, they might be depleted in common cis-functional polymorphisms; their common variants may be associated with other disorders that represent exclusion criteria for SCZ genetic studies, e.g., other central nervous system disorders; their effect may be nonadditive, for instance, because of epistasis with other variants masking their effect in GWAS, or their expression may be principally regulated by trans-eQTLs. Gene coexpression provides a handle on the role of potentially relevant genes undetected by GWAS, thus highlighting previously unreported targets “guilty by association” and related regulatory elements. These non-GWAS genes thus may be part of the cellular environment in which GWAS-positive genes exert their pathogenic function, potentially accounting at least in part for how diverse GWAS risk genes of small clinical effect converge on a common phenotype that shows varying expressivity over the life span.

In terms of limitations of our methods, although we analyzed the largest developmental series available to date, the absolute cohort sizes are still limited, especially for perinatal and juvenile networks. We endeavored to mitigate this shortcoming by including a minimum of 40 samples per network; however, this choice came at the cost of diluting biological specificity associated with age by extending the age intervals beyond the desirable limits, e.g., up to 6 years of age in the perinatal window. The networks identified also depend on the nonuniform age sampling with fewer samples

between 1 and 20 years. For these reasons, we confirmed findings pertaining to perinatal and juvenile windows in independent RNA-seq data and also included a sliding window approach, which supported the conclusions of the fixed age period study.

A relevant methodological aspect of this work is that we obtained networks comparable across the life span via statistical corrections for cell type abundance and by matching connectivity between different brain regions and age periods. These operations reduce the possibility that these potential confounding variables drive the findings. Cell specificity enrichments computed on SCZ-enriched modules showed that removing interindividual variation in cell abundance estimates did not compromise module specificity, which appears equal to or greater than those of published reports. Despite our efforts, some confounding effects may still affect the results, particularly regarding ancestry and sex. Statistical correction of the effects of these variables on expression data cannot address the potential different genetic architecture underlying the expression of SCZ risk genes across ancestry groups and sex. The population in our cohort differed from the reference PGC3 GWAS that we used to identify SCZ risk genes. It is thus possible that these variables still affect our findings, although no evidence has been reported of differential genome-wide association with SCZ in males and females (2).

In conclusion, while the convergence of SCZ risk genes into coexpression pathways has been observed before, here, we identify spatiotemporal dimensions of this convergence such that the relationships between genes change along the life span, especially between perinatal and postnatal life, perhaps in parallel to the seminal transition from the preclinical to the manifest illness. Age-parsed networks afford superior predictions in terms of potential gene association with SCZ. The granular quantification of SCZ risk gene convergence across age reveals a difference between neurotypical adults and patients with SCZ specifically in caudate nucleus data. Most notably, SCZ risk genes converge especially in perinatal and juvenile DLPFC, as shown by enrichment and prediction analyses, and keep being coexpressed in a shifting context. Together, these findings highlight that the genetic architecture of SCZ is embedded in shifting coexpression patterns across brain regions over the life span. Leveraging this shifting pattern of coexpression, we identified novel associations with SCZ in a set of genes consistently coexpressed with GWAS-positive genes that may be part of the cellular environment that mediates the translation of genetic risk into biological risk for SCZ, an environment that can be readily studied in an *in vitro* model to investigate the potential contribution of these genes for the development of future treatments.

MATERIALS AND METHODS

Subjects and materials

This study was based on brain tissue of NC and individuals diagnosed with SCZ of European or African American ancestry, all with RNA integrity number (RIN) ≥ 6 (table S1). All tissue donations were made with informed consent from next of kin. A number of brains in the LIBD Human Brain Repository were transferred from the National Institute of Mental Health (NIMH) Clinical Brain Disorders Branch under a material transfer agreement, after having been collected under NIMH Protocol 90-M-0142 and processed and approved by the NIMH/National Institutes of Health Institutional Review Board. Additional cases in the LIBD repository

were collected from the Office of the Chief Medical Examiner for the State of Maryland under State of Maryland Department of Health and Mental Hygiene Protocol 12-24 and from the Kalamazoo County Michigan Medical Examiners' Office under Western Institutional Review Board Protocol 20111080. All samples were collected and processed using a standardized protocol specifically developed to minimize sample heterogeneity and technical artifacts. All included NC subjects had minimal age-associated neuropathology (determined from postmortem histopathological examination) and no substance or drug use from toxicology and were free from any psychiatric or neurological disorder from clinical histories. Postmortem clinical information was gathered by conducting family interviews with the next of kin, as previously described (61). After psychiatric record reviews and postmortem family interviews were completed, brief psychiatric narratives were prepared on each case, summarizing the demographic, clinical, medical, and death information obtained from as many sources as possible (i.e., multiple psychiatric records, police reports, neuropathology reports, medical examiner's information, toxicology screen, and postmortem family interview). Each case was then independently reviewed by two board-certified psychiatrists, who arrived at consensus DSM-IV Axis I lifetime diagnoses or consulted with a third reviewer to reach a final diagnosis.

DLPFC samples were derived from Brodmann area 9/46 at the level of the rostrum of corpus callosum. Hippocampus samples were derived from the mid-hippocampus proper (all dissections included the dentate gyrus, CA3, CA2, and CA1) plus the subicular complex (61). For caudate nucleus, samples were taken from the anterior "head" portion, which is the caudate part most tightly connected with the prefrontal cortex. Dentate gyrus samples were obtained from the granule cell layer of the dentate gyrus (*DG-GCL*). *DG-GCL* samples were isolated from neighboring polymorphic and molecular layers by using laser capture microdissection (23). For all tissues, RNA-seq was performed via the Illumina Ribo-Zero Kit. In the Age study, we determined the age periods as a trade-off between achieving a minimum sample size for each age period to identify networks (minimum of 40 subjects) and the identification of biologically meaningful age windows.

Replication datasets. The replication study was carried out using the (i) UCLA-ASD data from the PsychENCODE collection (36); here, we selected 16 NC ranging in age from 6 to 25 years; (ii) BrainSpan data from 21 perinatal and 9 juvenile samples; and (iii) PolyA data from the LIBD repository, not overlapping the LIBD samples already used in the main analysis, from 29 perinatal and 19 juvenile individuals.

Quantification and statistical analysis

Data preprocessing

All downstream analyses have been performed primarily using the R statistical software (version 4.0.0+). Specific R packages used are also listed in the Key Resource Table in the Supplementary Materials.

Regional coexpression and age studies

We performed a common preprocessing for non-age-parsed tissue samples to ensure the use of the same samples and settings across ages (the only difference is, thus, the subsequent age parsing). Computing preprocessing parameters on the entire cohorts also allowed the best-informed estimation of confounding effects on a large sample. Specific challenges in this approach concern the

preservation of age effects during preprocessing, the comparability between the networks obtained, and the variation of cell population abundance between early development and adulthood.

Gene-level mRNA expression was quantified as reads per kilobase per million mapped reads (RPKMs) and annotated as total gene expression separately for each brain region, regardless of alternative transcript quantification (61) using GENCODE release 25 (GRCh38.p7). We selected genes above median RPKM ≥ 0.1 and free of floor effects (maximum 20% of zeroes per gene); we log-transformed RPKM values with an offset of 1, i.e., $\log_2(\text{RPKM} + 1)$. We used inter-array distance to identify outlier subjects deviating more than 3 SDs from the mean (14). Outliers were identified relative to their age period in the fixed period analysis. Outliers identified in this step were also excluded from the subsequent sliding windows analysis. After removing mitochondrial genes that were highly expressed in some samples, datasets included a variable number of genes for different regions (table S1).

Comparing networks across different age periods requires the preservation of the biological effects of age while removing unwanted confounding effects. Standard noise removal techniques, such as principal components analysis, are effective at modeling noise into latent factors while preserving biological effects (62, 63). However, these latent variables were strongly associated with age in our samples (>80% shared variance), likely because of the inclusion of prenatal samples. This finding extended to latent variables derived from experimental quantification of degradation effects conducted on adult samples, i.e., qSVA (quantitative surrogate variable analysis) (64). To reduce the risk of removing age-related signal in the Age study, we used correction for explicit confounders, a practice that has been suggested as optimal for gene coexpression networks identified in brain tissue (11). In particular, we regressed out the effect of the following variables from each brain region cohort separately on the basis of availability: Confounding variables regressed out for BrainSpan were sex, self-reported ancestry, and RIN. Confounding variables regressed out for UCLA-ASD were sex and RIN. Confounders used for PolyA were the same as those used for the primary LIBD networks, namely, sex, mitochondrial mapping rate, ribosomal RNA (rRNA) rate, gene mapping rate, RIN, ancestry estimated via the first 10 genomic eigenvariates [computed as in the work of Collado-Torres *et al.* (61)], and estimated individual neuronal proportion; the latter confounder was included to obtain an appropriate comparison between subjects of different age. Individual-level estimated neuronal proportion was computed using the R package BRETIGEA (species: human, scaling: FALSE, dataset: kelley) (65) and regressed out because it could bias estimated gene-gene connectivity due to interindividual variation in cell abundance across age periods. We estimated and protected linear and quadratic terms of age using the function `cleaningY` from the `Jaffelab` package (66). We rank-normalized residuals using Blom formula (13, 54, 56, 58) to limit the impact of deviations from normality in expression data (67).

Network identification across all studies

Blom-normalized residuals obtained using the linear models described above were entered as input in the `blockwiseModules` function from the package `WGCNA` to construct "signed hybrid" network(s), i.e., negative correlation were set to zero and positive correlations were soft-thresholded. We obtained the similarity matrix using Pearson's R correlation index (`minModuleSize`: 40, `maxBlockSize`: 3000, `deepSplit`: 4, `mergeCutHeight`: 0.15, `pamStage`:

TRUE, reassignThreshold: 1e-06). The parameter used for soft-thresholding is the exponent β to which the correlation matrix is raised to obtain the adjacency matrix. The standard procedure is to pick the lowest possible β that is high enough to satisfy the scale invariance criterion, defined as the $R^2 > 0.8$ in the correlation between original and log-transformed values (34). Lower β values are often associated with greater network median connectivity. For each network, we selected the parameter β such that connectivity was matched across all networks (see below; table S3).

Exploratory analyses on our datasets revealed that perinatal networks yielded systematically greater connectivity than other age periods, thus undermining comparisons between brain regions and time windows. Outcome variables such as the number of non-clustered genes are influenced by β and by network connectivity; thus, the standard criterion to identify networks generated large differences between networks. To obtain comparable networks, we assessed the median network connectivity for each sample and selected β to meet the scale invariance requirement in all networks while also matching connectivity. To achieve this effect, we used larger β values for early, highly connected networks at a level of connectivity compatible with later life-span periods while still meeting scale invariance. Following this procedure, we obtained comparable graph structures with small variations in connectivity and in number of nonclustered genes (fig. S10).

For each replication dataset, we rank-transformed and normalized expression values to construct a network for each age group. For the construction of networks, genes were matched between all sets and with the main networks that we generated (n genes = 14,553). A soft power (β) was first identified for each such that the network was scale invariant (scale invariant index > 0.8) and the median connectivity is not greater than 1.123 (the median connectivity across LIBD network). The BrainSpan juvenile network did not match these criteria under $\beta = 30$ (the maximum allowed by the WGCNA package) and was thus not used further. The blockwiseConsensusModules function from the WGCNA package was used to detect consensus modules across perinatal networks polyA_Perinatal ($\beta = 17$) and Brainspan_Perinatal ($\beta = 29$) and across juvenile networks polyA_Juvenile ($\beta = 22$) and UCLA_Juvenile ($\beta = 25$).

Other previously published networks were used "as is," except for the network published by Pergola *et al.* (13). This network used BrainCloud, a microarray dataset featuring constitutive and alternative probes. Therefore, we reannotated it to obtain a total gene expression quantification and make it comparable with the other networks, which were all annotated to GENCODE release 25 (GRCh38.p7) to be consistent with the annotation database used in the network identification. For each probeID, first, the corresponding accession number was converted into Ensembl ID using clusterprofiler package (68). When an accession number mapped to multiple Ensembl IDs, a single Ensembl ID was randomly selected. If no Ensembl ID was returned via accession number, then the gene symbol corresponding to the probeID was used to derive an Ensembl ID. Only one probeID per Ensembl ID was retained in a preferential order of probe types: constitutive exonic (hHC), then alternative exonic (hHA), and then human mRNA (hHR).

Network characterization and association with SCZ risk

Definition of SCZ risk genes. We used the PGC3 definition of 120 priority genes as the list called "PGC." Then, to obtain genomic regions agnostic of PGC3 fine mapping and reflecting a consensus

between SCZ risk gene lists, we evaluated enrichment at multiple thresholds of genomic distance from the index PGC3 GWAS-significant SNP. We have previously tested a similar method in another dataset using PGC2 GWAS-significant loci (14). The rationale for this choice was to generate lists of genes in the proximity of significant SNPs that survived genome-wide correction for multiple comparisons, assuming that in these genetic regions, the probability of finding causative genes is highest. On the other hand, as we do not know whether the index SNP is causative and gene regulation is also possible at a certain distance from the genes, we evaluated increasing extension thresholds to include increasing numbers of genes. The number of genes included was still much smaller than all genes included in all loci identified by PGC3, which are often larger than 1 million base pairs (Mbp) and thus are proximal to thousands of genes (possibly with the dilution of the probability that each gene is causative). By assessing the consensus between gene lists, we attempted to use a conservative approach in which the choice of parameters does not drive results. We considered nine extension windows at increasing distances in both directions from each PGC3 GWAS significant SNP: PGC [120 genes], 0 kbp [178 genes] (meaning, genes that overlapped with the index SNP), 20 kbp [299 genes] [extension from the loci used in PGC2 (69)], 50 kbp [456 genes] [extension from the loci used in PGC3 (2)], 100 kbp [705 genes], 150 kbp [963 genes], 200 kbp [1196 genes], 250 kbp [1436 genes], 500 kbp [2475 genes] [500 kbp was the maximum extension where enrichments were found significant by (14)].

The ancestry composition of our expression data included European and African American individuals, hence different from the PGC core cohort (130,644 European and 30,761 Asian ancestry individuals), as well as from the PGC primary cohort including European, Asian, African, and Latino ancestries (2). The addition of African and Latino ancestry in the PGC primary results has modest effects on the overall summary statistics because the sample size is below 10% of the PGC core dataset but has a sizable effect on the clumping, arguably because ancestry in the genome largely takes the form of LD blocks. As the purpose of our analysis was only to identify index SNPs associated with SCZ and not for genetic trait associations, we based our definition of risk genes on the PGC core summary statistics (2). These summary statistics are well powered, genetically more homogeneous than the PGC primary results, and largely consistent with previous SCZ-GWAS. It is also noteworthy that the sex distribution in our expression data privileged males over females, representing another difference from the PGC core cohort (female-to-male ratio = 0.81), in which the overrepresentation of males was less overwhelming. A chi-square test on the proportion of males and females returned a highly significant difference [χ^2 (df = 1, $N = 164,061$) = 34, $P = 5.5 \times 10^{-9}$].

