BMC Genomics

RESEARCH Open Access

Genomic insights into genes expressed specifically during infancy highlight their dominant influence on the neuronal system

Weidi Wang[1†], Zhe Liu[1†], Daihui Peng[1], Guan Ning Lin[1*] and Zhen Wang[1,2,3*]

Abstract

Background Elucidating the dynamics of gene expression across developmental stages, including the genomic characteristics of brain expression during infancy, is pivotal in deciphering human psychiatric and neurological disorders and providing insights into developmental disorders.

Results Leveraging comprehensive human GWAS associations with temporal and spatial brain expression data, we discovered a distinctive co-expression cluster comprising 897 genes highly expressed specifically during infancy, enriched in functions related to the neuronal system. This gene cluster notably harbors the highest ratio of genes linked to psychiatric and neurological disorders. Through computational analysis, MYT1L emerged as a potential central transcription factor governing these genes. Remarkably, the infancy-specific expressed genes, including *SYT1*, exhibit prominent colocalization within human accelerated regions. Additionally, chromatin state analysis unveiled prevalent epigenetic markers associated with enhancer-specific modifications. In addition, this cluster of genes has demonstrated to be specifically highly expressed in cell-types including excitatory neurons, medial ganglionic eminence and caudal ganglionic eminence.

Conclusions This study comprehensively characterizes the genomics and epigenomics of genes specifically expressed during infancy, identifying crucial hub genes and transcription factors. These findings offer valuable insights into early detection strategies and interventions for psychiatric and neurological disorders.

Keywords Infancy, Psychiatric and neurological disorders disorders, Co-expression clustering, Evolutionary conservation

[†]Weidi Wang and Zhe Liu contributed equally to this work.

*Correspondence:
Guan Ning Lin
nickgnlin@sjtu.edu.cn
Zhen Wang
wangzhen@smhc.org.cn
[1]Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine, Shanghai 200030, China
[2]Shanghai Key Laboratory of Psychotic Disorders, Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine, Shanghai 201108, China
[3]Shanghai Intelligent Psychological Evaluation and Intervention Engineering Technology Research Center, Shanghai 200030, China

Background

Infancy, which spans from birth to 18 months [1], is a critical phase for accelerated brain development. By the end of this phase, the brain reaches 80–90% of the adult brain volume [2]. Cellular events drive the rapid growth of the infant brain, resulting in structural changes and the reorganization of neural circuits [3]. At the cognitive and behavioral levels during this stage, various abilities such as language, memory, social cognition, emotional regulation, and executive functions emerge and continuously develop [4]. Brain development during this period is highly susceptible to both genetic and non-genetic factors, which can lead to neuronal degeneration, reduced integrity of white matter fiber bundles, and insufficient synaptic pruning [5, 6]. These deficits can result in long-term cognitive, behavioral, and emotional problems, increasing the risk of developing psychiatric disorders [7].

For psychiatric disorders, there are a complex set of conditions that can be particularly challenging to diagnose and manage due to their atypical manifestations, phenotypic heterogeneity, and sometimes idiopathic nature [8]. The existing diagnostic systems for mental disorders (Diagnostic and Statistical Manual of Mental Disorders, 5th Edition, DSM-5) are based on clinical phenomenological classification, with diagnostic criteria that are relatively vague and overlapping, resulting in poor diagnostic consistency. In addition, for infants or toddlers, it is much more difficult to make a diagnosis because the behavioral impairments of psychiatric disorders normally show after the age of four years old [9, 10]. Early identification of abnormal blood biomarkers during infancy or early childhood can serve as a warning sign for susceptibility to psychiatric and neurological disorders in later life. For instance, accumulating evidence links an altered plasma proteome and metabolome with mental health issues in young individuals [11]. Developing genetic markers for psychiatric disorders is essential for earlier and more effective interventions, understanding the biological basis of these conditions, and enabling personalized treatment strategies that improve patient outcomes.

With the advancement of sequencing technologies, Omics-based approaches have become feasible for identifying risk genes or variants associated with psychiatric and neurological disorders. For instance, through genome-wide association studies (GWAS), Trubetskoy et al. [12] identified 120 genes associated with schizophrenia, 106 of which are protein-coding genes, such as the glutamate receptor subunit *GRIN2A* and the transcription factor *SP4*. These schizophrenia-associated genes are enriched in brain neurons and are predominantly involved in physiological processes like synaptic organization, differentiation, and transmission. Another study utilizing exome sequencing identified *AKAP11* as a risk gene for both bipolar disorder and schizophrenia [13]. However, while GWAS-based methods have provided a wealth of risk biomarkers for psychiatric and neurological disorders in adults, identifying genetic biomarkers for abnormal developmental stages in infancy has been challenging due to difficulties in obtaining population-scale phenotypic data. Moreover, it is not feasible to guide diagnoses in a spatial and temporal manner with this approach. Genes specifically expressed in different brain regions at different stages normally act distinctly in human brain development, especially during infancy when the brain is undergoing significant development [14]. Previous research has generated large bulk and single-cell transcriptomic datasets from the brain at different developmental stages [1, 15]. By integrating bulk transcriptomic data, co-expression networks could be constructed. Utilizing the co-expression information, researchers can prioritize GWAS hits that are more likely to be biologically relevant. For instance, if the GWAS variant is located near a gene that is highly co-expressed with other genes known to be associated with the disease, it may be considered a stronger candidate for further investigation. In addition, by identifying gene modules associated with specific traits in the GWAS, researchers can investigate whether these modules are enriched for particular biological functions or pathways, providing insights into the mechanism of the trait. In addition, by linking GWAS data to single-cell expression profiles, researchers can determine if genetic associations are driven by particular cell types, enhancing the biological relevance of GWAS findings. The large-scale transcriptomic data provided complementary power to GWASs in identification of psychiatric and neurological disorder genes and determination of the cell-type.

To identify genetic biomarkers for abnormal developmental stages during infancy, this study leveraged current comprehensive GWAS disorder associations and temporal and spatial expression datasets to pinpoint potential risk genes for psychiatric and neurological disorders (Fig. 1). We discovered a gene cluster with a specific high expression pattern during infancy and notably annotated it using known human psychiatric and neurological disorders risk loci from the GWAS and family *de novo* mutation databases. For the identified potential early risk genes, in silico hub transcription factor (TF) prediction, chromatin state comparisions and overlappings with human accelerated regions were performed. Further research on this risk gene set revealed that these genes are also highly expressed and conserved in rats during infancy and respond to trauma in infancy. Additionally, based on single-cell data from the brain at different developmental stages, we excavated the cell types of these genes, finding that they are highly expressed in