We generated negative control gene sets equal in size to the ones described above used in the enrichment analysis by selecting the 500 least significant SNPs in the PGC3 core summary statistics and defining negative loci at the above-considered distances from each tagging negative SNP. We excluded any region located at ± 3 Mbp from SCZ SNPs from this selection to rule out any overlaps between positive and negative gene sets.

Coexpression of SCZ risk genes regardless of module membership

To test the hypothesis that putative SCZ risk genes are more highly coexpressed than expected by chance, we took as connectivity

measure the median of the gene-wise mean connectivity with all other SCZ risk genes from the adjacency matrix of each network identified and compared it to a null distribution of randomly permuted non-SCZ genes.

As GC gene content and gene length may bias random permutations [SCZ risk genes are GC richer and longer than random genes (2)] together with gene expression within each network, we ran 10,000 permutations with the restriction to match both gene set size and GC content, besides gene length and average expression distributions of the SCZ risk genes. The universe from which random genes were pooled consisted of genes expressed in each network, excluding SCZ risk genes (the most extensive list that we generated at 500 kbp from GWAS-significant SNPs was considered for this analysis). Last, we defined the empirical *P* value as the number of occurrences in which the connectivity measure of the random genes exceeded the one of the SCZ risk genes divided by the number of permutations. We set the significance threshold at an empirical *P* value < 0.001. We repeated this procedure for the nine extension windows previously created.

MAGMA

To calculate MAGMA and H-MAGMA, we used the MAGMA tool v1.09b (39), PGC3 summary statistics as SNP *P* value data (2) and 1000G European as the reference data file for a European ancestry population to estimate linkage disequilibrium between SNPs. We took the following steps: (i) We mapped 1000G SNPs to genes encompassed in each network module [a window of 35 kb upstream and 10 kb downstream of each gene; for H-MAGMA, we used adult brain Hi-C annotation files already computed in the H-MAGMA publication (28)], (ii) calculated gene *P* values on the basis of PGC3 SNPs *P* values, and (iii) performed gene set enrichment analysis using the modules detected in the network that assigns a gene set *P* value (universe used by MAGMA consists of all network genes with at least one SNP mapped). We computed this analysis for each network for both all biotypes and protein-coding genes and corrected for the number of modules in each network using Bonferroni's rule.

Enrichment methods

We identified modules in each network. Then, we assessed the overrepresentation of SCZ risk genes for each module except gray in each network relative to a universe composed of all non-gray genes. We corrected results for multiple comparisons via Bonferroni (number of non-gray modules in each network). This statistic is independent of the underrepresentation in the gray module and represents the competitive enrichment of one module relative to the background. We labeled as SCZ risk modules those significantly enriched in at least three of the nine extension windows (hypergeometric test, Bonferroni-corrected *P* value < 0.05). In addition to hypergeometric tests, we also computed permutation statistics (creating for each module in each network a null distribution of 10,000 sets of randomly sampled genes using network-specific genes as the universe) to obtain an unbiased empirical *P* value. We performed these same analyses in networks from published reports to provide a benchmark (5, 10–14, 26, 27, 31, 35, 36). Moreover, we tested whether SCZ risk genes were significantly underrepresented in the gray module of nonclustered genes relative to a universe represented by all genes expressed in a network. Significant underrepresentation suggests that SCZ risk genes tend to be coexpressed rather than not (see results in fig. S2).

Using the same methods with other GWAS summary statistics, we computed gene enrichment for several psychiatric, neurologic, and immune disorders [attention deficit hyperactivity disorder (70), autism spectrum disorder (71), bipolar disorder (72), major depressive disorder (73), obsessive compulsive disorder (74), post-traumatic stress disorder (75), suicide attempt (76), Alzheimer's disease (77), amyotrophic lateral sclerosis (78), multiple sclerosis (79), Parkinson's disease (80), Crohn's disease and ulcerative colitis (81), and rheumatoid arthritis (82)].

To mitigate the risk of biased enrichments due to the inclusion of many genes mapping onto few loci, which may depend on RNA-seq read annotation for transcripts overlapping in the DNA sequence or due to common regulatory sequence elements between close genes, we also performed a permutation test of loci overrepresentation. In this test, all genes within a module located in the same locus (defined according to the genes located in the proximity of GWAS-significant SNPs) were counted as a single hit.

For the cell type enrichment analysis, we used gene sets reported by Skene and colleagues (83) including cell type specificity indices available at www.hjerling-leffler-lab.org/data/scz_singlecell/. They computed specificity indices for each gene and range between 1 (high specificity for a given cell type) and 0 (low specificity). We used specificity indices derived from single-nuclei RNA-seq of human brains [DroNc-seq in prefrontal cortex and hippocampus (84)]. We used mean-rank gene set test in the limma R package (85) to evaluate the enrichment of our components for the specificity indices of each cell type.

We characterized the modules of interest using the Gene Ontology Database (PANTHER) (86) and clusterProfiler package (annotation package: org.Hs.eg.db 3.16). We also computed enrichments for cell specificity (83); transcriptome-wide association study (TWAS) genes obtained from caudate nucleus (52), hippocampus (61), and DLPFC data (7, 61, 87); DEGs obtained from caudate nucleus (52), hippocampus, and DLPFC data (61); DEGs in each brain region (5, 6, 52, 61, 88), genes proximal to differentially methylated CpG islands (DMGs) in PFC, and blood (89–94); DEGs in humans relative to great apes (88); SCZ drug target genes (95–98); and loss-of-function intolerant genes (99) in the selected modules. For TWAS and DEGs, we performed a brain region-specific enrichment using the appropriate gene list of each tissue. The "universe" of TWAS enrichment was adjusted by intersecting the list of genes associated with a cis-eQTL, called heritable genes, with the brain region-specific list of expressed genes. Moreover, for DLPFC TWAS and DMGs, as well as for drug target enrichment, we computed a meta-analysis of the papers that we retrieved target genes from to obtain module-wise enrichment *P* values (sum-log Fisher's method, agnostic to developmental stage).

MAGMA linear models

To reduce biases associated with genes uniquely expressed in a single brain region or age period, in this analysis, we only considered autosomal genes included in all Age study networks, i.e., the intersection of genes expressed in the DLPFC, hippocampus, and caudate nucleus. In particular, we performed the gene analysis using MAGMA annotation files (35-kb upstream and 10-kb downstream window was used), PGC3 summary statistics as SNP *P* value data, and 1000G European as the reference data file for a European ancestry population, as reported above. Thus, we considered the gene-wise *z* scores obtained in this analysis as the MAGMA importance score (i.e., the measure of gene annotation to SCZ) and

assessed the association of coexpression in different brain regions and age periods with gene importance for SCZ derived by MAGMA. Because H-MAGMA Hi-C annotation was performed by authors of the paper using hg19/GRCh37 as genomic reference, we only kept SNP and genes locations consistent with H-MAGMA reference when generating H-MAGMA scores for the consensus genes.

We computed linear models to associate this score with coexpression. In these analyses, the outcome variables were represented by the MAGMA importance score obtained for each gene, previously marginalized by gene position, number of transcripts, GC content, and gene. We assessed the improvement in model fit deriving from the inclusion of coexpression variables in the model relative to a null model only, including confounders, via maximum likelihood estimation. This analysis implied an analysis of variance (ANOVA) comparing the null model with the test model. We then tested the following hypotheses: (i) Age-parsed networks explain more variance in gene importance for SCZ than using the same samples to compute a single network per tissue; we tested this hypothesis with the Vuong test for non-nested linear models (comparing full sample versus age-parsed networks). (ii) SCZ risk convergence is region and time specific, i.e., networks computed at critical age periods are reproducibly related to SCZ risk importance.

We considered as confounders chromosome location, gene start (to index location within the chromosome), gene width, GC content, strand, number of isoforms, and gene type. These variables are potentially related to SCZ risk importance, but they do not necessarily reflect coexpression-related signal, which is why their biological association with SCZ is considered confounding here. The poorly represented gene types in our dataset were excluded to avoid biasing the models; the minimum WGCNA module size served as the threshold for exclusion (40). We thus excluded genes annotated as 3prime_overlapping_ncRNA (4), polymorphic_pseudogene (3), IG_V_gene (2), non_coding (1), rybozyme (2), scaRNA (15), scRNA (1), TR_V_pseudogene (1), transcribed_unitary_pseudogene (15), and unitary_pseudogene (15). The remaining gene types had over 41 entries each. The dataset included 17,996 genes.

To test the second hypothesis, we compared prediction performances for the same machine learning workflow on parsed and nonparsed datasets. The two datasets differed in expression attributes and coexpression properties, specifically median gene expression (3 features in nonparsed and 11 in age-parsed) and kTotal (connectivity of a gene to all other genes in the network) and KME (module membership of a gene) connectivity variables for the modules of the coexpression networks considered (3 kTotal + 146 KME features in nonparsed and 11 kTotal + 409 KME features in age-parsed). The workflow of the analysis is described in Fig. 4A and in detail in the methods S2.2. Briefly, we used a cross-validation approach stratified by chromosome location to identify statistically important features for each gene using Boruta's algorithm (100). Such features were then used to train an XGBoost regressor and evaluate its performance using R^2 . Reiterating the pipeline for 200 runs of fivefold stratified cross-validation resulted in a distribution of 1000 R^2 values (Fig. 4B). Figure 4 (C and D) shows the features selected by Boruta in the absolute majority of runs for the two datasets. The dataset included 21,751 genes because the algorithm was able to handle missing entries.

SCZ risk gene coexpression within-set connectivity

We subset the original adjacency matrix for each SCZ risk gene set to include only the in-set genes. For each gene in a set, we then computed the mean connectivity (connectivity having been defined as gene coexpression raised to an exponent β) to all other in-set genes. Last, the median of the gene-wise mean connectivity was computed for each set, as a measure of set-wise within set gene connectivity/coexpression. Connectivity values were compared to a null distribution of gene sets of equal set size, gene length, GC content, and average expression.

Risk gene flow

We constructed Sankey plots illustrating the flow of risk across ages. We considered the extension threshold of 200 kbp because it captured most of the enriched modules compared to all other extension windows (Fig. 5). We studied the composition of gray modules of nonclustered genes and of SCZ risk modules obtained considering all biotypes, not just protein-coding genes, to maintain the networks comparable with each other without using multiple criteria (DLPFC perinatal black, juvenile blue, adult pink, older adult green-yellow; hippocampal perinatal red, juvenile turquoise, and adult turquoise and brown). For additional results about the composition and ontologies of risk gene flow, see the Supplementary Materials (note S1.5).

Jl analyses

In the analysis of module continuity across age periods, JIs were computed as the intersection/union of the considered sets. To draw statistical inferences on the dissimilarity between gray modules across age periods, we pooled all genes that at least in one of the age-parsed networks belonged to the gray module and resampled them 10,000 times to create a null distribution for each network. Then, using the earlier network as a reference, we computed a distribution of JIs with the null distribution and an empirical P value. Figure 7C thus represents the null distribution of JIs for each transition, holding the earlier network as a reference.

Sliding windows

In the work outlined above, we examined age periods that necessarily include quite diverse biological processes. For example, the perinatal window ranged up to 6 years of age, thus reflecting different biology compared with fetal samples (although only a handful of samples were available between 0 and 6 years). In addition, keeping age windows fixed does not allow us to estimate to what extent results are an attribute of the age period considered and not an attribute of the specific samples analyzed or rather of sample size available. Furthermore, in the context of the age study, a comparison between patients and controls cannot be made because diagnosed patients are adults. Study 2 (sliding windows) was therefore aimed at addressing these points in two main steps. The first consisted in ruling out that the findings about convergence of SCZ risk genes across the life span are related to the specific samples analyzed. The second was to obtain a comparison between patients and controls. We identified networks in chronological sequential windows of 40 subjects to control for these potential confounding effects. This way, we could observe the continuous variation of parameters of interest (like fold-enrichment mean and variance and others) across different tissues. For this analysis, we added patients with SCZ to the previously used neurotypical individuals and dentate gyrus data (table S1).

We first compared neuronal proportion estimates between patients with SCZ and controls. Neuronal proportion was estimated using the BRETIGEA package using the BRETIGEA::brainCells function [BRETIGEA parameters: nmarkers = 50 (default) and method = "svd" (default), species = "human," scaling = "F"] and then marginalized by total assigned genes, RNA_rate, and mitoRate. Residualized neuronal proportions were compared between age-matched diagnoses for each region using two-sided *t* tests. Neuronal proportion estimates were significantly lower in patients with SCZ than in controls across all regions except the caudate (CN $t_{264} = -1.07$, $P = 0.284$; DG $t_{132} = 4.27$, $P = 0.000038$; DLPFC $t_{287} = 3.74$, $P = 0.00022$; HP $t_{243} = 3.53$, $P = 0.00049$).

To compute sliding window networks, we sorted subjects by increasing age for each tissue and identified the first network with the 1st to 40th subjects, the second with the 2nd to 41st, and so on. We examined primarily the extension window at 200 kbp that in the age period analyses provided the strongest convergence (Fig. 5). For plotting, we used the ggplot2 package in R with smoothing method "loess" and span parameter 0.5, i.e., using 50% of the points to fit the local regressions (for dentate, the span parameter was 0.65 to accommodate for the lower number of windows compared to other tissues). For caudate NC windows, the first window is not plotted, as it is far from the second window in terms of median age.

Gene-level mRNA expression was quantified as RPKMs and annotated as total gene expression separately for each brain region, regardless of alternative transcript quantification (61) using GENCODE release 25 (GRCh38.p7). We selected genes above median RPKM ≥ 0.1 and free of floor effects (maximum 20% of zeroes per gene) and log-transformed RPKM values with an offset of 1, i.e., $\log_2(\text{RPKM} + 1)$. We used inter-array distance to identify outlier subjects deviating more than 3 SDs from the mean (14). Outliers were identified relative to their age period in the fixed period analysis. Mitochondrial genes were removed. In addition, we performed an intersample comparison on the rank-normalized expression data of each window to further identify potential outlier samples. Visual inspection of scale invariance index (R^2 of the correlation between raw and log-transformed connectivity values) and β values across windows revealed that specific samples caused disproportionate effects on these parameters, hence influencing power and connectivity. Only the SCZ group had samples with such an impact, and therefore, we removed three SCZ samples to reach the desired median connectivity with the minimum number of excluded individuals.

We also found that specific windows deviated from the proximal windows in the WGCNA parameter that we assessed. Note that consecutive sliding windows have 95% of the subjects in common; hence, there is a considerable redundancy to accommodate a slight decrease in the number of windows. For each window of 40 subjects, scale invariance index and median connectivity were estimated for β values ranging from 1 to 25. We selected β for each network such that the scale invariance index was greater than 80, and the connectivity was matched to the connectivity value (1.123) found in the age study. We tolerated connectivity values between one-third and three times the target connectivity. Windows where no β value returned connectivity values within these boundaries were removed (nine windows in NC [CN: 1, DLPFC: 3, HP: 5] and one in SCZ [HP: 1]). Thus, we could match NC and SCZ connectivity with the same value used in the

age study. Notice that there was no outlier sample removal in the NC group; therefore, the only difference between studies relates to the windowing procedure.

Moreover, we wanted to assess the association of SCZ with functional biology along sliding window networks. Therefore, we computed an across module association of cell type and gene ontology with MAGMA enrichment for each network. The mean rank gene set test-based geneSetTest (alternative = "up", type = "t", ranks.only = T) function from the limma package was used to derive *P* values for cell type and MAGMA enrichment for all modules in all sliding window networks. For each sliding window network, the rlm function was used to predict MAGMA enrichment $[-\log_{10}(p_{\text{MAGMAenr_module}})]$ from cell type enrichment $[-\log_{10}(p_{\text{CellTypeEnr_module}})]$, accounting for module size ($n_{\text{genes_module}}$) using robust linear modeling. See below

$$\begin{aligned} & \text{rlm}[-\log_{10}(p_{\text{MAGMAenr_module}})] \\ & \sim -\log_{10}(p_{\text{CellTypeEnr_module}}) + n_{\text{genes_module}} \end{aligned}$$

For GO term enrichment, the ratios of genes found in specific GO term lists were computed for all modules in all networks. As with cell type enrichment, this ratio (ratio_GOterm_module) was used to predict MAGMA enrichment using rlm, accounting for module size. See below

$$\begin{aligned} & \text{rlm}[-\log_{10}(p_{\text{MAGMAenr_module}})] \\ & \sim \text{ratio_GOterm_module} + n_{\text{genes_module}} \end{aligned}$$

The Z.stats of cell type or GO term enrichment for predicting MAGMA enrichment was extracted with their respective robust linear models and are depicted on the *y* axis of Fig. 8 (C and D).

Consensus

For each DLPFC or published network featuring SCZ risk modules according to the criteria that we defined, as shown in Figs. 2 and 5, we pooled SCZ risk genes from all SCZ risk modules. As our SCZ risk modules were identified using protein-coding gene lists and the report by Li *et al.* (31) had no module significantly enriched, we also considered for this analysis modules enriched for all biotype genes. For the negative lists, which were generated at 200 kbp from the negative SNP list, several networks had no significant enrichment; thus, we used the most significant module to generate the consensus, although it was not significant. Note that this procedure makes the test conservative, as it goes to the "advantage" of the negative list. We then intersected lists across networks to count how many times each gene appeared in an enriched module (minimum three modules). To compute the fraction of overlapping networks over the total for each gene, we divided the number of overlapping modules for each gene by the number of networks in which that gene was included: If a gene had not been annotated in a particular network, then that network did not count toward reaching this ratio. We only considered genes annotated in at least three networks (six for the postnatal consensus).

To obtain a *P* value for each consensus gene identified, we randomly permuted the network-specific SCZ risk modules and repeated the aforementioned procedure. The universe from which random genes were pooled consisted of genes expressed in the specific DLPFC or published networks. Last, we obtained a list of random consensus genes and defined the empirical *P* value as the

number of occurrences in which the real consensus gene appeared in the random consensus list divided by the number of permutations (10,000). We set the significance threshold at an empirical P value < 0.05 . We repeated this procedure for peri- and postnatal consensus lists separately.

Gene pLI scores were downloaded from the ExAC consortium release (99). A gene is generally considered likely to be loss-of-function intolerant if its pLI score is 0.9 or higher.

Regulomic analysis

For the regulomic analysis, we used the gProfileR package to search for transcription factors and miRNA whose potential targets were enriched in the postnatal consensus list. We restricted the universe of transcription factors to those expressed above the minimum threshold in at least six of the networks considered. We corrected results for multiple comparisons using false discovery rate.

iPSC data

The iPSC generation pipeline has been described in detail elsewhere (33). Ninety-five human neuronal samples with 56 to 70 days in vitro belonging to 26 male participants of European ancestry were available. The original report included 15 NC and 13 patients with SCZ. Of these subjects, RNA-seq data were available for 14 NC and 12 patients with SCZ. NC had a relatively low polygenic risk score for SCZ (bottom 37% of the European ancestry male distribution), and patients with SCZ had relatively high polygenic risk score (top 45% of the European ancestry male distribution). Using the criteria described above for gene filtering and outlier detection, we removed one outlier sample (inter-array correlation for this sample exceeded 3 SD; hence, we processed 94 samples). Next, keeping only the human genes, we gene-wise averaged all samples corresponding to the same subject (logRPKM expression) to obtain a robust estimate of gene expression and remove within-dataset interdependency. Similarly, the variables rRNA rate ratio of all reads aligned to rRNA regions to total reads, total assigned gene (proportion of aligned reads that were assigned to genes during read counting), and mitochondrial rate for human genes (proportion of reads that aligned to chrM) were also averaged. We regressed out along with the first principal component (PC1) the effect of these confounders, whereas we protected the effect of diagnosis, which, in these iPSCs, simply reflects the difference in terms of genetic risk for SCZ between patients and controls. We identified a network using WGCNA with the settings described above [except that we used the minimum beta granting scale invariance ($\beta = 7$) as per the standard WGCNA settings, as we did not have multiple networks to compare]. Then, for each module, we computed the hypergeometric overrepresentation of SCZ genes at various extension windows and 22 non-GWAS consensus genes (5 consensus genes of 28 that were already in the PGC3 list were excluded to avoid double dipping, whereas 1 gene was not expressed in iPSC) as target gene lists.

To test the hypothesis that the non-GWAS consensus genes are more highly coexpressed with SCZ risk genes than expected by chance, we took as connectivity measure the median of the gene-wise mean connectivity of consensus genes with all SCZ risk genes from the adjacency matrix of the iPSC network identified and compared it to a null distribution of randomly permuted non-SCZ and nonconsensus genes, after matching for gene set

size, GC content, gene length, and average expression as previously described.

The universe from which random genes were pooled consisted of genes expressed in the iPSC network, excluding SCZ risk genes (the most extensive list that we generated at 500 kbp from GWAS-significant SNPs was considered for this analysis) and consensus genes. Last, we defined the empirical P value as the number of occurrences in which the connectivity measure of the random genes exceeded the one of the consensus genes with SCZ risk genes divided by the number of permutations (10,000). We set the significance threshold at a one-sided empirical P value < 0.05 . We repeated this procedure for the nine extension windows previously created.

Supplementary Materials

This PDF file includes:

Figs. S1 to S11

Table S1 to S5

References

[View/request a protocol for this paper from Bio-protocol.](#)

REFERENCES AND NOTES

1. I. I. Gottesman, J. Shields, A polygenic theory of schizophrenia. *Proc. Natl. Acad. Sci. U.S.A.* **58**, 199–205 (1967).
2. V. Trubetskoy, A. F. Pardinas, T. Qi, G. Panagiotaropoulou, S. Awasthi, T. B. Bigdeli, J. Bryois, C. Y. Chen, C. A. Dennison, L. S. Hall, M. Lam, K. Watanabe, O. Frei, T. Ge, J. C. Harwood, F. Koopmans, S. Magnusson, A. L. Richards, J. Sidorenko, Y. Wu, J. Zeng, J. Grove, M. Kim, Z. Li, G. Voloudakis, W. Zhang, M. Adams, I. Agartz, E. G. Atkinson, E. Agerbo, M. Al Eissa, M. Albus, M. Alexander, B. Z. Alizadeh, K. Alptekin, T. D. Als, F. Amin, V. Arolt, M. Arrojo, L. Athanasiu, M. H. Azevedo, S. A. Bacanu, N. J. Bass, M. Begemann, R. A. Belliveau, J. Bene, B. Benyamin, S. E. Bergen, G. Blasi, J. Bobes, S. Bonassi, A. Braun, R. A. Bressan, E. J. Bromet, R. Bruggeman, P. F. Buckley, R. L. Buckner, J. Bybjerg-Grauholm, W. Cahn, M. J. Cairns, M. E. Calkins, V. J. Carr, D. Castle, S. V. Catts, K. D. Chambert, R. C. K. Chan, B. Chaumette, W. Cheng, E. F. C. Cheung, S. A. Chong, D. Cohen, A. Consoli, Q. Cordeiro, J. Costas, C. Curtis, M. Davidson, K. L. Davis, L. de Haan, F. Degenhardt, L. E. DeLisi, D. Demontis, F. Dickerson, D. Dikeos, T. Dinan, S. Djurovic, J. Duan, G. Ducci, F. Dudbridge, J. G. Eriksson, L. Fananas, S. V. Faraone, A. Fiorentino, A. Forstner, J. Frank, N. B. Freimer, M. Fromer, A. Frustaci, A. Gadelha, G. Genovese, E. S. Gershon, M. Giannitelli, I. Giegling, P. Giusti-Rodriguez, S. Godard, J. I. Goldstein, J. Gonzalez Penas, A. Gonzalez-Pinto, S. Gopal, J. Gratten, M. F. Green, T. A. Greenwood, O. Guillin, S. Guloksuz, R. E. Gur, R. C. Gur, B. Gutierrez, E. Hahn, H. Hakonarson, V. Haroutunian, A. M. Hartmann, C. Harvey, C. Hayward, F. A. Henskens, S. Herms, P. Hoffmann, D. P. Howrigan, M. Ikeda, C. Iyegbe, I. Joa, A. Julia, A. K. Kahler, T. Kam-Thong, Y. Kamatani, S. Karachanak-Yankova, O. Kebir, M. C. Keller, B. J. Kelly, A. Khrunin, S. W. Kim, J. Klivins, N. Kondratiev, B. Korta, J. Kraft, M. Kubo, V. Kucinskis, Z. A. Kucinskiene, A. Kusumawardhani, H. Kuzelova-Platkova, S. Landi, L. C. Lazzeroni, P. H. Lee, S. E. Legge, D. S. Lehrer, R. Lencer, B. Lerer, M. Li, J. Lieberman, G. A. Light, S. Limborska, C. M. Liu, J. Lonnqvist, C. M. Loughland, J. Lubinski, J. J. Luykx, A. Lynham, M. Macek Jr., A. Mackinnon, P. K. E. Magnusson, B. S. Maher, W. Maier, D. Malaspina, J. Mallet, S. R. Marder, S. Marsal, A. R. Martin, L. Martorell, M. Mattheisen, R. W. McCarley, C. McDonald, J. J. McGrath, H. Medeiros, S. Meier, B. Melegh, I. Melle, R. I. Meshulam-Gately, A. Metspalu, P. T. Michie, L. Milani, V. Milanova, M. Mitjans, E. Molden, E. Molina, M. D. Molto, V. Mondelli, C. Moreno, C. P. Morley, G. Muntane, K. C. Murphy, I. Myin-Germers, I. Nenadic, G. Nestadt, L. Nikitina-Zake, C. Noto, K. H. Nuechterlein, N. L. O'Brien, F. A. O'Neill, S. Y. Oh, A. Olincy, V. K. Ota, C. Pantelis, G. N. Papadimitriou, M. Parellada, T. Paunio, R. Pellegrino, S. Periyasamy, D. O. Perkins, B. Pfuhlmann, O. Pietilainen, J. Pimm, D. Porteous, J. Powell, D. Quattrone, D. Quedest, A. D. Radant, A. Rampino, M. H. Rapaport, A. Rautanen, A. Reichenberg, C. Roe, J. L. Roffman, J. Roth, M. Rothermundt, B. P. F. Rutten, S. Saker-Delye, V. Salomaa, J. Sanjuan, M. L. Santoro, A. Savitz, U. Schall, R. J. Scott, L. J. Seidman, S. I. Sharp, J. Shi, L. J. Siever, E. Sigurdsson, K. Sim, N. Skarabis, P. Slominsky, H. C. So, J. L. Sobell, E. Soderman, H. J. Stain, N. E. Steen, A. A. Steixner-Kumar, E. Stogmann, W. S. Stone, R. E. Straub, F. Streit, E. Strengman, T. S. Stroup, M. Subramaniam, C. A. Sugar, J. Suvisaari, D. M. Svrakic, N. R. Swerdlow, J. P. Szatkiewicz, T. M. T. Ta, A. Takahashi, C. Terao, F. Thibaut, D. Toncheva, P. A. Tooney, S. Torretta, S. Tosato, G. B. Tura, B. I. Turetsky, A. Ucol, A. Vaaler, T. van Amelsvoort, R. van Winkel, J. Veijola, J. Waddington, H. Walter, A. Waterreus,