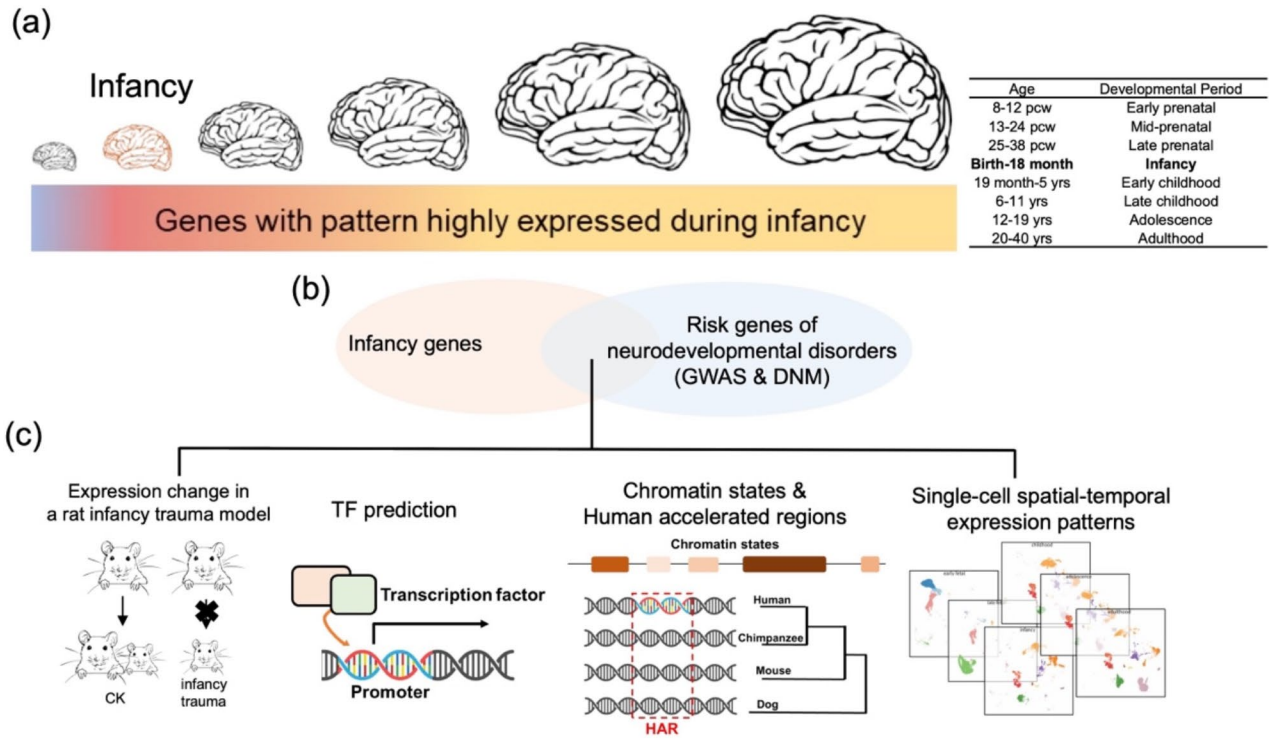

Fig. 1 Outline of this study. (**a**) Using temporal gene expression patterns from the BrainSpan project, a cluster of genes with a distinctly high expression pattern during infancy was identified. The list on the right shows the developmental stages of the brain in this study, as cited from the supplementary materials of Kang et al. [1]; (**b**) This cluster overlaps with the list of risk genes for neurodevelopmental disorders identified through public GWAS and DNM databases; (**c**) For these overlapping genes, further in-depth studies were conducted. The investigations, from left to right, include: the examination of expression changes in a rat infancy trauma model, with a comparison between rats that experienced maternal separation and their standard-reared littermates (CK); the prediction of transcription factors (TFs) that may potentially bind to these genes; the characterization of their chromatin states and the examination of whether any of the genes are colocalized with human accelerated regions; Additionally, the spatial and temporal expression patterns of the gene cluster were investigated

excitatory neurons, the medial ganglionic eminence, and the caudal ganglionic eminence during infancy. This study provides a valuable repertoire of genes associated with early mental disorders and insights for the early diagnosis of psychiatric and neurological disorder risks.

Methods

Clustering of co-expressed genes

We utilized publicly available brain gene expression data obtained from the Allen Institute for Brain Science [1] in our study. This data included information from various age groups, spanning prenatal and postnatal stages. After acquiring the expression data matrix, we filtered out gene entries with expression values of 0 and retained the remaining genes. To ensure appropriate grouping of data and align with our research focus on infancy, we divided the dataset into eight distinct developmental stages. These stages were determined based on the donor's age and developmental stage at the time of sample collection. The sample groups were categorized as follows: early prenatal, mid-prenatal, late prenatal, infancy, early childhood, late childhood, adolescence, and adulthood. Supplementary Table 1 provides more detailed information on these developmental stages. We employed Clust [16] to identify co-expressed gene clusters specific to each of the eight developmental stages. Furthermore, we conducted Gene Ontology (GO) and functional enrichment analysis for genes within these co-expressed clusters using Enrichr [17]. GO terms were considered significant if they met the criteria of bonferroni adjusted p-value < 0.01.

Expression of ortholgs in rat genome

The rat orthologous genes of the human genes in different co-expressed clusters were obtained from the Rat Genome Database (https://rgd.mcw.edu/rgdweb/ortholog/start.html). The transcriptomic data of the prefrontal cortex from rats, which had undergone maternal separation or served as control subjects, have previously been released [18]. The expression level of each rat orthologous gene was collected and compared between different clustered groups.

Annotation of clustered genes using GWAS and DNM databases

The GWAS association data used in our study was obtained from the GWAS Catalog [19, 20]. Specifically, we narrowed down our focus to a specific set of genes associated with 12 psychiatric and neurological disorders, as listed in Supplementary Table 6. In addition to this, we also obtained a list of genes affected by *de novo* mutations that have the potential to result in psychiatric and neurological disorders from the PsyMuKB database [21]. Furthermore, we curated genes that were affected by loss-of-function (LoF) mutations or had a CADD_phred score of ≥30 from a total of 21 psychiatric and neurological disorders, which are detailed in Supplementary Table 7 [22]. We then proceeded to map the genes within the different co-expression clusters to these two gene datasets, which are relevant to psychiatric and neurological disorders.

TF prediction for hub genes

The TF enrichment analysis was conducted using ChEA3, a tool designed to perform transcription factor enrichment analysis by integrating data from various sources, including TF-target gene associations from ENCODE, ReMap, ARCHS4, GTEx, Enrichr, and curated results from the literature [23]. We opted for the 'Mean Rank' setting in ChEA3, to identify the most significant TFs for genes in the C3 cluster. The 'Mean Rank' method calculates the average rank across the integrated libraries (ARCHS4, GTEx, and Enrichr) within ChEA3. The construction of the TF co-expression network involved a Weighted Gene Co-expression Network Analysis (WGCNA) applied to human TF expression data. The WGCNs based on the Genotype-Tissue Expression (GTEx) project was illustrated. We used TF co-expression networks to further analyze the top 15 enriched transcription factors to identify the hubs through ChEA3.

Cell-type specific expression analysis

Cell-type Specific Expression Analysis (CSEA) [24] was used in our study to identify potential cell populations that may be disrupted in a group of patients with various genetic abnormalities. This analysis utilizes a dataset of mouse transcriptomic profiles to identify sets of transcripts that are specifically expressed in particular mouse cells. We subjected the 154 genes that showed signals in GWAS/DNM and the top 15 associated regulatory TF genes to an investigation of their expression in specific cell types (as shown in Fig. 4d).

Chromatin state analysis

Based on various epigenomic marks, the genomic sequences of the human genome have previously been classified into 15 core chromatin states [25]. In our study, we employed bedtools [26] to determine the base-level overlap between different chromatin states and all human genes. This allowed us to annotate each gene with the proportion covered by the 15 distinct chromatin states. Subsequently, we compared the proportion of the 154 genes associated with psychiatric and neurological disorders in GWAS and DNM with the proportion of all genes in chromatin states C0 to C5.

Colocalization with human accelerated regions

Human Accelerated Regions (HARs) demonstrate intriguing patterns of conservation within human populations while showing significant divergence from other mammals, highlighting their crucial role in human-specific evolution and diversification. In our study, we obtained the HAR regions previously identified by Doan et al. [27] and downloaded them for further analysis. To explore the relationship between HAR regions and gene expression, we utilized bedtools [26] to screen the base-level overlapping rate of genes belonging to different co-expressed clusters.

Evaluation of the spatiotemporal expression pattern at single-cell level

The single-nucleus RNA sequencing (snRNA-seq) data from Zhu et al. (2023) was employed in this study, offering a detailed single-cell RNA sequencing dataset of 45,549 cortical nuclei sampled across six developmental stages, from fetal to adult life. This dataset, accessible as a Seurat object in an R package, was downloaded and analyzed using the Seurat package in R. To evaluate the activity of the C3 gene set at the single-cell level during various developmental stages, we utilized the R package AUCell v1.26 [28]. Our methodology began with the application of the 'AUCell_buildRankings' function, employing its default parameters to rank gene expressions for each cell based on an expression matrix. Subsequently, for each gene set and cell, we determined the area-under-the-curve (AUC) values through the 'AUCell_calcAUC' function. These AUC values denote the percentage of genes within the pathway gene set that are ranked highly for each cell.

We used DoRothEA v3.19 [29] to assess the comparative activity levels of transcription factors (TFs). DoRothEA determines TF activity levels by analyzing the expression levels of their target genes, not the own expression levels of TFs. For the statistical evaluation of these TF activity levels, we applied VIPER [30], a method that accounts for the specific mode of each TF-target interaction and has been proven effective for single-cell data analysis. We conducted a comparative analysis of TF activity across different cell types and identified the activity patterns of for MYT1L.