- B. T. Webb, M. Weiser, N. M. Williams, S. H. Witt, B. K. Wormley, J. Q. Wu, Z. Xu, R. Yolken, C. C. Zai, W. Zhou, F. Zhu, F. Zimprich, E. C. Atbasoglu, M. Ayub, C. Benner, A. Bertolino, D. W. Black, N. J. Bray, G. Breen, N. G. Buccola, W. F. Byerley, W. J. Chen, C. R. Cloninger, B. Crespo-Facorro, G. Donohoe, R. Freedman, C. Galletly, M. J. Gandal, M. Gennarelli, D. M. Hougaard, H. G. Hwu, A. V. Jablensky, S. A. McCarroll, J. L. Moran, O. Mors, P. B. Mortensen, B. Muller-Myhsok, A. L. Neil, M. Nordentoft, M. T. Pato, T. L. Petryshen, M. Pirinen, A. E. Pulver, T. G. Schulze, J. M. Silverman, J. W. Smoller, E. A. Stahl, D. W. Tsuang, E. Vilella, S. H. Wang, S. Xu; Indonesia Schizophrenia Consortium; PsychENCODE; Psychosis Endophenotypes International Consortium; The SynGO Consortium, R. Adolfsson, C. Arango, B. T. Baune, S. I. Belanger, A. D. Borglum, D. Braff, E. Bramer, J. D. Buxbaum, D. Campion, J. A. Cervilla, S. Cichon, D. A. Collier, A. Corvin, D. Curtis, M. D. Forti, E. Domenici, H. Ehrenreich, V. Escott-Price, T. Esko, A. H. Fanous, A. Gareeva, M. Gawlik, P. V. Gejman, M. Gill, S. J. Glatt, V. Golimbet, K. S. Hong, C. M. Hultman, S. E. Hyman, N. Iwata, E. G. Jonsson, R. S. Kahn, J. L. Kennedy, E. Khusnutdinova, G. Kirov, J. A. Knowles, M. O. Krebs, C. Laurent-Levinson, J. Lee, T. Lencz, D. F. Levinson, Q. S. Li, J. Liu, A. K. Malhotra, D. Malhotra, A. McIntosh, A. McQuillin, P. R. Menezes, V. A. Morgan, D. W. Morris, B. J. Mowry, R. M. Murray, V. Nimgaonkar, M. M. Nothen, R. A. Ophoff, S. A. Paciga, A. Palotie, C. N. Pato, S. Qin, M. Rietschel, B. P. Riley, M. Rivera, D. Rujescu, M. C. Saka, A. R. Sanders, S. G. Schwab, A. Serretti, P. C. Sham, Y. Shi, D. S. Clair, H. Stefansson, K. Stefansson, M. T. Tsuang, J. van Os, M. P. Vawter, D. R. Weinberger, T. Werge, D. B. Wildenauer, X. Yu, W. Yue, P. A. Holmans, A. J. Pocklington, P. Roussos, E. Vassos, M. Verhage, P. M. Visscher, J. Yang, D. Posthuma, O. A. Andreassen, K. S. Kendler, M. J. Owen, N. R. Wray, M. J. Daly, H. Huang, B. M. Neale, P. F. Sullivan, S. Ripke, J. T. R. Walters, M. C. O'Donovan; Schizophrenia Working Group of the Psychiatric Genomics Consortium, Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature* **604**, 502–508 (2022).
3. E. A. Boyle, Y. I. Li, J. K. Pritchard, An expanded view of complex traits: From polygenic to omnigenic. *Cell* **169**, 1177–1186 (2017).
 4. M. Li, A. E. Jaffe, R. E. Straub, R. Tao, J. H. Shin, Y. Wang, Q. Chen, C. Li, Y. Jia, K. Ohi, B. J. Maher, N. J. Brandon, A. Cross, J. G. Chenoweth, D. J. Hoepfner, H. Wei, T. M. Hyde, R. McKay, J. E. Kleinman, D. R. Weinberger, A human-specific AS3MT isoform and BORCS7 are molecular risk factors in the 10q24.32 schizophrenia-associated locus. *Nat. Med.* **22**, 649–656 (2016).
 5. M. Fromer, P. Roussos, S. K. Sieberts, J. S. Johnson, D. H. Kavanagh, T. M. Perumal, D. M. Ruderfer, E. C. Oh, A. Topol, H. R. Shah, L. L. Klei, R. Kramer, D. Pinto, Z. H. Gumus, A. E. Cicek, K. K. Dang, A. Browne, C. Lu, L. Xie, B. Readhead, E. A. Stahl, J. Xiao, M. Parvizi, T. Hamamsy, J. F. Fullard, Y. C. Wang, M. C. Mahajan, J. M. Derry, J. T. Dudley, S. E. Hemby, B. A. Logsdon, K. Talbot, T. Raj, D. A. Bennett, P. L. De Jager, J. Zhu, B. Zhang, P. F. Sullivan, A. Chess, S. M. Purcell, L. A. Shinobu, L. M. Mangravite, H. Toyoshiba, R. E. Gur, C. G. Hahn, D. A. Lewis, V. Haroutunian, M. A. Peters, B. K. Lipska, J. D. Buxbaum, E. E. Schadt, K. Hirai, K. Roeder, K. J. Brennan, N. Katsanis, E. Domenici, B. Devlin, P. Sklar, Gene expression elucidates functional impact of polygenic risk for schizophrenia. *Nat. Neurosci.* **19**, 1442–1453 (2016).
 6. A. E. Jaffe, R. E. Straub, J. H. Shin, R. Tao, Y. Gao, L. Collado-Torres, T. Kam-Thong, H. S. Xi, J. Quan, Q. Chen, C. Colantuoni, W. S. Ulrich, B. J. Maher, A. Deep-Soboslay, C. BrainSeq, A. J. Cross, N. J. Brandon, J. T. Leek, T. M. Hyde, J. E. Kleinman, D. R. Weinberger, Developmental and genetic regulation of the human cortex transcriptome illuminate schizophrenia pathogenesis. *Nat. Neurosci.* **21**, 1117–1125 (2018).
 7. A. Gusev, N. Mancuso, H. Won, M. Kousi, H. K. Finucane, Y. Reshef, L. Song, A. Safi; Schizophrenia Working Group of the Psychiatric Genomics Consortium, S. McCarroll, B. M. Neale, R. A. Ophoff, M. C. O'Donovan, G. E. Crawford, D. H. Geschwind, N. Katsanis, P. F. Sullivan, B. Pasaniuc, A. L. Price, Transcriptome-wide association study of schizophrenia and chromatin activity yields mechanistic disease insights. *Nat. Genet.* **50**, 538–548 (2018).
 8. L. M. Huckins, A. Dobbyn, D. M. Ruderfer, G. Hoffman, W. Wang, A. F. Pardinas, V. M. Rajagopal, T. D. Als, H. T. Nguyen, K. Girdhar, J. Boockck, P. Roussos, M. Fromer, R. Kramer, E. Domenici, E. R. Gamazon, S. Purcell; CommonMind Consortium; The Schizophrenia Working Group of the Psychiatric Genomics Consortium; iPSYCH-GEMS Schizophrenia Working Group, D. Demontis, A. D. Borglum, J. T. R. Walters, M. C. O'Donovan, P. Sullivan, M. J. Owen, B. Devlin, S. K. Sieberts, N. J. Cox, H. K. Im, P. Sklar, E. A. Stahl, Gene expression imputation across multiple brain regions provides insights into schizophrenia risk. *Nat. Genet.* **51**, 659–674 (2019).
 9. N. Schrodde, S. M. Ho, K. Yamamuro, A. Dobbyn, L. Huckins, M. R. Matos, E. Cheng, P. J. M. Deans, E. Flaherty, N. Barretto, A. Topol, K. Alganem, S. Abadali, J. Gregory, E. Hoelzli, H. Phatnani, V. Singh, D. Girish, B. Aronow, R. McCullumsmith, G. E. Hoffman, E. A. Stahl, H. Morishita, P. Sklar, K. J. Brennan, Synergistic effects of common schizophrenia risk variants. *Nat. Genet.* **51**, 1475–1485 (2019).
 10. M. J. Gandal, J. R. Haney, N. N. Parikhshak, V. Leppa, G. Ramaswami, C. Hartl, A. J. Schork, V. Appadurai, A. Buil, T. M. Werge, C. Liu, K. P. White; CommonMind Consortium; PsychENCODE Consortium; iPSYCH-BROAD Working Group, S. Horvath, D. H. Geschwind, Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. *Science* **359**, 693–697 (2018).
 11. C. L. Hartl, G. Ramaswami, W. G. Pembroke, S. Muller, G. Pintacuda, A. Saha, P. Parsana, A. Battle, K. Lage, D. H. Geschwind, Coexpression network architecture reveals the brain-wide and multiregional basis of disease susceptibility. *Nat. Neurosci.* **24**, 1313–1323 (2021).
 12. E. Radulescu, A. E. Jaffe, R. E. Straub, Q. Chen, J. H. Shin, T. M. Hyde, J. E. Kleinman, D. R. Weinberger, Identification and prioritization of gene sets associated with schizophrenia risk by co-expression network analysis in human brain. *Mol. Psychiatry* **25**, 791–804 (2020).
 13. G. Pergola, P. Di Carlo, E. D'Ambrosio, B. Gelao, L. Fazio, M. Papalino, A. Monda, G. Scozia, B. Pietrangelo, M. Attrotto, J. A. Apud, Q. Chen, V. S. Mattay, A. Rampino, G. Caforio, D. R. Weinberger, G. Blasi, A. Bertolino, DRD2 co-expression network and a related polygenic index predict imaging, behavioral and clinical phenotypes linked to schizophrenia. *Transl. Psychiatry* **7**, e1006 (2017).
 14. G. Pergola, P. Di Carlo, A. E. Jaffe, M. Papalino, Q. Chen, T. M. Hyde, J. E. Kleinman, J. H. Shin, A. Rampino, G. Blasi, D. R. Weinberger, A. Bertolino, Prefrontal coexpression of schizophrenia risk genes is associated with treatment response in patients. *Biol. Psychiatry* **86**, 45–55 (2019).
 15. U. Braun, A. Harneit, G. Pergola, T. Menara, A. Schafer, R. F. Betzel, Z. Zang, J. I. Schweiger, X. Zhang, K. Schwarz, J. Chen, G. Blasi, A. Bertolino, D. Durstewitz, F. Pasqualetti, E. Schwarz, A. Meyer-Lindenberg, D. S. Bassett, H. Tost, Brain network dynamics during working memory are modulated by dopamine and diminished in schizophrenia. *Nat. Commun.* **12**, 3478 (2021).
 16. D. R. Weinberger, Implications of normal brain development for the pathogenesis of schizophrenia. *Arch. Gen. Psychiatry* **44**, 660–669 (1987).
 17. E. K. Lambe, L. S. Krimer, P. S. Goldman-Rakic, Differential postnatal development of catecholamine and serotonin inputs to identified neurons in prefrontal cortex of rhesus monkey. *J. Neurosci.* **20**, 8780–8787 (2000).
 18. E. Walker, R. J. Lewine, Prediction of adult-onset schizophrenia from childhood home movies of the patients. *Am. J. Psychiatry* **147**, 1052–1056 (1990).
 19. D. R. Weinberger, P. J. Harrison, *Schizophrenia* (Wiley-Blackwell, 2011).
 20. Q. Chen, G. Ursini, A. L. Romer, A. R. Knodt, K. Mezeivitch, E. Xiao, G. Pergola, G. Blasi, R. E. Straub, J. H. Callicott, K. F. Berman, A. R. Hariri, A. Bertolino, V. S. Mattay, D. R. Weinberger, Schizophrenia polygenic risk score predicts mnemonic hippocampal activity. *Brain* **141**, 1218–1228 (2018).
 21. O. D. Howes, R. McCutcheon, M. J. Owen, R. M. Murray, The role of genes, stress, and dopamine in the development of schizophrenia. *Biol. Psychiatry* **81**, 9–20 (2017).
 22. A. Abi-Dargham, From "bedside" to "bench" and back: A translational approach to studying dopamine dysfunction in schizophrenia. *Neurosci. Biobehav. Rev.* **110**, 174–179 (2020).
 23. A. E. Jaffe, D. J. Hoepfner, T. Saito, L. Blanpain, J. Ukaigwe, E. E. Burke, L. Collado-Torres, R. Tao, K. Tajinda, K. R. Maynard, M. N. Tran, K. Martinowich, A. Deep-Soboslay, J. H. Shin, J. E. Kleinman, D. R. Weinberger, M. Matsumoto, T. M. Hyde, Profiling gene expression in the human dentate gyrus granule cell layer reveals insights into schizophrenia and its genetic risk. *Nat. Neurosci.* **23**, 510–519 (2020).
 24. D. Cameron, D. Mi, N. N. Vinh, C. Webber, M. Li, O. Marin, M. C. O'Donovan, N. J. Bray, Single-Nuclei RNA sequencing of 5 regions of the human prenatal brain implicates developing neuron populations in genetic risk for schizophrenia. *Biol. Psychiatry*, (2022).
 25. A. J. Willsey, S. J. Sanders, M. Li, S. Dong, A. T. Tebbenkamp, R. A. Muhle, S. K. Reilly, L. Lin, S. Fertuzinhos, J. A. Miller, M. T. Murtha, C. Bichsel, W. Niu, J. Cotney, A. G. Ercan-Sencicek, J. Gockley, A. R. Gupta, W. Han, X. He, E. J. Hoffman, L. Klei, J. Lei, W. Liu, L. Liu, C. Lu, X. Xu, Y. Zhu, S. M. Mane, E. S. Lein, L. Wei, J. P. Noonan, K. Roeder, B. Devlin, N. Sestan, M. W. State, Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. *Cell* **155**, 997–1007 (2013).
 26. R. L. Walker, G. Ramaswami, C. Hartl, N. Mancuso, M. J. Gandal, L. de la Torre-Ubieta, B. Pasaniuc, J. L. Stein, D. H. Geschwind, Genetic control of expression and splicing in developing human brain informs disease mechanisms. *Cell* **179**, 750–771.e22 (2019).
 27. D. M. Werling, S. Pocharedddy, J. Choi, J. Y. An, B. Sheppard, M. Peng, Z. Li, C. Dastmalchi, G. Santpere, A. M. M. Sousa, A. T. N. Tebbenkamp, N. Kaur, F. O. Gulden, M. S. Breen, L. Liang, M. C. Gilson, X. Zhao, S. Dong, L. Klei, A. E. Cicek, J. D. Buxbaum, H. Adle-Biassette, J. L. Thomas, K. A. Aldinger, D. R. O'Day, I. A. Glass, N. A. Zaitlen, M. E. Talkowski, K. Roeder, M. W. State, B. Devlin, S. J. Sanders, N. Sestan, Whole-genome and RNA sequencing reveal variation and transcriptomic coordination in the developing human prefrontal cortex. *Cell Rep.* **31**, 107489 (2020).
 28. N. Y. A. Sey, B. Hu, W. Mah, H. Fauni, J. C. McAfee, P. Rajarajan, K. J. Brennan, S. Akbarian, H. Won, A computational tool (H-MAGMA) for improved prediction of brain-disorder risk genes by incorporating brain chromatin interaction profiles. *Nat. Neurosci.* **23**, 583–593 (2020).

29. A. E. Jaffe, J. Shin, L. Collado-Torres, J. T. Leek, R. Tao, C. Li, Y. Gao, Y. Jia, B. J. Maher, T. M. Hyde, J. E. Kleinman, D. R. Weinberger, Developmental regulation of human cortex transcription and its clinical relevance at single base resolution. *Nat. Neurosci.* **18**, 154–161 (2015).
30. C. Colantuoni, B. K. Lipska, T. Ye, T. M. Hyde, R. Tao, J. T. Leek, E. A. Colantuoni, A. G. Elkhalloun, M. M. Herman, D. R. Weinberger, J. E. Kleinman, Temporal dynamics and genetic control of transcription in the human prefrontal cortex. *Nature* **478**, 519–523 (2011).
31. M. Li, G. Santpere, Y. Imamura Kawasawa, O. V. Evgrafov, F. O. Gulden, S. Pochareddy, S. M. Sunkin, Z. Li, Y. Shin, Y. Zhu, A. M. M. Sousa, D. M. Werling, R. R. Kitchen, H. J. Kang, M. Pletikos, J. Choi, S. Muchnik, X. Xu, D. Wang, B. Lorente-Galdos, S. Liu, P. Giusti-Rodriguez, H. Won, C. A. de Leeuw, A. F. Pardinas; BrainSpan Consortium; PsychENCODE Consortium; PsychENCODE Developmental Subgroup, M. Hu, F. Jin, Y. Li, M. J. Owen, M. C. O'Donovan, J. T. R. Walters, D. Posthuma, M. A. Reimers, P. Levitt, D. R. Weinberger, T. M. Hyde, J. E. Kleinman, D. H. Geschwind, M. J. Hawrylycz, M. W. State, S. J. Sanders, P. F. Sullivan, M. B. Gerstein, E. S. Lein, J. A. Knowles, N. Sestan, Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science* **362**, eaat7615 (2018).
32. D. Wang, S. Liu, J. Warrell, H. Won, X. Shi, F. C. P. Navarro, D. Clarke, M. Gu, P. Emami, Y. T. Yang, M. Xu, M. J. Gandal, S. Lou, J. Zhang, J. J. Park, C. Yan, S. K. Rhie, K. Manakongtreecheep, H. Zhou, A. Nathan, M. Peters, E. Mattei, D. Fitzgerald, T. Brunetti, J. Moore, Y. Jiang, K. Girdhar, G. E. Hoffman, S. Kalayci, Z. H. Gumus, G. E. Crawford; PsychENCODE Consortium, P. Roussos, S. Akbarian, A. E. Jaffe, K. P. White, Z. Weng, N. Sestan, D. H. Geschwind, J. A. Knowles, M. B. Gerstein, Comprehensive functional genomic resource and integrative model for the human brain. *Science* **362**, eaat8464 (2018).
33. S. C. Page, S. R. Sripathy, F. Farinelli, Z. Ye, Y. Wang, D. J. Hiler, E. A. Pattie, C. V. Nguyen, M. T. Tippani, R. L. Moses, H.-Y. Chen, M. N. Tran, N. J. Eagles, J. M. Stolz, J. L. Cattalini, O. R. Soudry, D. Dickinson, K. F. Berman, J. A. Apud, D. R. Weinberger, K. Martinowich, A. E. Jaffe, R. E. Straub, B. J. Maher, Electrophysiological measures from human iPSC-derived neurons are associated with schizophrenia clinical status and predict individual cognitive performance. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2109395119 (2022).
34. B. Zhang, S. Horvath, A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.* **4**, Article17 (2005).
35. G. Pergola, A. Rampino, P. Di Carlo, A. Marakhovskaia, T. Quarto, L. Fazio, M. Papalino, S. Torretta, N. Amoroso, M. N. Castro, E. Domenici, J. Dukart, J. Khilghatyan, A. Monaco, T. Popolizio, R. Romano, L. Sportelli, H. Zunzer, G. Blasi, J. M. Beaulieu, A. Bertolino, A miR-137-related biological pathway of risk for schizophrenia is associated with human brain emotion processing. *bioRxiv* 2020.08.03.230227, (2020).
36. M. J. Gandal, P. Zhang, E. Hadjimiichael, R. L. Walker, C. Chen, S. Liu, H. Won, H. van Bakel, M. Varghese, Y. Wang, A. W. Shieh, J. Haney, S. Parhami, J. Belmont, M. Kim, P. Moran Losada, Z. Khan, J. Mleczo, Y. Xia, R. Dai, D. Wang, Y. T. Yang, M. Xu, K. Fish, P. R. Hof, J. Warrell, D. Fitzgerald, K. White, A. E. Jaffe, E. C. Psych, M. A. Peters, M. Gerstein, C. Liu, L. M. Iakoucheva, D. Pinto, D. H. Geschwind, Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. *Science* **362**, (2018).
37. R. Birnbaum, D. R. Weinberger, Genetic insights into the neurodevelopmental origins of schizophrenia. *Nat. Rev. Neurosci.* **18**, 727–740 (2017).
38. R. Yurko, K. Roeder, B. Devlin, M. G'Sell, H-MAGMA, inheriting a shaky statistical foundation, yields excess false positives. *Ann. Hum. Genet.* **85**, 97–100 (2021).
39. C. de Leeuw, N. Y. A. Sey, D. Posthuma, H. Won, A response to Yurko et al: H-MAGMA, inheriting a shaky statistical foundation, yields excess false positives. *bioRxiv* 2020.09.25.310722, (2020).
40. T. Singh, T. Poterba, D. Curtis, H. Akil, M. Al Eissa, J. D. Barchas, N. Bass, T. B. Bigdely, G. Breen, E. J. Bromet, P. F. Buckley, W. E. Bunney, J. Bybjerg-Grauholm, W. F. Byerley, S. B. Chapman, W. J. Chen, C. Churchhouse, N. Craddock, C. M. Cusick, L. DeLisi, S. Dodge, M. A. Escamilla, S. Eskelinen, A. H. Fanous, S. V. Faraone, A. Fiorentino, L. Francioli, S. B. Gabriel, D. Gage, S. A. Gagliano Taliun, A. Ganna, G. Genovese, D. C. Glahn, J. Grove, M. H. Hall, E. Hamalainen, H. O. Heyne, M. Holi, D. M. Hougaard, D. P. Howrigan, H. Huang, H. G. Hwu, R. S. Kahn, H. M. Kang, K. J. Karczewski, G. Kirov, J. A. Knowles, F. S. Lee, D. S. Lehrer, F. Lescai, D. Malaspina, S. R. Marder, S. A. McCarroll, A. M. McIntosh, H. Medeiros, L. Milani, C. P. Morley, D. W. Morris, P. B. Mortensen, R. M. Myers, M. Nordentoft, N. L. O'Brien, A. M. Olivares, D. Ongur, W. H. Ouwehand, D. S. Palmer, T. Paunio, D. Quested, M. H. Rapoport, E. Rees, B. Rollins, F. K. Satterstrom, A. Schatzberg, E. Scolnick, L. J. Scott, S. I. Sharp, P. Sklar, J. W. Smoller, J. L. Sobell, M. Solomonson, E. A. Stahl, C. R. Stevens, J. Suvisaari, G. Tiao, S. J. Watson, N. A. Watts, D. H. Blackwood, A. D. Borglum, B. M. Cohen, A. P. Corvin, T. Esko, N. B. Freimer, S. J. Glatt, C. T. M. Hultman, A. McQuillin, A. Palotie, C. N. Pato, M. T. Pato, A. E. Pulver, D. S. Clair, M. T. Tsuang, M. P. Pawtew, J. T. Walters, T. M. Werge, R. A. Ophoff, P. F. Sullivan, M. J. Owen, M. Boehnke, M. C. O'Donovan, B. M. Neale, M. J. Daly, Rare coding variants in ten genes confer substantial risk for schizophrenia. *Nature* **604**, 509–516 (2022).
41. M. N. Tran, K. R. Maynard, A. Spangler, L. A. Huuki, K. D. Montgomery, V. Sadashivaiah, M. T. Tippani, B. K. Barry, D. B. Hancock, S. C. Hicks, J. E. Kleinman, T. M. Hyde, L. Collado-Torres, A. E. Jaffe, K. Martinowich, Single-nucleus transcriptome analysis reveals cell-type-specific molecular signatures across reward circuitry in the human brain. *Neuron* **109**, 3088–3103.e5 (2021).
42. D. R. Weinberger, From neuropathology to neurodevelopment. *Lancet* **346**, 552–557 (1995).
43. D. R. Weinberger, Thinking about schizophrenia in an era of genomic medicine. *Am. J. Psychiatry* **176**, 12–20 (2019).
44. H. Tost, A. Meyer-Lindenberg, Puzzling over schizophrenia: Schizophrenia, social environment and the brain. *Nat. Med.* **18**, 211–213 (2012).
45. Y. Zhu, A. M. M. Sousa, T. Gao, M. Skarica, M. Li, G. Santpere, P. Esteller-Cucala, D. Juan, L. Ferrandez-Peral, F. O. Gulden, M. Yang, D. J. Miller, T. Marques-Bonet, Y. Imamura Kawasawa, H. Zhao, N. Sestan, Spatiotemporal transcriptomic divergence across human and macaque brain development. *Science* **362**, (2018).
46. L. de la Torre-Ubieta, J. L. Stein, H. Won, C. K. Opland, D. Liang, D. Lu, D. H. Geschwind, The dynamic landscape of open chromatin during human cortical neurogenesis. *Cell* **172**, 289–304.e18 (2018).
47. H. Won, L. de la Torre-Ubieta, J. L. Stein, N. N. Parikshak, J. Huang, C. K. Opland, M. J. Gandal, G. J. Sutton, F. Hormozdiari, D. Lu, C. Lee, E. Eskin, I. Voineagu, J. Ernst, D. H. Geschwind, Chromosome conformation elucidates regulatory relationships in developing human brain. *Nature* **538**, 523–527 (2016).
48. G. Ursini, G. Punzi, Q. Chen, S. Marengo, J. F. Robinson, A. Porcelli, E. G. Hamilton, M. Mitjans, G. Maddalena, M. Begemann, J. Seidel, H. Yanamori, A. E. Jaffe, K. F. Berman, M. F. Egan, R. E. Straub, C. Colantuoni, G. Blasi, R. Hashimoto, D. Rujescu, H. Ehrenreich, A. Bertolino, D. R. Weinberger, Convergence of placenta biology and genetic risk for schizophrenia. *Nat. Med.* **24**, 792–801 (2018).
49. J. H. Callicott, A. Bertolino, V. S. Mattay, F. J. Langheim, J. Duyn, R. Coppola, T. E. Goldberg, D. R. Weinberger, Physiological dysfunction of the dorsolateral prefrontal cortex in schizophrenia revisited. *Cereb. Cortex* **10**, 1078–1092 (2000).
50. J. H. Callicott, M. F. Egan, V. S. Mattay, A. Bertolino, A. D. Bone, B. Verchinski, D. R. Weinberger, Abnormal fMRI response of the dorsolateral prefrontal cortex in cognitively intact siblings of patients with schizophrenia. *Am. J. Psychiatry* **160**, 709–719 (2003).
51. Y. Patel, J. Shin, C. Abe, I. Agartz, C. Alloza, D. Alnaes, S. Ambrogio, L. A. Antonucci, C. Arango, V. Arolt, G. Auzias, R. Ayesa-Arriola, N. Banaj, T. Banaschewski, C. Bandeira, Z. Basgoze, R. B. Cupertino, C. H. D. Bau, J. Bauer, S. Baumeister, F. Bernardoni, A. Bertolino, C. D. M. Bonnín, D. Brandeis, S. Brem, J. Bruggemann, R. Bulow, J. R. Bustillo, S. Calderoni, R. Calvo, E. J. Canales-Rodriguez, D. M. Cannon, S. Carmona, V. J. Carr, S. V. Catts, S. Chenji, Q. H. Chew, D. Coghill, C. G. Connolly, A. Conzelmann, A. R. Craven, B. Crespo-Facorro, K. Cullen, A. Dahl, U. Dannlowski, C. G. Davey, C. Deruelle, C. M. Diaz-Caneja, K. Dohm, S. Ehrlich, J. Epstein, T. Erwin-Grabner, L. T. Eyer, J. Fedor, J. Fitzgerald, W. Foran, J. M. Ford, L. Fortea, P. Fuentes-Claramonte, J. Fullerton, L. Furlong, L. Gallagher, B. Gao, S. Gao, J. M. Goikolea, I. Gotlib, R. Goya-Maldonado, H. J. Grabe, M. Green, E. H. Grevet, N. A. Groenewold, D. Grotegerd, O. Gruber, J. Haavik, T. Hahn, B. J. Harrison, W. Heindel, F. Henskens, D. J. Heslenfeld, E. Hilland, P. J. Hoekstra, S. Hohmann, N. Holz, F. M. Howells, J. C. Ipser, N. Jahanshad, B. Jakobi, A. Jansen, J. Janssen, R. Jonassen, A. Kaiser, V. Kaleda, J. Karantonis, J. A. King, T. Kircher, P. Kochunov, S. M. Koopowitz, M. Landen, N. I. Landro, S. Lawrie, I. Lebedeva, B. Luna, A. J. Lundervold, F. P. MacMaster, L. A. Maglanoc, D. H. Mathalon, C. McDonald, A. McIntosh, S. Meinert, P. T. Michie, P. Mitchell, A. Moreno-Alcazar, B. Mowry, F. Muratori, L. Nabulsi, I. Nenadic, R. O'Gorman Tuura, J. Oosterlaan, B. Overs, C. Pantelis, M. Parellada, J. C. Pariente, P. Pauli, G. Pergola, F. M. Piarulli, F. Picon, F. Piras, E. Pomarol-Clotet, C. Pretus, Y. Quide, J. Radua, J. A. Ramos-Quiroga, P. E. Rasser, A. Reif, A. Retico, G. Roberts, S. Rossell, D. L. Rovaris, K. Rubia, M. Sacchet, J. Salavert, R. Salvador, S. Sarro, A. Sawa, U. Schall, R. Scott, P. Selvaggi, T. Silk, K. Sim, A. Skoch, G. Spalletta, F. Spaniel, D. J. Stein, O. Steinstrater, A. Stolicyn, Y. Takayanagi, L. Tamm, M. Tavares, A. Teumer, K. Thiel, S. I. Thomopoulos, D. Tomecek, A. S. Tomyshev, D. Tordesillas-Gutierrez, M. Tosetti, A. Uhlmann, T. Van Rheenen, J. Vazquez-Bourgoin, M. W. Vernooij, E. Vieta, O. Vilarroya, C. Weickert, T. Weickert, L. T. Westlye, H. Whalley, D. Willinger, A. Winter, K. Wittfeld, T. T. Yang, Y. Yoncheva, J. L. Zijlmans, M. Hoogman, B. Franke, D. van Rooij, J. Buitelaar, C. R. K. Ching, O. A. Andreassen, E. Pozzi, D. Veltman, L. Schmaal, T. G. M. van Erp, J. Turner, F. X. Castellanos, Z. Suskova, P. Thompson, T. Paus, Virtual ontogeny of cortical growth preceding mental illness. *Biol. Psychiat.* **92**, 299–313 (2022).
52. K. J. M. Benjamin, Q. Chen, A. E. Jaffe, J. M. Stolz, L. Collado-Torres, L. A. Huuki-Myers, E. E. Burke, R. Arora, A. S. Feltrin, A. R. Barbosa, E. Radulescu, G. Pergola, J. H. Shin, W. S. Ulrich, A. Deep-Soboslay, R. Tao, C. BrainSeq, T. M. Hyde, J. E. Kleinman, J. A. Erwin, D. R. Weinberger, A. C. M. Paquola, Analysis of the caudate nucleus transcriptome in individuals with schizophrenia highlights effects of antipsychotics and new risk genes. *Nat. Neurosci.* **25**, 1559–1568 (2022).