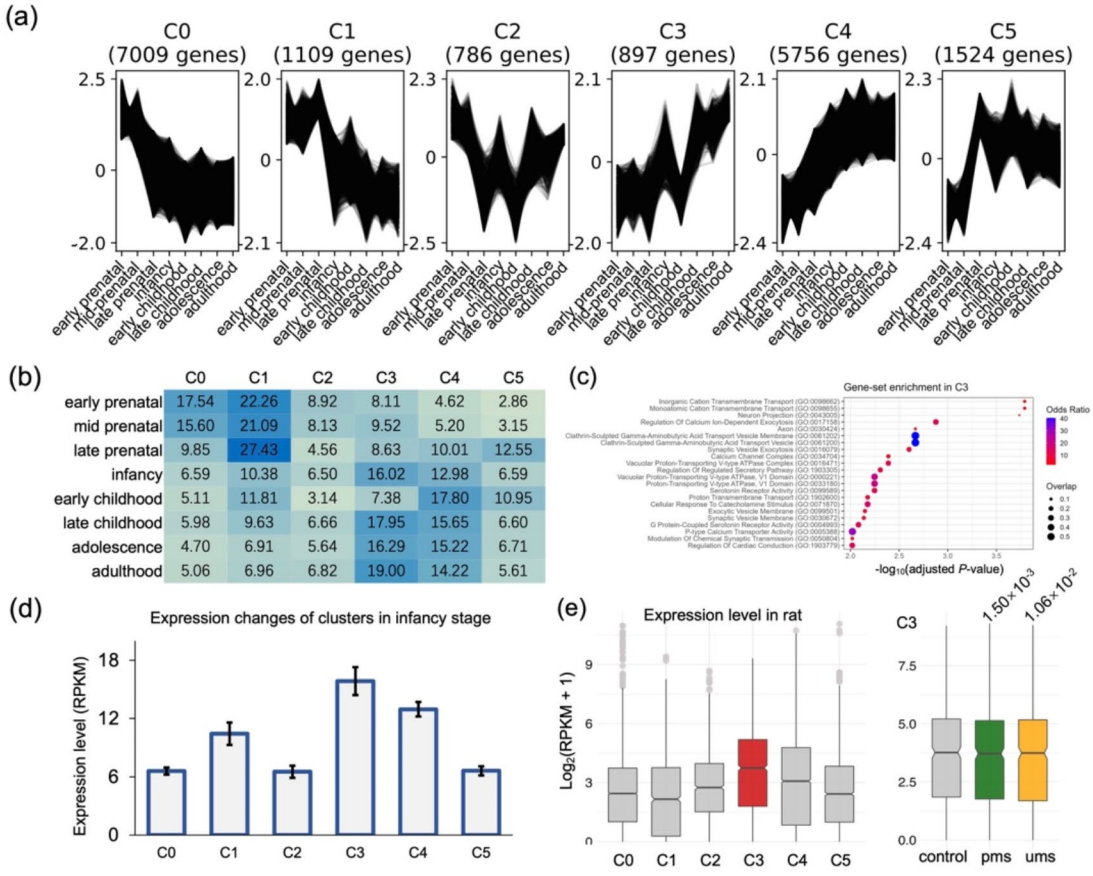

Fig. 2 Expression pattern clustering and functional enrichment of genes highly expressed during the infancy period. (**a**) Clustering of gene expression pattern during eight human developmental stages. The number of genes in each cluster is displayed at the top of the image. The six clusters contain 7009, 1109, 786, 897, 5756, and 1524 genes, respectively; (**b**) Heatmaps for average expression values for genes of six clusters across different stages; (**c**) GO enrichment of 897 genes in C3; (**d**) Expression level of different cluster genes at infancy stage; (**e**) Left: the expression level of homologous genes of different clusters in rat brain; Right: comparisons of expression levels of C3 cluster genes in rats brain subjected to predictable (pms) and unexpected (ums) maternal separation with those control rats

Results

The neuronal signature of infant-specific highly expressed genes

To identify the gene set expressed explicitly during infancy, we utilized the transcriptome dataset in the BrainSpan project [1]. We define infancy as the period from birth to 18 months, a critical developmental window for brain growth and neural circuit establishment, where disruptions can significantly increase the risk of psychiatric and neurological disorders later in life [21]. We partitioned the data into eight developmental periods according to the classification system of human brain development and adulthood created by Kang et al. [1] (Supplementary Table 1). These expression data of brain tissues, defined for different developmental stages, including early prenatal, mid-prenatal, late prenatal, infancy early childhood, late childhood, adolescence, and adulthood, will be used for subsequent analysis. To ensure robust statistical analysis and representation, we have ensured that each developmental stage includes expression data from at least three donor samples (Supplementary Table 2). Co-expressed clusters of genes were identified by Clust [16], which generated six gene expression clusters (C0 – C5). The smallest cluster (C2) includes 786 genes, while the largest cluster (C0) includes 7009 genes.

The expression pattern varied among the six clusters. Genes in C0 and C1 showed a generally declining expression pattern throughout eight developmental periods from early prenatal to adulthood, while genes in C4 showed a continuously increased expression pattern (Fig. 2a). The other three clusters (C2, C3, C5) showed fluctuated pattern during development. Notably, we observed that the expression of 897 genes in the

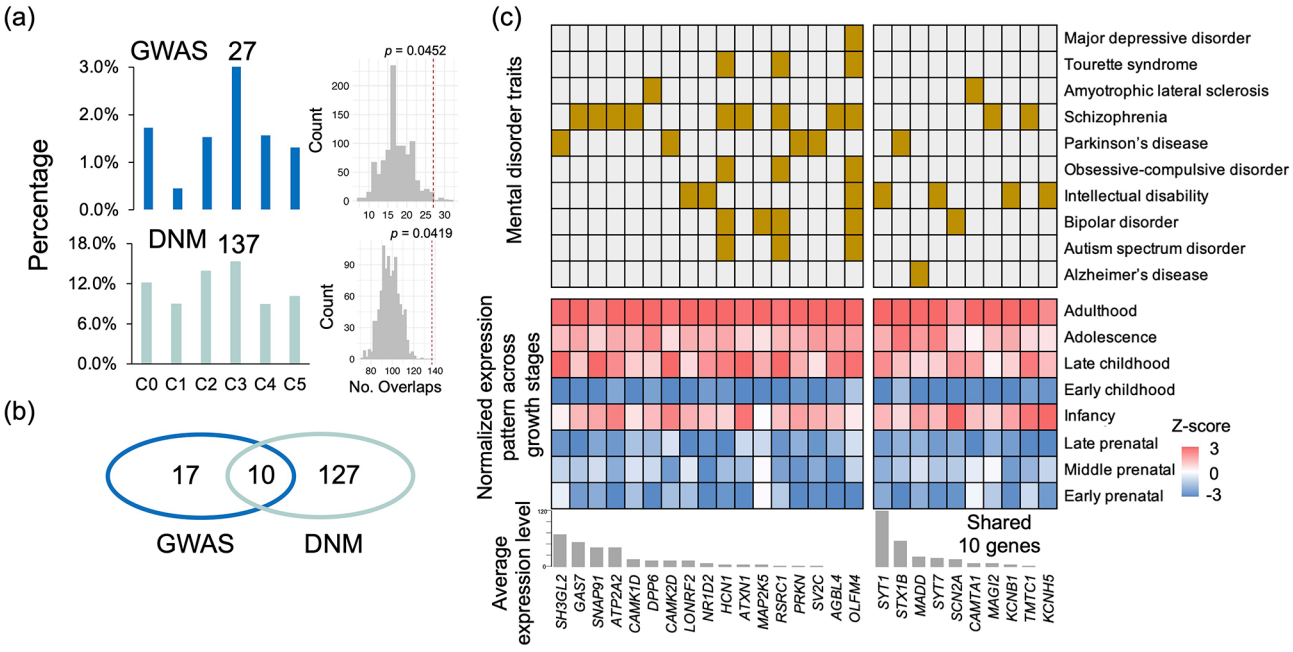

Fig. 3 Exploration of C3 genes residing in GWAS signals associated with psychiatric and neurological disorders. (**a**) Percentage of genes in six different clusters with GWAS signals associated with psychiatric and neurological disorders or with loss of function *de novo* mutation (DNM) burden based on trio studies. The right figures illustrate the comparison between the observed number of overlaps (marked with red dashed lines) and the distribution of overlap numbers for 1000 simulated genes with comparable gene sizes; (**b**) Number of overlapped genes of C3 cluster with both GWAS and DNM; (**c**) Associated psychiatric and neurological disorder traits for the 27 genes with GWAS signals. The color represents the expression levels of the genes, with red indicating relatively high expression and blue indicating relatively low expression, based on a standardized scale for each gene