53. E. D'Ambrosio, G. Pergola, A. F. Pardini, D. Dahoun, M. Veronese, L. Sportelli, P. Taurisano, M. Griffiths, S. Jauhar, M. Rogdaki, M. A. P. Bloomfield, S. Froud-Rust-Walsh, I. Bonoldi, J. T. R. Walters, G. Blasi, A. Bertolino, O. D. Howes, A polygenic score indexing a DRD2-related co-expression network is associated with striatal dopamine function. *Sci. Rep.* **12**, 12610 (2022).
54. P. Taurisano, G. Pergola, A. Monda, L. A. Antonucci, P. Di Carlo, F. Piarulli, R. Passiatore, M. Papalino, R. Romano, A. Monaco, A. Rampino, A. Bonvino, A. Porcelli, T. Popolizio, R. Bellotti, A. Bertolino, G. Blasi, The interaction between cannabis use and a CB1-related polygenic co-expression index modulates dorsolateral prefrontal activity during working memory processing. *Brain Imaging Behav.* **15**, 288–299 (2021).
55. S. Torretta, A. Rampino, M. Basso, G. Pergola, P. Di Carlo, J. H. Shin, J. E. Kleinman, T. M. Hyde, D. R. Weinberger, R. Masellis, G. Blasi, M. Pennuto, A. Bertolino, NURR1 and ERR1 modulate the expression of genes of a DRD2 coexpression network enriched for schizophrenia risk. *J. Neurosci.* **40**, 932–941 (2020).
56. L. A. Antonucci, P. Di Carlo, R. Passiatore, M. Papalino, A. Monda, N. Amoroso, S. Tangaro, P. Taurisano, A. Rampino, F. Sambataro, T. Popolizio, A. Bertolino, G. Pergola, G. Blasi, Thalamic connectivity measured with fMRI is associated with a polygenic index predicting thalamo-prefrontal gene co-expression. *Brain Struct. Funct.* **224**, 1331–1344 (2019).
57. P. Selvaggi, G. Pergola, B. Gelao, P. Di Carlo, M. A. Nettis, G. Amico, L. Fazio, A. Rampino, F. Sambataro, G. Blasi, A. Bertolino, Genetic variation of a DRD2 co-expression network is associated with changes in prefrontal function after D2 receptors stimulation. *Cereb. Cortex* **29**, 1162–1173 (2019).
58. L. Fazio, G. Pergola, M. Papalino, P. Di Carlo, A. Monda, B. Gelao, N. Amoroso, S. Tangaro, A. Rampino, T. Popolizio, A. Bertolino, G. Blasi, Transcriptomic context of DRD1 is associated with prefrontal activity and behavior during working memory. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 5582–5587 (2018).
59. A. Monaco, A. Monda, N. Amoroso, A. Bertolino, G. Blasi, P. Di Carlo, M. Papalino, G. Pergola, S. Tangaro, R. Bellotti, A complex network approach reveals a pivotal sub-structure of genes linked to schizophrenia. *PLoS ONE* **13**, e0190110 (2018).
60. G. Pergola, N. Penzel, L. Sportelli, A. Bertolino, Lessons learned from parsing genetic risk for schizophrenia into biological pathways. *Biol. Psychiatry*, in press (2022).
61. L. Collado-Torres, E. E. Burke, A. Peterson, J. Shin, R. E. Straub, A. Rajpurohit, S. A. Semick, W. S. Ulrich, C. BrainSeq, A. J. Price, C. Valencia, R. Tao, A. Deep-Soboslay, T. M. Hyde, J. E. Kleinman, D. R. Weinberger, A. E. Jaffe, Regional heterogeneity in gene expression, regulation, and coherence in the frontal cortex and hippocampus across development and schizophrenia. *Neuron* **103**, 203–216.e8 (2019).
62. P. Parsana, C. Ruberman, A. E. Jaffe, M. C. Schatz, A. Battle, J. T. Leek, Addressing confounding artifacts in reconstruction of gene co-expression networks. *Genome Biol.* **20**, 94 (2019).
63. S. Freytag, J. Gagnon-Bartsch, T. P. Speed, M. Bahlo, Systematic noise degrades gene co-expression signals but can be corrected. *BMC Bioinformatics* **16**, 309 (2015).
64. A. E. Jaffe, R. Tao, A. L. Norris, M. Kealhofer, A. Nellore, J. H. Shin, D. Kim, Y. Jia, T. M. Hyde, J. E. Kleinman, R. E. Straub, J. T. Leek, D. R. Weinberger, qSVA framework for RNA quality correction in differential expression analysis. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 7130–7135 (2017).
65. A. T. McKenzie, M. Wang, M. E. Hauberg, J. F. Fullard, A. Kozlenkov, A. Keenan, Y. L. Hurd, S. Dracheva, P. Casaccia, P. Roussos, B. Zhang, Brain cell type specific gene expression and co-expression network architectures. *Sci. Rep.* **8**, 8868 (2018).
66. L. Collado-Torres, A. Jaffe, E. Burke, LieberInstitute/Jaffelab: Zenodo integration (v0.99.23), Zenodo (2019); <https://doi.org/10.5281/zenodo.3376221>.
67. S. Kumari, J. Nie, H. S. Chen, H. Ma, R. Stewart, X. Li, M. Z. Lu, W. M. Taylor, H. Wei, Evaluation of gene association methods for coexpression network construction and biological knowledge discovery. *PLoS ONE* **7**, e50411 (2012).
68. G. Yu, L. G. Wang, Y. Han, Q. Y. He, clusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).
69. C. Schizophrenia Working Group of the Psychiatric Genomics, Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427 (2014).
70. D. Demontis, R. K. Walters, J. Martin, M. Mattheisen, T. D. Als, E. Agerbo, G. Baldursson, R. Belliveau, J. Bybjerg-Grauholm, M. Baekvad-Hansen, F. Cerrato, K. Chambert, C. Churchhouse, A. Dumont, N. Eriksson, M. Gandalf, J. I. Goldstein, K. L. Grasby, J. Grove, O. O. Gudmundsson, C. S. Hansen, M. E. Hauberg, M. V. Hollegaard, D. P. Howrigan, H. Huang, J. B. Maller, A. R. Martin, N. G. Martin, J. Moran, J. Pallesen, D. S. Palmer, C. B. Pedersen, M. G. Pedersen, T. Poterba, J. B. Poulsen, S. Ripke, E. B. Robinson, F. K. Satterstrom, H. Stefansson, C. Stevens, P. Turley, G. B. Walters, H. Won, M. J. Wright; ADHD Working Group of the Psychiatric Genomics Consortium (PGC); Early Lifecourse & Genetic Epidemiology (EAGLE) Consortium; 23andMe Research Team, O. A. Andreassen, P. Asherson, C. L. Burton, D. I. Boomsma, B. Cormand, S. Dalsgaard, B. Franke, J. Gelernter, D. Geschwind, H. Hakonarson, J. Haavik, H. R. Kranzler, J. Kuntsi, K. Langley, K. P. Lesch, C. Middeldorp, A. Reif, L. A. Rohde, P. Roussos, R. Schachar, P. Sklar, E. J. S. Sonuga-Barke, P. F. Sullivan, A. Thapar, J. Y. Tung, I. D. Waldman, S. E. Medland, K. Stefansson, M. Nordentoft, D. M. Hougaard, T. Werge, O. Mors, P. B. Mortensen, M. J. Daly, S. V. Faraone, A. D. Borglum, B. M. Neale, Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat. Genet.* **51**, 63–75 (2019).
71. J. Grove, S. Ripke, T. D. Als, M. Mattheisen, R. K. Walters, H. Won, J. Pallesen, E. Agerbo, O. A. Andreassen, R. Anney, S. Awasthi, R. Belliveau, F. Bettella, J. D. Buxbaum, J. Bybjerg-Grauholm, M. Baekvad-Hansen, F. Cerrato, K. Chambert, J. H. Christensen, C. Churchhouse, K. Dellennvall, D. Demontis, S. De Rubeis, B. Devlin, S. Djurovic, A. L. Dumont, J. I. Goldstein, C. S. Hansen, M. E. Hauberg, M. V. Hollegaard, S. Hope, D. P. Howrigan, H. Huang, C. M. Hultman, L. Klei, J. Maller, J. Martin, A. R. Martin, J. L. Moran, M. Nyegaard, T. Naerland, D. S. Palmer, A. Palotie, C. B. Pedersen, M. G. Pedersen, T. Poterba, J. B. Poulsen, B. S. Pourcain, P. Qvist, K. Rehnstrom, A. Reichenberg, J. Reichert, E. B. Robinson, K. Roeder, P. Roussos, E. Saemundsen, S. Sandin, F. K. Satterstrom, G. D. Smith, H. Stefansson, S. Steinberg, C. R. Stevens, P. F. Sullivan, P. Turley, G. B. Walters, X. Xu; Autism Spectrum Disorder Working Group of the Psychiatric Genomics Consortium; BUPGEN; Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium; 23andMe Research Team, K. Stefansson, D. H. Geschwind, M. Nordentoft, D. M. Hougaard, T. Werge, O. Mors, P. B. Mortensen, B. M. Neale, M. J. Daly, A. D. Borglum, Identification of common genetic risk variants for autism spectrum disorder. *Nat. Genet.* **51**, 431–444 (2019).
72. E. A. Stahl, G. Breen, A. J. Forstner, A. McQuillin, S. Ripke, V. Trubetskoy, M. Mattheisen, Y. Wang, J. R. I. Coleman, H. A. Gaspar, C. A. de Leeuw, S. Steinberg, J. M. W. Pavlides, M. Trzaskowski, E. M. Byrne, T. H. Pers, P. A. Holmans, A. L. Richards, L. Abbott, E. Agerbo, H. Akil, D. Albani, N. Alliey-Rodriguez, T. D. Als, A. Anjorin, V. Antilla, S. Awasthi, J. A. Badner, M. Baekvad-Hansen, J. D. Barchas, N. Bass, M. Bauer, R. Belliveau, S. E. Bergen, C. B. Pedersen, E. Boen, M. P. Boks, J. Boocock, M. Budde, W. Bunney, M. Burmeister, J. Bybjerg-Grauholm, W. Byerley, M. Casas, F. Cerrato, P. Cervantes, K. Chambert, A. W. Charney, D. Chen, C. Churchhouse, T. K. Clarke, W. Coryell, D. W. Craig, C. Cruceanu, D. Curtis, P. M. Czerski, A. M. Dale, S. de Jong, F. Degenhardt, J. Del-Favero, J. R. DePaulo, S. Djurovic, A. L. Dobbyn, A. Dumont, T. Elvsashagen, V. Escott-Price, C. C. Fan, S. B. Fischer, M. Flickinger, T. M. Foroud, L. Forty, J. Frank, C. Fraser, N. B. Freimer, L. Frisen, K. Gade, D. Gage, J. Garnham, C. Giambartolomei, M. G. Pedersen, J. Goldstein, S. D. Gordon, K. Gordon-Smith, E. K. Green, M. J. Green, M. J. Greenwood, J. Grove, W. Guan, J. Guzman-Parra, M. L. Hamshere, M. Hautzinger, U. Heilbronner, S. Herms, M. Hipolito, P. Hoffmann, D. Holland, L. Huckins, S. Jamain, J. S. Johnson, A. Jureus, R. Kandawaswamy, R. Karlsson, J. L. Kennedy, S. Kittel-Schneider, J. A. Knowles, M. Kogevinas, A. C. Koller, R. Kupka, K. Lavebratt, J. Lawrence, W. B. Lawson, M. Leber, P. H. Lee, S. E. Levy, J. Z. Li, C. Liu, S. Lucae, A. Maaser, D. J. MacIntyre, P. B. Mahon, W. Maier, L. Martinsson, S. McCarrroll, P. McGuffin, M. G. McInnis, J. D. McKay, H. Medeiros, S. E. Medland, F. Meng, L. Milani, G. W. Montgomery, D. W. Morris, T. W. Muhleisen, N. Mullins, H. Nguyen, C. M. Nievergelt, A. N. Adolfsson, E. A. Nwulia, C. O'Donovan, L. M. O. Loohuis, A. P. S. Ori, L. Oruc, U. Osby, R. H. Perlis, A. Perry, A. Pfenning, J. B. Potash, S. M. Purcell, E. J. Regeer, A. Reif, C. S. Reinbold, J. P. Rice, F. Rivas, M. Rivera, P. Roussos, D. M. Ruderfer, E. Ryu, C. Sanchez-Mora, A. F. Schatzberg, W. A. Scheftner, N. J. Schork, C. S. Weickert, T. Shekhtman, P. D. Shilling, E. Sigurdsson, C. Slaney, O. B. Smeland, J. L. Sobell, C. Soholm Hansen, A. T. Spjiker, D. S. Clair, M. Steffens, J. S. Strauss, F. Streit, J. Strohmaier, S. Szlinger, R. C. Thompson, T. E. Thorgerisson, J. Treutlein, H. Vedder, W. Wang, S. J. Watson, T. W. Weickert, S. H. Witt, S. Xi, W. Xu, A. H. Young, P. Zandi, P. Zhang, S. Zollner; eQTLGen Consortium; BIOS Consortium, R. Adolfsson, I. Agartz, M. Alda, L. Backlund, T. B. Baune, F. Bellivier, W. H. Berrettini, J. M. Biernacka, D. H. R. Blackwood, M. Boehnke, A. D. Borglum, A. Corvin, N. Craddock, M. J. Daly, U. Dannlowski, T. Esko, B. Etain, M. Frye, J. M. Fullerton, E. S. Gershon, M. Gill, F. Goes, M. Grigoriou-Serbanescu, J. Hauser, D. M. Hougaard, C. M. Hultman, I. Jones, L. A. Jones, R. S. Kahn, G. Kirov, M. Landen, M. Leboyer, C. M. Lewis, Q. S. Li, J. Lissowska, N. G. Martin, F. Mayoral, S. L. McElroy, A. M. McIntosh, F. J. McMahon, I. Melle, A. Metspalu, P. B. Mitchell, G. Morken, O. Mors, P. B. Mortensen, B. Muller-Myhsok, R. M. Myers, B. M. Neale, V. Nimgaonkar, M. Nordentoft, M. M. Nothen, M. C. O'Donovan, K. J. Oedegaard, M. J. Owen, S. A. Paciga, C. Pato, M. T. Pato, D. Posthuma, J. A. Ramos-Quiroga, M. Ribases, M. Rietschel, G. A. Rouleau, M. Schalling, P. R. Schofield, T. G. Schulze, A. Serretti, J. W. Smoller, H. Stefansson, K. Stefansson, E. Stordal, P. F. Sullivan, G. Turecki, A. E. Vaaler, E. Vieta, J. B. Vincent, T. Werge, J. I. Nurnberger, N. R. Wray, A. Di Florio, H. J. Edenberg, S. Cichon, R. A. Ophoff, L. J. Scott, O. A. Andreassen, J. Kelsoe, P. Sklar; Bipolar Disorder Working Group of the Psychiatric Genomics Consortium, Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat. Genet.* **51**, 793–803 (2019).
73. D. M. Howard, M. J. Adams, T. K. Clarke, J. D. Hafferty, J. Gibson, M. Shirali, J. R. I. Coleman, S. P. Hagenaars, J. Ward, E. M. Wigmore, C. Alloza, X. Shen, M. C. Barbu, E. Y. Xu, H. C. Whalley, R. E. Marioni, D. J. Porteous, G. Davies, I. J. Deary, G. Hemani, K. Berger, H. Teismann, R. Rawal, V. Arolt, B. T. Baune, U. Dannlowski, K. Domschke, C. Tian, D. A. Hinds; 23andMe Research Team; Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium, M. Trzaskowski, E. M. Byrne, S. Ripke, D. J. Smith, P. F. Sullivan, N. R. Wray, G. Breen, C. M. Lewis, A. M. McIntosh, Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nat. Neurosci.* **22**, 343–352 (2019).

74. International Obsessive Compulsive Disorder Foundation Genetics Collaborative (IOCDF-GC) and OCD Collaborative Genetics Association Studies (OCGAS). Revealing the complex genetic architecture of obsessive-compulsive disorder using meta-analysis. *Mol. Psychiatry* **23**, 1181–1188 (2018).
75. C. M. Nievergelt, A. X. Maihofer, T. Klengel, E. G. Atkinson, C. Y. Chen, K. W. Choi, J. R. I. Coleman, S. Dalvie, L. E. Duncan, J. E. Gelernter, D. F. Levey, M. W. Logue, R. Polimanti, A. C. Provest, A. Ratanatharathorn, M. B. Stein, K. Torres, A. E. Aiello, L. M. Almli, A. B. Amstadter, S. B. Andersen, O. A. Andreassen, P. A. Arbisi, A. E. Ashley-Koch, S. B. Austin, E. Avdibegovic, D. Babic, M. Baekvad-Hansen, D. G. Baker, J. C. Beckham, L. J. Bierut, J. I. Bisson, M. P. Boks, E. A. Bolger, A. D. Borglum, B. Bradley, M. Brashear, G. Breen, R. A. Bryant, A. C. Bustamante, J. Bybjerg-Grauholm, J. R. Calabrese, J. M. Caldas-de-Almeida, A. M. Dale, M. J. Daly, N. P. Daskalakis, J. Deckert, D. L. Delahanty, M. F. Dennis, S. G. Disner, K. Domschke, A. Dzubur-Kulenovic, C. R. Erbes, A. Evans, L. A. Farrer, N. C. Feeny, J. D. Flory, D. Forbes, C. E. Franz, S. Galea, M. E. Garrett, B. Gelaye, E. Geuze, C. Gillespie, A. G. Uka, S. D. Gordon, G. Guffanti, R. Hammamieh, S. Harnal, M. A. Hauser, A. C. Heath, S. M. J. Hemmings, D. M. Hougaard, M. Jakovljevic, M. Jett, E. O. Johnson, I. Jones, T. Jovanovic, X. J. Qin, A. G. Junglen, K. I. Karstoft, M. L. Kaufman, R. C. Kessler, A. Khan, N. A. Kimbrel, A. P. King, N. Koen, H. R. Kranzler, W. S. Kremen, B. R. Lawford, L. A. M. Lebois, C. E. Lewis, S. D. Linnstaedt, A. Lori, B. Lugonja, J. J. Luykx, M. J. Lyons, J. Maples-Keller, C. Marmar, A. R. Martin, N. G. Martin, D. Maurer, M. R. Mavissakalian, A. McFarlane, R. E. McGlinchey, K. A. McLaughlin, S. A. McLean, S. McLeay, D. Mehta, W. P. Millberg, M. W. Miller, R. A. Morey, C. P. Morris, O. Mors, P. B. Mortensen, B. M. Neale, E. C. Nelson, M. Nordentoft, S. B. Norman, M. O'Donnell, H. K. Orcutt, M. S. Panizzon, E. S. Peters, A. L. Peterson, M. Peverill, R. H. Pietrzak, M. A. Polusny, J. P. Rice, S. Ripke, V. B. Risbrough, A. L. Roberts, A. O. Rothbaum, B. O. Rothbaum, P. Roy-Byrne, K. Ruggiero, A. Rung, B. P. F. Rutten, N. L. Saccone, S. E. Sanchez, D. Schijven, S. Seedat, A. V. Seligowski, J. S. Seng, C. M. Sheerin, D. Silove, A. K. Smith, J. W. Smoller, S. R. Spohnheim, D. J. Stein, J. S. Stevens, J. A. Sumner, M. H. Teicher, W. K. Thompson, E. Trapido, M. Uddin, R. J. Ursano, L. L. van den Heuvel, M. Van Hooff, E. Vermetten, C. H. Vinkers, J. Voisey, Y. Wang, Z. Wang, T. Werge, M. A. Williams, D. E. Williamson, S. Winternitz, C. Wolf, E. J. Wolf, J. D. Wolff, R. Yehuda, R. M. Young, K. A. Young, H. Zhao, L. A. Zoellner, I. Liberzon, K. J. Ressler, M. Haas, K. C. Koenen. International meta-analysis of PTSD genome-wide association studies identifies sex- and ancestry-specific genetic risk loci. *Nat. Commun.* **10**, 4558 (2019).
76. A. Erlangsen, V. Appadurai, Y. Wang, G. Turecki, O. Mors, T. Werge, P. B. Mortensen, A. Starnawska, A. D. Borglum, A. Schork, R. Nudel, M. Baekvad-Hansen, J. Bybjerg-Grauholm, D. M. Hougaard, W. K. Thompson, M. Nordentoft, E. Agerbo. Genetics of suicide attempts in individuals with and without mental disorders: a population-based genome-wide association study. *Mol. Psychiatry* **25**, 2410–2421 (2020).
77. I. E. Jansen, J. E. Savage, K. Watanabe, J. Bryois, D. M. Williams, S. Steinberg, J. Sealock, I. K. Karlsson, S. Hagg, L. Athanasiu, N. Voyle, P. Proitsi, A. Witoelar, S. Stringer, D. Aarsland, I. S. Almdahl, F. Andersen, S. Bergh, F. Bettella, S. Bjornsson, A. Braekhus, G. Brathen, C. de Leeuw, R. S. Desikan, S. Djurovic, L. Dumitrescu, T. Fladby, T. J. Hohman, P. V. Jonsson, S. J. Kiddle, A. Rongve, I. Saltvedt, S. B. Sando, G. Selbaek, M. Shoen, N. G. Skene, J. Snaedal, E. Stordal, I. D. Ulstein, Y. Wang, L. R. White, J. Hardy, J. Hjerling-Leffler, P. F. Sullivan, W. M. van der Flier, R. Dobson, L. K. Davis, H. Stefansson, K. Stefansson, N. L. Pedersen, S. Ripke, O. A. Andreassen, D. Posthuma. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat. Genet.* **51**, 404–413 (2019).
78. W. van Rheenen, A. Shatunov, A. M. Dekker, R. L. McLaughlin, F. P. Dijkstra, S. L. Pulit, R. A. van der Spek, U. Vosa, S. de Jong, M. R. Robinson, J. Yang, I. Fogh, P. T. van Doormaal, G. H. Tazelaar, M. Koppers, A. M. Blokhuis, W. Sproviero, A. R. Jones, K. P. Kenna, K. R. van Eijk, O. Harschnitz, R. D. Schellevis, W. J. Brands, J. Medic, A. Menelaou, A. Vajda, N. Ticozzi, K. Lin, B. Rogelj, K. Vrabec, M. Ravnik-Glavac, B. Koritnik, J. Zidar, L. Leonards, L. D. Groseelj, S. Millecamps, F. Salachas, V. Meisinger, M. de Carvalho, S. Pinto, J. S. Mora, R. Rojas-Garcia, M. Polak, S. Chandran, S. Colville, R. Swingler, K. E. Morrison, P. J. Shaw, J. Hardy, R. W. Orrell, A. Pittman, K. Sidle, P. Fratta, A. Malaspina, S. Topp, S. Petri, S. Abdulla, C. Drepper, M. Sendtner, T. Meyer, R. A. Ophoff, K. A. Staats, M. Wiedau-Pazos, C. Lomen-Hoerth, V. M. Van Deerlin, J. Q. Trojanowski, L. Elman, L. McCluskey, A. N. Basak, C. Tunca, H. Hamzeiy, Y. Parman, T. Meitinger, P. Lichtner, M. Radivojkov-Bлагоjevic, C. R. Andres, C. Maurel, G. Bensimon, B. Landwehrmeyer, A. Brice, C. A. Payan, S. Saker-Delye, A. Durr, N. W. Wood, L. Tittmann, W. Lieb, A. Franke, M. Rietschel, S. Cichon, M. M. Nothen, P. Amouyel, C. Tzourio, J. F. Dartigues, A. G. Uitterlinden, F. Rivadeneira, K. Estrada, A. Hofman, C. Curtis, H. M. Blauw, A. J. van der Kooij, M. de Visser, A. Goris, M. Weber, C. E. Shaw, B. N. Smith, O. Pansarasa, C. Cereda, R. Del Bo, G. P. Comi, S. D'Alfonso, C. Bertolin, G. Soraru, L. Mazzini, V. Pensato, C. Gellera, C. Tiloca, A. Ratti, A. Calvo, C. Moglia, M. Brunetti, S. Arcuti, R. Capozzo, C. Zecca, C. Lunetta, S. Penco, N. Riva, A. Padovani, M. Filosto, B. Muller, R. J. Stuit; PARALS Registry; SLALOM Group; SLAP Registry; FALS Sequencing Consortium; SLAGEN Consortium; NNIPPS Study Group, I. Blair, K. Zhang, E. P. McCann, J. A. Fifta, G. A. Nicholson, D. B. Rowe, R. Pamphlett, M. C. Kiernan, J. Grosskreutz, O. W. Witte, T. Ringer, T. Prell, B. Stubendorff, I. Kurth, C. A. Hubner, P. N. Leigh, F. Casale, A. Chio, E. Beghi, E. Pupillo, R. Tortelli, G. Logroscino, J. Powell, A. C. Ludolph, J. H. Weishaupt, W. Robberecht, P. Van Damme, L. Franke, T. H. Pers, R. H. Brown, J. D. Glass, J. E. Landers, O. Hardiman, P. M. Andersen, P. Corcia, P. Vourc'h, V. Silani, N. R. Wray, P. M. Visscher, P. I. de Bakker, M. A. van Es, R. J. Pasterkamp, C. M. Lewis, G. Breen, A. Al-Chalabi, L. H. van den Berg, J. H. Veldink. Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis. *Nat. Genet.* **48**, 1043–1048 (2016).
79. T. F. Andlauer, D. Buck, G. Antony, A. Bayas, L. Bechmann, A. Berthele, A. Chan, C. Gasperi, R. Gold, C. Graetz, J. Haas, M. Hecker, C. Infante-Duarte, M. Knop, T. Kumpfel, V. Limmroth, R. A. Linker, V. Loleit, F. Luessi, S. G. Meuth, M. Muhlau, S. Nischwitz, F. Paul, M. Putz, T. Ruck, A. Salmen, M. Stangel, J. P. Stellmann, K. H. Sturmer, B. Tackenberg, F. Then Bergh, H. Tumani, C. Warnke, F. Weber, H. Wiendl, B. Wildemann, U. K. Zettl, U. Ziemann, F. Zipp, J. Arloth, P. Weber, M. Radivojkov-Bлагоjevic, M. O. Scheinhardt, T. Dankowski, T. Bettecken, P. Lichtner, D. Czamara, T. Carrillo-Roa, E. B. Binder, K. Berger, L. Bertram, A. Franke, C. Gieger, S. Herms, G. Homuth, M. Ising, K. H. Jockel, T. Kacprowski, S. Kloiber, M. Laudes, W. Lieb, C. M. Lill, S. Lucae, T. Meitinger, S. Moebus, M. Muller-Nurasyid, M. M. Nothen, A. Petersmann, R. Rawal, U. Schminke, K. Strauch, H. Volzke, M. Waldenberger, J. Wellmann, E. Porcu, A. Mulas, M. Pitzalis, C. Sidore, I. Zara, F. Cucca, M. Zoledziwska, A. Ziegler, B. Hemmer, B. Muller-Myhok. Novel multiple sclerosis susceptibility loci implicated in epigenetic regulation. *Sci. Adv.* **2**, e1501678 (2016).
80. M. A. Nalls, C. Blauwendraat, C. L. Vallerga, K. Heilbron, S. Bandres-Ciga, D. Chang, M. Tan, D. A. Kia, A. J. Noyce, A. Xue, J. Bras, E. Young, R. von Coelln, J. Simon-Sanchez, C. Schulte, M. Sharma, L. Krohn, L. Pihlstrom, A. Siitonen, H. Iwaki, H. Leonard, F. Faghri, J. R. Gibbs, D. G. Hernandez, S. W. Scholz, J. A. Botia, M. Martinez, J. C. Corvol, S. Lesage, J. Jankovic, L. M. Shulman, M. Sutherland, P. Tienari, K. Majamaa, M. Toft, O. A. Andreassen, T. Bangale, A. Brice, J. Yang, Z. Gan-Or, T. Gasser, P. Heutink, J. M. Shulman, N. W. Wood, D. A. Hinds, J. A. Hardy, H. R. Morris, J. Gratten, P. M. Visscher, R. R. Graham, A. B. Singleton; 23andMe Research Team; System Genomics of Parkinson's Disease Consortium; International Parkinson's Disease Genomics Consortium. Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: A meta-analysis of genome-wide association studies. *Lancet Neurol.* **18**, 1091–1102 (2019).
81. J. Z. Liu, S. van Sommeren, H. Huang, S. C. Ng, R. Alberts, A. Takahashi, S. Ripke, J. C. Lee, L. Jostins, T. Shah, S. Abedian, J. H. Cheon, J. Cho, N. E. Dayani, L. Franke, Y. Fuyuno, A. Hart, R. C. Juyal, G. Juyal, W. H. Kim, A. P. Morris, H. Poustchi, W. G. Newman, V. Midha, T. R. Orchard, H. Vahedi, A. Sood, J. Y. Sung, R. Malekzadeh, H. J. Westra, K. Yamazaki, S. K. Yang; International Multiple Sclerosis Genetics Consortium; International IBD Genetics Consortium, J. C. Barrett, B. Z. Alizadeh, M. Parkes, T. B. K. M. J. Daly, M. Kubo, C. A. Anderson, R. K. Weersma. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat. Genet.* **47**, 979–986 (2015).
82. Y. Okada, D. Wu, G. Trynka, T. Raj, C. Terao, K. Ikari, Y. Kochi, K. Ohmura, A. Suzuki, S. Yoshida, R. R. Graham, A. Manoharan, W. Ortman, T. Bhargale, J. C. Denny, R. J. Carroll, A. E. Eyler, J. D. Greenberg, J. M. Kremer, D. A. Pappas, L. Jiang, J. Yin, L. Ye, D. F. Su, J. Yang, G. Xie, E. Keystone, H. J. Westra, T. Esko, A. Metspalu, X. Zhou, N. Gupta, D. Mirel, E. A. Stahl, D. Diogo, J. Cui, K. Liao, M. H. Guo, K. Myouzen, T. Kawaguchi, M. J. Coenen, P. L. van Riel, M. A. van de Laar, H. J. Guchelaar, T. W. Huizinga, P. Dieude, X. Mariette, S. L. Bridges Jr., A. Zernakova, R. E. Toes, P. P. Tak, C. Miceli-Richard, S. Y. Bang, H. S. Lee, J. Martin, M. A. Gonzalez-Gay, L. Rodriguez-Rodriguez, S. Rantapaa-Dahlqvist, L. Arlestig, H. K. Choi, Y. Kamatani, P. Galan, M. Lathrop; RACI consortium; GARNET consortium, S. Eyre, J. Bowes, A. Barton, N. de Vries, L. W. Moreland, L. A. Criswell, E. W. Karlson, A. Taniguchi, R. Yamada, M. Kubo, J. S. Liu, S. C. Bae, J. Worthington, L. Padyukov, L. Klareskog, P. K. Gregersen, S. Raychaudhuri, B. E. Stranger, P. L. De Jager, L. Franke, P. M. Visscher, M. A. Brown, H. Yamanaka, T. Mimori, A. Takahashi, H. Xu, T. W. Behrens, K. A. Siminovitch, S. Momohara, F. Matsuda, K. Yamamoto, R. M. Plenge. Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature* **506**, 376–381 (2014).
83. N. G. Skene, J. Bryois, T. E. Bakken, G. Breen, J. J. Crowley, H. A. Gaspar, P. Giusti-Rodriguez, R. D. Hodge, J. A. Miller, A. B. Munoz-Manchado, M. C. O'Donovan, M. J. Owen, A. F. Pardinas, J. Ryge, J. T. R. Walters, S. Linnarsson, E. S. Lein; Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium, P. F. Sullivan, J. Hjerling-Leffler. Genetic identification of brain cell types underlying schizophrenia. *Nat. Genet.* **50**, 825–833 (2018).
84. N. Habib, I. Avraham-Davidi, A. Basu, T. Burks, K. Shekhar, M. Hofree, S. R. Choudhury, F. Aguet, E. Gelfand, K. Ardlie, D. A. Weitz, O. Rozenblatt-Rosen, F. Zhang, A. Regev. Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nat. Methods* **14**, 955–958 (2017).
85. M. E. Ritchie, B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi, G. K. Smyth. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
86. P. D. Thomas, M. J. Campbell, A. Kejariwal, H. Mi, B. Karlak, R. Daverman, K. Diemer, A. Muruganujan, A. Narechania. PANTHER: A library of protein families and subfamilies indexed by function. *Genome Res.* **13**, 2129–2141 (2003).