C3 showed a peak in infancy compared to the three prenatal and early childhood stages (Fig. 2a-b). To identify the general functions of genes in each cluster, we applied Enrichr [17] to explore the functional enrichment. Results from Enrichr were used for all clusters (Supplementary Table 3). Interestingly, for C3, we found the significant GO enrichment terms are like 'neuron projection' (GO:0043005), 'axon' (GO:0030424) and 'synaptic vesicle exocytosis (GO:0016079) (Fig. 2c).

When focusing on the infancy stage, C3 genes were most highly expressed by comparing the average expression values of the six clusters of genes (Fig. 2d), implying the important roles of C3 genes in neuronal development at this stage. To examine whether genes in this cluster could respond to psychological trauma stress during infancy, we utilized the rat model developed in our previous research [18], and compared whether the rat orthologs of these genes exhibit significant expression differences in cases of maternal-infant separation anxiety during infancy. Transcriptomic expression data from rats subjected to predictable (pms) and unpredictable (ums) maternal separation during infancy were compared [18]. We identified the best homologous rat genes related to those of humans and observed that the C3 cluster genes that could be mapped in rat transcriptome data also had the highest expression level compared to other cluster genes (Fig. 2e), indicating their generally functionally conserved roles across different species. Moreover, we compared the expression levels of C3 cluster genes in rats subjected to predictable (pms) and unexpected (ums) maternal separation with those control rats. Significantly reduced expression in both maternal separation treated rats was observed (P_{pms}: 1.50×10^{-3}; P_{ums}: 1.06×10^{-2}) (Fig. 2e), supporting that the C3 genes may be involved in critical roles in neuronal development, and may respond to psychological trauma stress during infancy.

C3 genes differentially modulates susceptibility to various psychiatric and neurological disorders

Since genes in C3 cluster were shown to have a functional enrichment in the neuronal system (Fig. 2c), we further explored whether these genes resided in genomic regions previously identified as GWAS signals associated with psychiatric and neurological disorders. In addition, we also scanned the genes in the PsyMuKB database, which collects potential neuropsychiatric genes with *de novo* mutations (DNM) [31]. Interestingly, a total of 27 in C3 cluster overlapped with the psychiatric and neurological disorder-related GWAS signals, which presented the largest proportion among all clusters (Fig. 3a). Consistently, we also observed the highest percentage of genes with *de novo* Loss of function (LoF) mutations in C3 (Fig. 3a). To account for the potential bias in the overlap percentage due to gene length, we randomly selected genes with

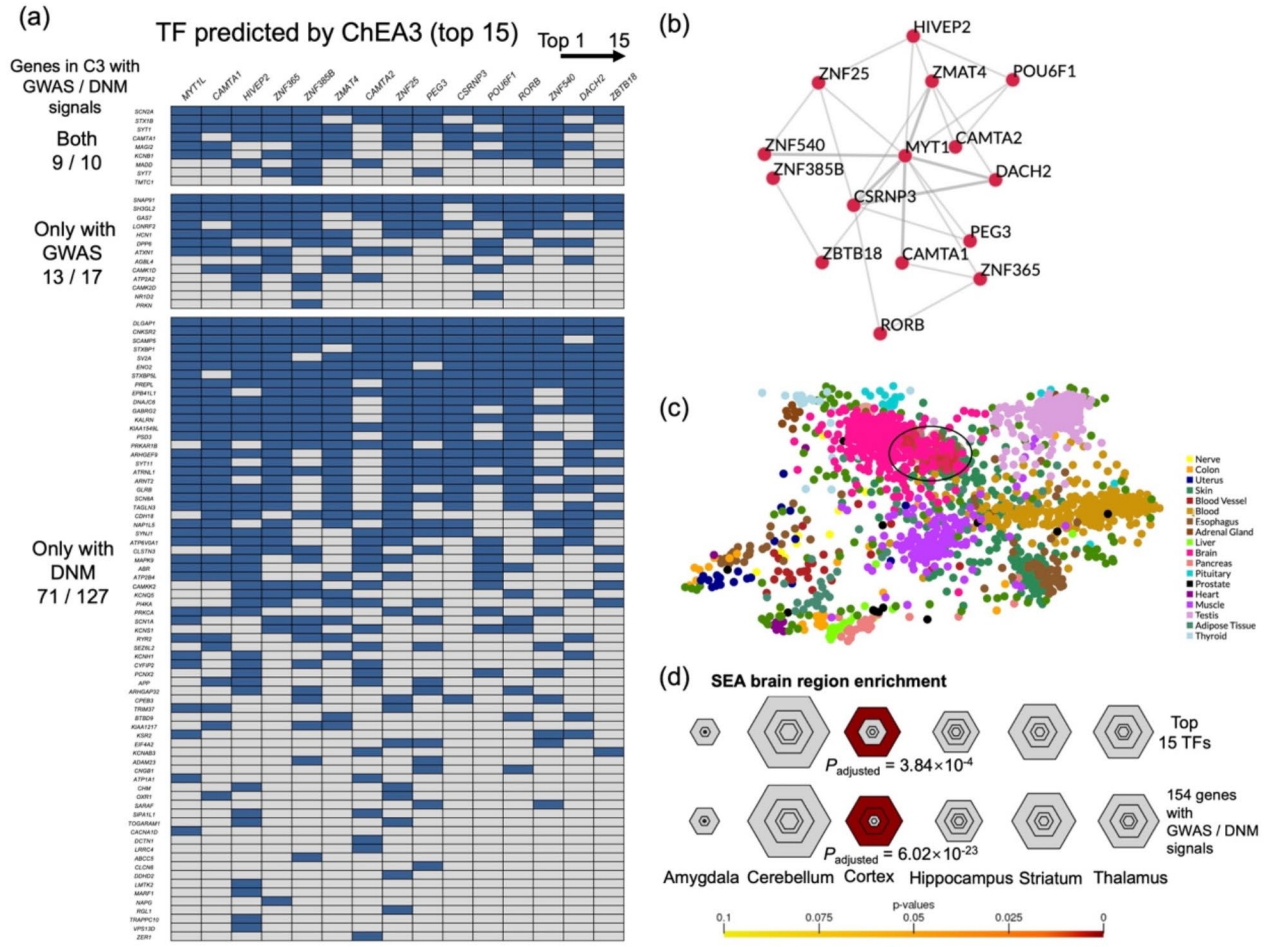

Fig. 4 Transcriptional profiling of TFs regulating neural genes in the C3 cluster. (**a**) TFs predicted by ChEA3 potentially regulating genes in C3 with GWAS or DNM signals; (**b**) Local network of the top 15 TFs. The edges within this network are defined by evidence from the ChEA3 libraries and are directed where ChIP-seq evidence supports the interaction. The relationships between each TF are based on data from the GTEx, TCGA, and ARCHS4 datasets in the ChEA3 database; (**c**) The global TF network presented by ChEA3 is constructed using Weighted Gene Co-expression Network Analysis (WGCNA) on human TF expression data from GTEx, TCGA, and ARCHS4. Global clustering of all TFs with different tissues coloring differently. The region of top 15 TFs is highlighted with black circle, indicating that the top 15 regulating TFs are tissue-specifically expressed in the brain; (**d**) SEA brain region enrichment for top the top 15 TFs and 154 genes in C3 with GWAS or DNM signals. Varying stringencies for enrichment are represented by the size of the hexagons, ranging from the least specific lists (outer hexagons) to the most specific (center). Hexagons are scaled to the size of the gene lists

similar length distributions 1000 times and consistently observed that the overlap ratio of C3 genes is significantly higher than that of the simulated sets, with *p*-values of 4.52×10^{-2} for the GWAS dataset and 4.19×10^{-2} for the DNM dataset. Notably, 10 of the 27 C3 genes overlapped with psychiatric and neurological disorder GWAS signals also showed to be carrying LoF *de novo* mutations (Fig. 3b). For the 27 genes, 24 were potentially associated with one single disease trait, while the rest three genes (*HCN1*, *RSRC1*, *OLFM4*) showing pleiotropic effects (Fig. 3c). All the 27 genes showed an elevated expression level during infancy, and three later developmental periods (late childhood, adolescence, adulthood). Among the ten genes that overlapped with both GWAS and DNM datasets, *SYT1*, may play a vital role in infancy development. *SYT1* is a synaptic binding protein that plays a crucial role in neurotransmitter release. Mutations in the *SYT1* gene are associated with a newly identified neurodevelopmental disorder called *SYT1*-associated neurodevelopmental disorder. *De novo* missense *SYT1* variants in the C2B domain were reported to be associated with severe intellectual disability [32].

Profiling of transcription factors regulating neural genes in C3 cluster

Transcription factors (TFs) are proteins that regulate gene expression by binding to DNA sequences and play a crucial role in the modification of gene expression [33]. To investigate potential TFs regulating C3 neural genes and examine whether there are TFs acting as hubs in these networks, the TF enrichment analysis was performed. ChEA3 (ChIP-X Enrichment Analysis

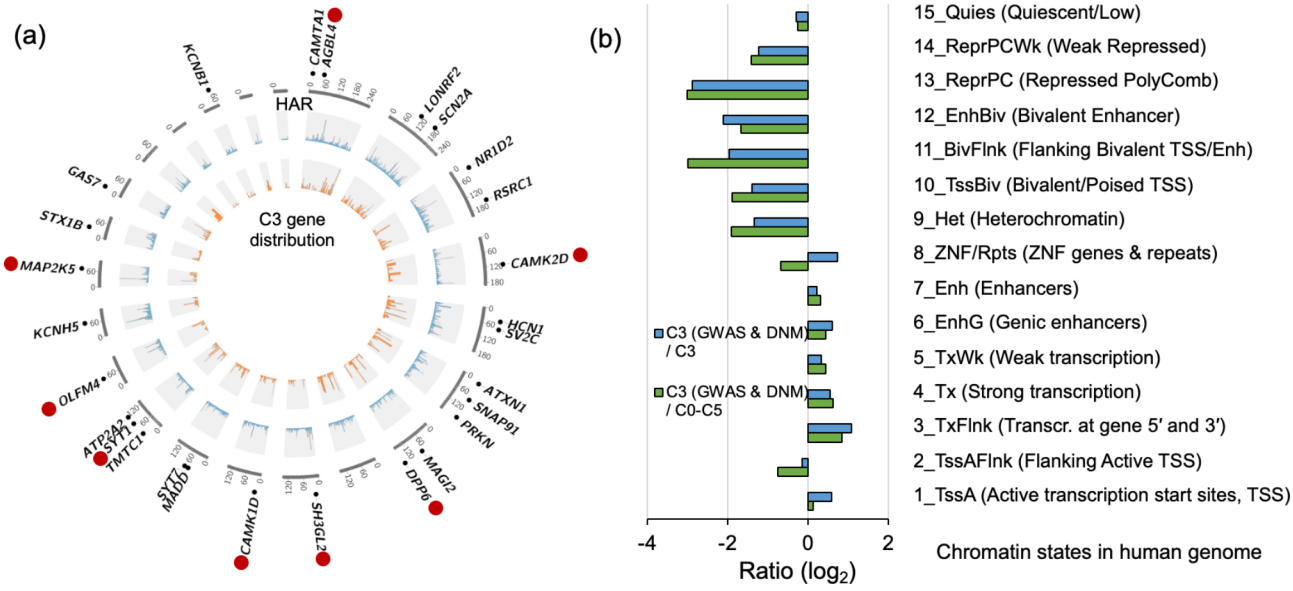

Fig. 5 Genomic characteristics of neuronal genes in C3 cluster. (**a**) Circos plot of the distribution of human accelerated regions (HAR) and C3 genes. The 27 genes with GWAS signals were marked outside, with genes colocalized with HARs highlighted in red dot; (**b**) Comparision of chromatin state of 154 genes with psychiatric and neurological disorder GWAS and DNM signals against all C3 genes and all C0-C5 genes

Version 3) was developed to conduct transcription factor enrichment analysis, which integrates data about TF/target–gene associations from multiple assay types and other sources of evidence [23]. A total of 154 genes in C3 with mental GWAS or DNM signals (Fig. 4b) were subjected to this analysis by ChEA3. We found that 93 of the 154 genes could potentially be regulated by the top 15 ranking TFs predicted by ChEA3 (Fig. 3a; Supplementary Table 4). We randomly selected 154 genes from the genome 1000 times and analyzed each set using the ChEA3, and found none of the permutations showed 93 or more genes targeted by the 15 transcription factors (p-value < 0.001). Among them, genes with both GWAS and DNM signals presented with the highest rate (90.0%) of being regulated by the top 15 TFs, followed by genes with only GWAS signals (76.9%), and those with only DNMs (55.9%) (Fig. 4a). The observation may suggest that genes with both GWAS and DNM signals are more likely to be functional crucial genes and play a central role in mental development.

Notably, we observed that the top 2 ranking TFs are MYT1L, CAMTA1, which are predicted to regulate 49, 47 of the 154 genes, respectively. TF Co-regulatory Networks are constructed using the ChEA3 database. Edges within these TF Co-regulatory Networks are defined by evidence from the ChEA3 libraries and are directed where ChIP-seq evidence supports the interaction. We utilized the TF Co-regulatory Networks from ChEA3 to evaluate the interconnectivity of the top 15 transcription factors (TFs). MYT1L is also presented as hub nodes in the local co-expression network constructed by the 15 TFs (Fig. 4b). *MYT1L* has been documented to be a crucial transcriptional activator and/or repressor, and functions in the developing mammalian central nervous system [34, 35]. Mutations in this gene lead to serious neurodevelopmental disorders, such as cognitive disability and with ASD [34]. For *CAMTA1*, *de novo* mutations in this gene cause a syndrome variably associated with spasticity, ataxia, and intellectual disability [36]. When examing the global TF network presented by ChEA3, we found the top 15 regulating TFs are tissue-specifically expressed in the brain (Fig. 4c). To further determine the specificity of the brain region, we used the CSEA tool [24], and found that both the 154 genes with GWAS/DNM signals ($P_{adjusted} = 6.02 \times 10^{-23}$) or the top 15 associated regulating TFs ($P_{adjusted} = 3.84 \times 10^{-4}$) are predicted to be dominantly functioning in the cortex (Fig. 4d).