87. L. S. Hall, C. W. Medway, O. Pain, A. F. Pardini, E. G. Rees, V. Escott-Price, A. Pocklington, N. J. Bray, P. A. Holmans, J. T. R. Walters, M. J. Owen, M. C. O'Donovan, A transcriptome-wide association study implicates specific pre- and post-synaptic abnormalities in schizophrenia. *Hum. Mol. Genet.* **29**, 159–167 (2020).
88. A. M. M. Sousa, Y. Zhu, M. A. Raghanti, R. R. Kitchen, M. Onorati, A. T. N. Tebbenkamp, B. Stutz, K. A. Meyer, M. Li, Y. I. Kawasawa, F. Liu, R. G. Perez, M. Mele, T. Carvalho, M. Skarica, F. O. Gulden, M. Pletikos, A. Shalbyata, A. R. Stephenson, M. K. Edler, J. J. Ely, J. D. Elsworth, T. L. Horvath, P. R. Hof, T. M. Hyde, J. E. Kleinman, D. R. Weinberger, M. Reimers, R. P. Lifton, S. M. Mane, J. P. Noonan, M. W. State, E. S. Lein, J. A. Knowles, T. Marques-Bonet, C. C. Sherwood, M. B. Gerstein, N. Sestan, Molecular and cellular reorganization of neural circuits in the human lineage. *Science* **358**, 1027–1032 (2017).
89. E. Hannon, E. Dempster, J. Viana, J. Burrage, A. R. Smith, R. Macdonald, D. St Clair, C. Mustard, G. Breen, S. Therman, J. Kaprio, T. Touloupoulou, H. E. Hulshoff Pol, M. M. Bohlken, R. S. Kahn, I. Nenadic, C. M. Hultman, R. M. Murray, D. A. Collier, N. Bass, H. Gurling, A. McQuillin, L. Schalkwyk, J. Mill, An integrated genetic-epigenetic analysis of schizophrenia: Evidence for co-localization of genetic associations and differential DNA methylation. *Genome Biol.* **17**, 176 (2016).
90. A. E. Jaffe, Y. Gao, A. Deep-Soboslay, R. Tao, T. M. Hyde, D. R. Weinberger, J. E. Kleinman, Mapping DNA methylation across development, genotype and schizophrenia in the human frontal cortex. *Nat. Neurosci.* **19**, 40–47 (2016).
91. M. Kinoshita, S. Numata, A. Tajima, K. Ohi, R. Hashimoto, S. Shimodera, I. Imoto, M. Takeda, T. Ohmori, Aberrant DNA methylation of blood in schizophrenia by adjusting for estimated cellular proportions. *Neuromolecular Med.* **16**, 697–703 (2014).
92. C. Montano, M. A. Taub, A. Jaffe, E. Briem, J. I. Feinberg, R. Trygvadottir, A. Idrizi, A. Runarsson, B. Berndsen, R. C. Gur, T. M. Moore, R. T. Perry, D. Fugman, S. Sabuncyan, R. H. Yolken, T. M. Hyde, J. E. Kleinman, J. L. Sobell, C. N. Pato, M. T. Pato, R. C. Go, V. Nimgaonkar, D. R. Weinberger, D. Braff, R. E. Gur, M. D. Fallin, A. P. Feinberg, Association of DNA methylation differences with schizophrenia in an epigenome-wide association study. *JAMA Psychiatry* **73**, 506–514 (2016).
93. S. Numata, T. Ye, M. Herman, B. K. Lipska, DNA methylation changes in the postmortem dorsolateral prefrontal cortex of patients with schizophrenia. *Front. Genet.* **5**, 280 (2014).
94. L. F. Wockner, E. P. Noble, B. R. Lawford, R. M. Young, C. P. Morris, V. L. Whitehall, J. Voisey, Genome-wide DNA methylation analysis of human brain tissue from schizophrenia patients. *Transl. Psychiatry* **4**, e339 (2014).
95. C. Finan, A. Gaulton, F. A. Kruger, R. T. Lumbers, T. Shah, J. Engmann, L. Galver, R. Kelley, A. Karlsson, R. Santos, J. P. Overington, A. D. Hingorani, J. P. Casas, The druggable genome and support for target identification and validation in drug development. *Sci. Transl. Med.* **9**, (2017).
96. R. Santos, O. Ursu, A. Gaulton, A. P. Bento, R. S. Donadi, C. G. Bologa, A. Karlsson, B. Al-Lazikani, A. Hersey, T. I. Oprea, J. P. Overington, A comprehensive map of molecular drug targets. *Nat. Rev. Drug Discov.* **16**, 19–34 (2017).
97. T. Sheils, S. L. Mathias, V. B. Siramshetty, G. Bocci, C. G. Bologa, J. J. Yang, A. Waller, N. Southall, D. T. Nguyen, T. I. Oprea, How to illuminate the druggable genome using Pharos. *Curr. Protoc. Bioinformatics* **69**, e92 (2020).
98. Q. Wang, R. Chen, F. Cheng, Q. Wei, Y. Ji, H. Yang, X. Zhong, R. Tao, Z. Wen, J. S. Sutcliffe, C. Liu, E. H. Cook, N. J. Cox, B. Li, A Bayesian framework that integrates multi-omics data and gene networks predicts risk genes from schizophrenia GWAS data. *Nat. Neurosci.* **22**, 691–699 (2019).
99. M. Lek, K. J. Karczewski, E. V. Minikel, K. E. Samocha, E. Banks, T. Fennell, A. H. O'Donnell-Luria, J. S. Ware, A. J. Hill, B. Cummings, T. Tukiainen, D. P. Birnbaum, J. A. Kosmicki, L. E. Duncan, K. Estrada, F. Zhao, J. Zou, E. Pierce-Hoffman, J. Berghout, D. N. Cooper, N. DeFlaux, M. DePristo, R. Do, J. Flannick, M. Fromer, L. Gauthier, J. Goldstein, N. Gupta, D. Howrigan, A. Kiezun, M. I. Kurki, A. L. Moonshine, P. Natarajan, L. Orozco, G. M. Peloso, R. Poplin, M. A. Rivas, V. Ruano-Rubio, S. A. Rose, D. M. Ruderfer, K. Shakir, P. D. Stenson, C. Stevens, B. P. Thomas, G. Tiao, M. T. Tusie-Luna, B. Weisburd, H. H. Won, D. Yu, D. M. Altshuler, D. Ardissono, M. Boehnke, J. Danesh, S. Donnelly, R. Elosua, J. C. Florez, S. B. Gabriel, G. Getz, S. J. Glatt, C. M. Hultman, S. Kathiresan, M. Laakso, S. McCarrroll, M. I. McCarthy, D. McGovern, R. McPherson, B. M. Neale, A. Palotie, S. M. Purcell, D. Saleheen, J. M. Scharf, P. Sklar, P. F. Sullivan, J. Tuomilehto, M. T. Tsuang, H. C. Watkins, J. G. Wilson, M. J. Daly, D. G. MacArthur, Exome Aggregation Consortium, Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016).
100. M. B. Kursu, W. R. Rudnicki, Feature selection with the Boruta package. *J. Stat. Softw.* **36**, 1–13 (2010).
101. J. A. Miller, S. L. Ding, S. M. Sunkin, K. A. Smith, L. Ng, A. Szafer, A. Ebbert, Z. L. Riley, J. J. Royall, K. Aiona, J. M. Arnold, C. Bennet, D. Bertagnolli, K. Brunner, S. Butler, S. Caldejon, A. Carey, C. Cuhaciyan, R. A. Dalley, N. Dee, T. A. Dolbeare, B. A. Facer, D. Feng, T. P. Fliss, G. Gee, J. Goldy, L. Gourley, B. W. Gregor, G. Gu, R. E. Howard, J. M. Jochim, C. L. Kuan, C. Lau, C. K. Lee, F. Lee, T. A. Lemon, P. Lesnar, B. McMurray, N. Mastan, N. Mosqueda, T. Nalua-Cecchini, N. K. Ngo, J. Nyhus, A. Oldre, E. Olson, J. Parente, P. D. Parker, S. E. Parry, A. Stevens, M. Pletikos, M. Reding, K. Roll, D. Sandman, M. Sraael, S. Shapouri, N. V. Shapovalova, E. H. Shen, N. Sjoquist, C. R. Slaughterbeck, M. Smith, A. J. Sodt, D. Williams, L. Zollei, B. Fischl, M. B. Gerstein, D. H. Geschwind, I. A. Glass, M. J. Hawrylycz, R. F. Hevner, H. Huang, A. R. Jones, J. A. Knowles, P. Levitt, J. W. Phillips, N. Sestan, P. Wohnoutka, C. Dang, A. Bernard, J. G. Hohmann, E. S. Lein, Transcriptional landscape of the prenatal human brain. *Nature* **508**, 199–206 (2014).
102. Z. Gu, R. Eils, M. Schlesner, Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016).
103. P. Langfelder, S. Horvath, WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).
104. C. A. de Leeuw, J. M. Mooij, T. Heskes, D. Posthuma, MAGMA: Generalized gene-set analysis of GWAS data. *PLOS Comput. Biol.* **11**, e1004219 (2015).
105. S. Durinck, P. T. Spellman, E. Birney, W. Huber, Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* **4**, 1184–1191 (2009).
106. R Core Team, R: A language and environment for statistical computing (R Foundation for Statistical Computing, Vienna, Austria, 2018); available online at <https://www.R-project.org/>.

Acknowledgments: We are grateful for the contributions of the Office of the Chief Medical Examiner of the State of Maryland, Office of the Chief Medical Examiner of Kalamazoo County Michigan, Office of the Chief Medical Examiner University of North Dakota School of Medicine, Gift of Life of Michigan, Office of the Chief Medical Examiner of Santa Clara County California, and Medical University of Sofia, Bulgaria in assisting the LIBD in the acquisition and curation of brain tissue donations for this study. All research at the LIBD is made possible by gifts from the families of S. and C. Lieber and M. and T. Maltz. We would like to thank all the family members of the donors for exceptional contribution. F.R. is a PhD student enrolled in the National PhD in Artificial Intelligence, XXXVII cycle, course on Health and Life Sciences, organized by Università Campus Bio-Medico di Roma (Italy). We are in debt to A. Jaffe, who has contributed to this work by offering data and software, along with many key methodological insights. We are grateful to P. Di Carlo, P. Di Palo, A. Gaudio, A. Monda, M. Papalino, N. Penzel, and F. Piarulli for exploratory analyses and computational tools used in this work. We thank R. Narurkar, who has given our team unflinching support to realize this work. S. Han, V. S. Mattay, R. McKay, and R. Straub have provided invaluable exchanges of ideas that have influenced this work, and the authors express the highest gratitude for their contribution. We express our gratitude to G. Thomas for contribution in generating Fig. 1 included in this article. Replication data were generated as part of the PsychENCODE Consortium, supported by U01DA048279, U01MH103339, U01MH103340, U01MH103346, U01MH103365, U01MH103392, U01MH116438, U01MH116441, U01MH116442, U01MH116488, 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, R56MH114899, R56MH114901, R56MH114911, R01MH125516, and P50MH106934 awarded to: A. Abzyov, N. Ahitov, S. Akbarian, A. Arguello, L. Bingham, K. Brennan, A. Chess, G. Cooper, G. Crawford, S. Dracheva, P. Farnham, M. Gerstein, D. Geschwind, F. Goes, V. Haroutunian, T. M. Hyde, A. Jaffe, P. Jin, M. Kellis, J. Kleinman, J. A. Knowles, A. Kriegstein, C. Liu, K. Martinowich, E. Mukamel, R. Myers, C. Nemeroff, M. Peters, D. Pinto, K. Pollard, K. Ressler, P. Roussos, S. Sanders, N. Sestan, P. Sklar, N. Sokol, M. State, J. Stein, P. Sullivan, F. Vaccarino, S. Warren, D. Weinberger, S. Weissman, Z. Weng, K. White, A. J. Willsey, H. Won, and P. Zandi. **Funding:** This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 798181 awarded to G.P. (PI), and A.B., D.R.W. M.P., L.S., R.B. were partially supported by NIH grant 5R21MH117432-02 awarded to D.R.W. (PI), G.P., R.B., and A.B.. This research has been partially supported by the project "Dopamine - dysbindin genetic interaction: a multidisciplinary approach to characterize cognitive phenotypes of schizophrenia and develop personalized treatments" (PRIN: Progetti di Ricerca di Rilevante Interesse Nazionale-Bando 2017 Prot. 2017K2NEF4) awarded to G.P., the funding initiative Horizon Europe Seeds 2021 (Next Generation EU-MUR D.M. 737/2021) for the project S68 to G.P., EXPRIVA for the research program: "Artificial intelligence, genetics and transcriptomics" to G.P., and the Apulian regional government for the project: "Early Identification of Psychosis Risk" to A.B., G.P. and L.B. have also obtained funding for this work under the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.4- Call for tender no. 3138 of 16 December 2021 of Italian Ministry of University and Research funded by the European Union-NextGenerationEU, award number: Project code: CN00000013, Concession Decree No. 1031 of 17 February 2022 adopted by the Italian Ministry of University and Research, CUP: D93C22000430001, Project title: "National Centre for HPC, Big Data and Quantum Computing". The LIBD supported tissue collection and maintenance, analysis, infrastructure, and personnel. **Author contributions:** Conceptualization: G.P., D.R.W., and A.B. Data curation: M.P., L.S., C.B., F.R., Y.W., S.R.S., and J.H.S. Formal analysis: G.P., M.P., L.S., C.B., F.R., L.B., and E.R. Funding acquisition: G.P., R.B., G.B., A.M., D.R.W., and A.B. Investigation: G.P., E.R., B.J.M., J.H.S., A.B., and D.R.W. Methodology: G.P., L.S., C.B., E.R., L.B., Q.C., and B.J.M. Project

administration: G.P. Resources: R.B., Q.C., J.E.K., Y.W., S.R.S., B.J.M., T.M.H., and D.R.W. Software: M.P., L.S., C.B., L.B., and F.R. Supervision: G.P., B.J.M., A.B., and D.R.W. Validation: G.P., L.B., and F.R. Visualization: M.P., L.S., C.B., E.R., and L.B. Writing (original draft): G.P., M.P., L.S., C.B., and D.R.W. Writing (review and editing): all authors. **Competing interests:** A.B. received consulting fees by Biogen and lecture fees by Otsuka, Janssen, Lundbeck. G.B. received lecture fees by Lundbeck. D.R.W. serves on the Scientific Advisory Boards of Sage Therapeutics and Pasithea Therapeutics. All other authors declare no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Preprocessed data files, WGCNA network information, and connectivity data along with interactive Sankey plots is publicly available at: <https://doi.org/10.5281/zenodo.5676480>.

5281/zenodo.5676480. This paper also analyzes existing, publicly available data. The accession numbers for the datasets are listed in the key resource table in the Supplementary Materials. The code used for this study is hosted in a GitHub repository and is publicly available at https://github.com/LieberInstitute/Brain_WGCNA. We will also continue to maintain the data at https://nets.libd.org/age_wgcna with possible updates on these research streams.

Submitted 6 August 2022

Accepted 10 March 2023

Published 14 April 2023

10.1126/sciadv.ade2812