Genomic and epigenomic characterization of psychiatric and neurological disorder related genes in C3

In order to explore whether the C3 genes possess human-specific evolutionary characteristics, we examined the extent to which C3 cluster genes are colocalized in human accelerated regions (HARs). HARs show intriguing conservation within human populations but display high divergence from other mammals. It has been documented several specific HARs likely to have essential functions in human brain that are potentially important targets of recent human brain evolution [27]. The results found that 8 of 27 genes overlapping with GWAS or DNM databases are coloalized with HARs, which includes *CAMTA1*, *SYT1*, *AGBL4*, *CAMK2D*, *DPP6*, *MAP2K5*, *OLFM4* and *SH3GL2* (Fig. 5a), indicating they

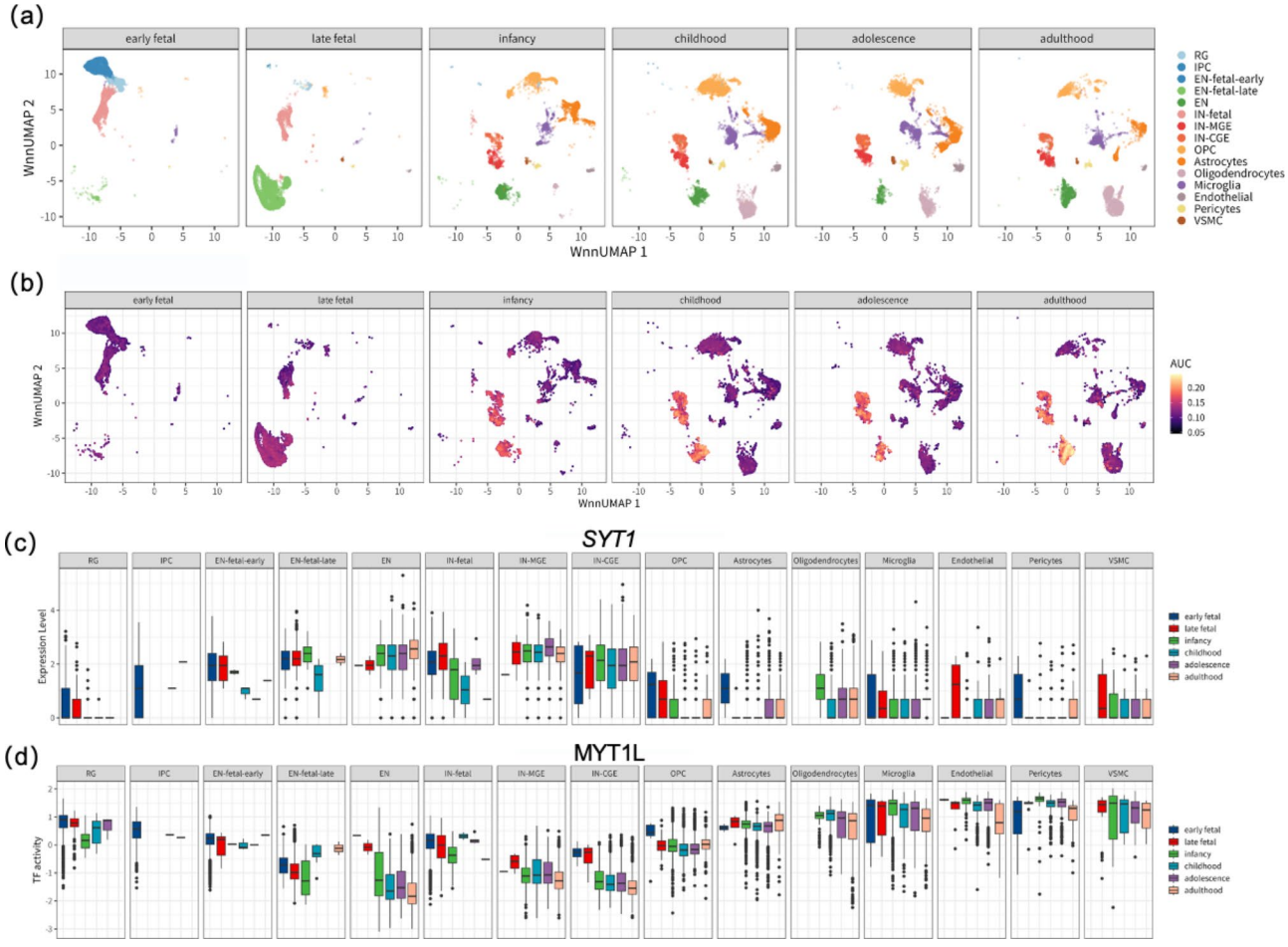

Fig. 6 Spatiotemporal Expression Patterns of C3 Genes. (**a**) UMAP visualizations depict single cells characterized by RNA-seq across six distinct developmental stages. The cell type annotations are based on the work of Zhu et al. [15]; (**b**) The relative expression levels of C3 genes are compared between different cell types at various stages; (**c**) The expression levels of *SYT1* are examined across various cell types and stages; (**d**) The activity of the MYT1L transcription factor is analyzed across different stages

may play a vital role in early neuronal development and under strong selection pressure.

Different human genomic regions are classified with 15 diverse core chromatin states based on multiple epigenomic marks, which can capture information describing different types of genomic elements such as promoters, enhancers, transcribed, repressed, and repetitive regions [25]. We first determined the chromatin states for all clustered genes at the gene-based level, and investigated chromatin states for the 154 genes with GWAS and DNM associated with psychiatric and neurological disorders. Using genes in all clusters and C3 as background, we found the chromatin states in the 154 genes were of higher presentation in states like '3_TxFlnk' (Transcription at gene 5' and 3'), '4_Tx' (Strong transcription), '5_TxWk' (weak transcription), '6_EnhG' (Genic enhancers), '7_Enh' (Enhancers) (Fig. 5b; Supplmentary Table 5). These states are dominant with epigenetic marks H3K4me1 and H3K27ac, which are enhancer-specific modifications.

Spatiotemporal expression dynamics of C3 gene set in brain development

In order to further refine the study of the spatiotemporal expression characteristics of the C3 gene set in brain tissues, we utilized the single-cell transcriptome data of brain tissues at various developmental stages from Zhu et al. [15]. The study includes six periods: Early fetal, Late fetal, Infancy, Childhood, Adolescence, and Adulthood. Our research found that the C3 gene set was inactive in most cell types during the fetal period before infancy. Interestingly, during the infancy period and subsequent developmental periods, the C3 gene was predominantly highly active in three types of brain cells, namely excitatory neurons (EN), medial ganglionic eminence (IN-MGE), and ganglionic eminence (IN-CGE) (Fig. 6a-b).

Focusing on the *SYT1* gene within the C3 gene set, it was found to be highly expressed in these three cell types, with a noticeable trend of increased expression during the infancy period in cell types EN and IN-CGE. Additionally, using the DoRothEA [29], we predicted that

MYT1L is a negative upstream regulatory transcription factor for *SYT1*. The spatiotemporal expression data of this batch of brain tissues also revealed that the activity of this transcription factor MYT1L is inversely related to the expression levels of *SYT1* across different cell types and developmental stages (Fig. 6c-d). This phenomenon suggests that the increase in expression levels of the *SYT1* gene in EN and CGE as the developmental period progresses may be due to the decreased activity of the transcription factor MYT1L, leading to the relief of inhibition.

Discussion

The development of the brain in infancy is crucial for the proper functioning of an individual throughout their life [37]. During this stage, the brain is rapidly growing and forming connections [38], which lay the foundation for future learning, behavior, and overall well-being. The research findings of this study reveal crucial insights into infancy-specific genes enriched in the neuronal system and their potential implications for psychiatric and neurological disorders. The identified unique gene cluster, exhibiting a peak in expression during infancy, demonstrates significant functional enrichment in the neuronal system and synaptic signaling. Furthermore, the study introduces a fascinating association between these specific C3 genes and psychiatric and neurological disorders, which is drawn from overlaps observed with GWAS signals and instances of *de novo* loss of function mutations. These overlaps suggest that these genes might play a substantial role in both the development and susceptibility to a broad spectrum of psychiatric and neurological disorders. For researchers in neurobiology and psychology, this presents an intriguing molecular terrain to explore towards a better understanding of psychiatric and neurological disorders.

The consistent and elevated expression levels of these genes in both human and rat datasets suggest a functionally conserved role across species, thus emphasizing their critical involvement in neuronal development during infancy. An intriguing aspect of these genes is their remarkable response to environmental factors, as depicted through experimental results involving maternal separation in rats [18]. In such scenarios, mammals display a significant reduction in the expression levels of these genes, emphasizing the conservation of these genes across species and their necessity in neural development. This insight deepens our understanding of how early-life experiences cast sharp effects on the neural development process.

Early detection of mental illness in childhood is crucial for timely intervention and better long-term outcomes. Identifying mental health issues at an early stage can lead to more effective treatment plans and improved quality of life. However, diagnosing mental illness in infants and young children poses significant challenges due to the subtlety of symptoms and the variability in developmental milestones [39]. These findings could drive the development of novel early intervention strategies for psychiatric and neurological disorders by targeting the specific genes within the C3 cluster that play crucial roles in infant neuronal development [40]. By gaining a deeper understanding of the genetic and epigenetic characteristics of these genes, researchers and clinicians could potentially identify biomarkers for early detection and diagnosis of psychiatric disorders [41, 42], leading to more effective and personalized treatment approaches [43]. Moreover, the identification of TFs that regulate the C3 neural genes, such as *MYT1L* and *CAMTA1*, provides potential targets for therapeutic interventions aimed at modulating gene expression in the context of psychiatric and neurological disorders. By understanding the regulatory networks and brain region specificity of these TFs, researchers could develop targeted approaches for modulating gene expression in the developing brain, with potential applications in mitigating the risk of psychiatric and neurological disorders.

Conclusions

In summary, this research illuminates the pivotal role of one co-expressed gene cluster in infancy-specific neuronal development and its potential implications for psychiatric and neurological disorders. The findings underscore the importance of these genes in shaping the trajectory of neuronal development during infancy, and their potential involvement in modulating susceptibility to a spectrum of psychiatric and neurological disorders. Further research focusing on the precise molecular mechanisms by which these genes contribute to psychiatric and neurological disorders, as well as their potential as therapeutic targets, holds significant promise for advancing our understanding and potential interventions in the field of mental health.

Abbreviations

DSM-5	Diagnostic and Statistical Manual of Mental Disorders, 5th Edition
GWAS	Genome-wide association studies
TF	Transcription factor
GO	Gene ontology
LoF	Loss-of-function
WGCNA	Weighted Gene Co-expression Network Analysis
GTEx	Genotype-Tissue Expression project
CSEA	Cell-type Specific Expression Analysis
HARs	Human Accelerated Regions
snRNA-seq	Single-nucleus RNA sequencing

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-024-10911-0.

Supplementary Material 1

Acknowledgements
The authors thank Dr. Jie Qiu for his valuable recommendations on this study. Additionally, the authors appreciate the reviewers for their constructive suggestions and revisions.

Author contributions
Z.W. conceived and managed the project. W.W. and Z.L. processed the data, W.W. designed and performed experiments for variant characterization and different components of analysis. D.P. and G-N.L. collected and analyzed the data. W.W. wrote the manuscript, All authors critically revised and approved the manuscript.

Funding
This work was supported by grants from National Natural Science Foundation of China (grant nos. 32200924, 82071518, 81971261), Shanghai Science and Technology Committee (grant no. 22YF1439000), Shanghai Municipal Education Commission (grant no. 2021-01-07-00-02-E0086), STI2030-Major Projects (grant no. 2021ZD0200600).

Data availability
All the data used in this study were downloaded from publicly accessible resources: expression data (http://www.brainspan.org/static/download.html), GWAS data (https://www.ebi.ac.uk/gwas/), de novo variants (https://www.psymukb.net/Download).

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 3 April 2024 / Accepted: 16 October 2024
Published online: 29 October 2024

References

1. Kang HJ, Kawasawa YI, Cheng F, Zhu Y, Xu X, Li M, Sousa AMM, Pletikos M, Meyer KA, Sedmak G, Guennel T, Shin Y, Johnson MB, Krsnik Ž, Mayer S, Fertuzinhos S, Umlauf S, Lisgo SN, Vortmeyer A, Weinberger DR, Mane S, Hyde TM, Huttner A, Reimers M, Kleinman JE, Šestan N. Spatio-temporal transcriptome of the human brain. Nature. 2011;478:483–9. https://doi.org/10.1038/nature10523.
2. Knickmeyer RC, Gouttard S, Kang C, Evans D, Wilber K, Smith JK, Hamer RM, Lin W, Gerig G, Gilmore JH. A structural MRI study of human Brain Development from Birth to 2 years. J Neurosci. 2008;28:12176–82. https://doi.org/10.1523/JNEUROSCI.3479-08.2008.
3. Gilmore JH, Knickmeyer RC, Gao W. Imaging structural and functional brain development in early childhood. Nat Rev Neurosci. 2018;19:123–37. https://doi.org/10.1038/nrn.2018.1.
4. Feldman R, Eidelman AI. Biological and environmental initial conditions shape the trajectories of cognitive and social-emotional development across the first years of life. Dev Sci. 2009;12:194–200. https://doi.org/10.1111/j.1467-7687.2008.00761.x.
5. Natu VS, Rosenke M, Wu H, Querdasi FR, Kular H, Lopez-Alvarez N, Grotheer M, Berman S, Mezer AA, Grill-Spector K. Infants' cortex undergoes microstructural growth coupled with myelination during development. Commun Biol. 2021;4:1191. https://doi.org/10.1038/s42003-021-02706-w.
6. Huang H, Shu N, Mishra V, Jeon T, Chalak L, Wang ZJ, Rollins N, Gong G, Cheng H, Peng Y, Dong Q, He Y. Development of human brain structural networks through infancy and childhood. Cereb Cortex. 2015;25:1389–404. https://doi.org/10.1093/cercor/bht335.
7. American Psychological Association. (n.d.), Child and adolescent mental and behavioral health., APA. (2022).
8. Krystal JH, State MW. Psychiatric disorders: diagnosis to Therapy. Cell. 2014;157:201–14. https://doi.org/10.1016/j.cell.2014.02.042.
9. Shen MD, Piven J. Brain and behavior development in autism from birth through infancy. Dialogues Clin Neurosci. 2017;19:325–33. https://doi.org/10.31887/dcns.2017.19.4/mshen.
10. Sacrey LAR, Zwaigenbaum L, Bryson S, Brian J, Smith IM, Roberts W, Szatmari P, Vaillancourt T, Roncadin C, Garon N. Screening for behavioral signs of Autism Spectrum Disorder in 9-Month-Old Infant siblings. J Autism Dev Disord. 2021;51:839–48. https://doi.org/10.1007/s10803-020-04371-0.
11. Madrid-Gambin F, Föcking M, Sabherwal S, Heurich M, English JA, O'Gorman A, Suvitaival T, Ahonen L, Cannon M, Lewis G, Mattila I, Scaife C, Madden S, Hyötyläinen T, Orešič M, Zammit S, Cagney G, Cotter DR, Brennan L. Integrated Lipidomics and Proteomics Point to early blood-based changes in Childhood Preceding later development of psychotic experiences: evidence from the Avon Longitudinal Study of parents and children. Biol Psychiatry. 2019;86:25–34. https://doi.org/10.1016/j.biopsych.2019.01.018.
12. činskas, Z.A. Kučinskiene, A. Kusumawardhani, H. Kuzelova-Ptackova, S. Landi, L.C. Lazzeroni, P.H.Lee, S.E. Legge, D.S. Lehrer, R. Lencer, B. Lerer, M. Li, J. Lieberman, G.A. Light,S. Limborska, C.-M. Liu, J. Lönnqvist, C.M. Loughland, J. Lubinski, J.J. Luykx, A.Lynham, M. Macek, A. Mackinnon, P.K.E. Magnusson, B.S. Maher, W. Maier, D. Malaspina,J. Mallet, S.R. Marder, S. Marsal, A.R. Martin, L. Martorell, M. Mattheisen, R.W.McCarley, C. McDonald, J.J. McGrath, H. Medeiros, S. Meier, B. Melegh, I. Melle, R.I.Mesholam-Gately, A. Metspalu, P.T. Michie, L. Milani, V. Milanova, M. Mitjans, E.Molden, E. Molina, M.D. Molto, V. Mondelli, C. Moreno, C.P. Morley, G. Muntané, K.C.Murphy, I. Myin-Germeys, I. Nenadić, G. Nestadt, L. Nikitina-Zake, C. Noto, K.H. Nuechterlein, N.L. O'Brien, F.A. O'Neill,S.-Y. Oh, A. Olincy, V.K. Ota, C. Pantelis, G.N. Papadimitriou, M. Parellada, T. Paunio,R. Pellegrino, S. Periyasamy, D.O. Perkins, B. Pfuhlmann, O. Pietiläinen, J. Pimm,D. Porteous, J. Powell, D. Quattrone, D. Quested, A.D. Radant, A. Rampino, M.H. Rapaport,A. Rautanen, A. Reichenberg, C. Roe, J.L. Roffman, J. Roth, M. Rothermundt, B.P.F.Rutten, S. Saker-Delye, V. Salomaa, J. Sanjuan, M.L. Santoro, A. Savitz, U. Schall,R.J. Scott, L.J. Seidman, S.I. Sharp, J. Shi, L.J. Siever, E. Sigurdsson, K. Sim,N. Skarabis, P. Slominsky, H.-C. So, J.L. Sobell, E. Söderman, H.J. Stain, N.E. Steen,A.A. Steixner-Kumar, E. Stögmann, W.S. Stone, R.E. Straub, F. Streit, E. Strengman,T.S. Stroup, M. Subramaniam, C.A. Sugar, J. Suvisaari, D.M. Svrakic, N.R. Swerdlow,J.P. Szatkiewicz, T.M.T. Ta, A. Takahashi, C. Terao, F. Thibaut, D. Toncheva, P.A.Tooney, S. Torretta, S. Tosato, G.B. Tura, B.I. Turetsky, A. Üçok, A. Vaaler, T. van Amelsvoort, R. van Winkel, J. Veijola, J. Waddington, H. Walter, A. Waterreus, B.T.Webb, M. Weiser, N.M. Williams, S.H. Witt, B.K. Wormley, J.Q. Wu, Z. Xu, R. Yolken,C.C. Zai, W. Zhou, F. Zhu, F. Zimprich, E.C. Atbaşoğlu, M. Ayub, C. Benner, A. Bertolino, D.W. Black, N.J. Bray, G. Breen, N.G. Buccola,W.F. Byerley, W.J. Chen, C.R. Cloninger, B. Crespo-Facorro, G. Donohoe, R. Freedman,C. Galletly, M.J. Gandal, M. Gennarelli, D.M. Hougaard, H.-G. Hwu, A. V. Jablensky,S.A. McCarroll, J.L. Moran, O. Mors, P.B. Mortensen, B. Müller-Myhsok, A.L. Neil,M. Nordentoft, M.T. Pato, T.L. Petryshen, M. Pirinen, A.E. Pulver, T.G. Schulze, J.M.Silverman, J.W. Smoller, E.A. Stahl, D.W. Tsuang, E. Vilella, S.-H. Wang, S. Xu, N.Dai, Q. Wenwen, D.B. Wildenauer, F. Agiananda, N. Amir, R. Antoni, T. Arsianti, A.Asmarahadi, H. Diatri, P. Djatmiko, I. Irmansyah, S. Khalimah, I. Kusumadewi, P. Kusumaningrum,P.R. Lukman, M.W. Nasrun, N.S. Safyuni, P. Prasetyawan, G. Semen, K. Siste, H. Tobing,N. Widiasih, T. Wiguna, D. Wulandari, N. Evalina, A.J. Hananto, J.H. Ismoyo, T.M.Marini, S. Henuhili, M. Reza, S. Yusnadewi, A. Abyzov, S. Akbarian, A. Ashley-Koch,H. van Bakel, M. Breen, M. Brown, J. Bryois, B. Carlyle, A. Charney, G. Coetzee, G.Crawford, S. Dracheva, P. Emani, P. Farnham, M. Fromer, T. Galeev, M. Gandal, M. Gerstein,G. Giase, K. Girdhar, F. Goes, K. Grennan, M. Gu, B. Guerra, G. Gursoy, G. Hoffman,T. Hyde, A. Jaffe, S. Jiang, Y. Jiang, A. Kefi, Y. Kim, R. Kitchen, J.A. Knowles,F. Lay, D. Lee, M. Li, C. Liu, S. Liu, E. Mattei, F. Navarro, X. Pan, M.A. Peters,D. Pinto, S. Pochareddy, D. Polioudakis, M. Purcaro, S. Purcell, H. Pratt, T. Reddy,S. Rhie, P. Roussos, J. Rozowsky, S. Sanders, N. Sestan, A. Sethi, X. Shi, A. Shieh,V. Swarup, A. Szekely, D. Wang, J. Warrell, S. Weissman, Z. Weng, K. White, J. Wiseman,H. Witt, H. Won, S. Wood, F. Wu, X. Xu, L. Yao, P. Zandi, M.J. Arranz, S. Bakker,S. Bender, E. Bramon, D.A. Collier, B. Crepo-Facorro, J. Hall, C. Iyegbe, R. Kahn,S. Lawrie, C. Lewis, K. Lin, D.H. Linszen, I. Mata, A. McIntosh, R.M. Murray, R.A.Ophoff, J. van Os, J. Powell, D. Rujescu, M. Walshe, M. Weisbrod, T. Achsel, M. Andres-Alonso,C. Bagni, À. Bayés, T. Biederer, N. Brose, T.C. Brown, J.J.E. Chua, M.P. Coba, L.N.Cornelisse, A.P.H. de Jong, J. de Juan-Sanz, D.C. Dieterich, G. Feng, H.L. Goldschmidt,E.D. Gundelfinger, C. Hoogenraad, R.L. Huganir, S.E. Hyman, C. Imig, R. Jahn, H. Jung,P.S. Kaeser, E. Kim, F. Koopmans, M.R. Kreutz, N. Lipstein, H.D. MacGillavry, R. Malenka,P.S. McPherson, V. O'Connor, R. Pielot, T.A. Ryan, D. Sahasrabudhe, C. Sala, M. Sheng,K.-H. Smalla, A.B. Smit, T.C. Südhof, P.D. Thomas, R.F. Toonen, J.R.T. van Weering,M. Verhage, C. Verpelli, R. Adolfsson, C. Arango, B.T. Baune, S.I. Belangero, A.D.Børglum, D. Braff, E. Bramon, J.D. Buxbaum, D. Campion, J.A. Cervilla, S. Cichon,D.A. Collier, A. Corvin, D. Curtis, M. Di 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. Jönsson, 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. Nöthen, 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. St Clair, H. Stefánsson, 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, L.de Haan, T. van Amelsvoort, R. van Winkel, A. Gareeva, P.C. Sham, Y. Shi, D. St Clair,J. van Os, Mapping genomic loci implicates genes and synaptic biology in schizophrenia,Nature 604 (2022) 502–508. https://doi.org/10.1038/s41586-022-04434-5.

13. Palmer DS, Howrigan DP, Chapman SB, Adolfsson R, Bass N, Blackwood D, Boks MPM, Chen C-Y, Churchhouse C, Corvin AP, Zandi PP, Neale BM. Exome sequencing in bipolar disorder identifies AKAP11 as a risk gene shared with schizophrenia. Nat Genet. 2022;54:541–7. https://doi.org/10.1038/s41588-022-01034-x.
14. Naumova OY, Lee M, Rychkov SY, Vlasova Nv, Grigorenko EL. Gene expression in the human brain: the current state of the study of specificity and Spatiotemporal dynamics. Child Dev. 2013;84:76–88. https://doi.org/10.1111/cdev.12014.
15. Zhu K, Bendl J, Rahman S, Vicari JM, Coleman C, Clarence T, Latouche O, Tsankova NM, Li A, Brennand KJ, Lee D, Yuan G-C, Fullard JF, Roussos P. Multi-omic profiling of the developing human cerebral cortex at the single-cell level. Sci Adv. 2023;9. https://doi.org/10.1126/sciadv.adg3754.
16. Abu-Jamous B, Kelly S. Clust: automatic extraction of optimal co-expressed gene clusters from gene expression data. Genome Biol. 2018;19:1–11. https://doi.org/10.1186/s13059-018-1536-8.
17. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A, McDermott MG, Monteiro CD, Gundersen GW. Ma'ayan, Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. 2016;44:W90–7. https://doi.org/10.1093/nar/gkw377.
18. Shi D-D, Zhang Y-D, Ren Y-Y, Peng S-Y, Yuan T-F, Wang Z. Predictable maternal separation confers adult stress resilience via the medial prefrontal cortex oxytocin signaling pathway in rats. Mol Psychiatry. 2021;26:7296–307. https://doi.org/10.1038/s41380-021-01293-w.
19. Sollis E, Mosaku A, Abid A, Buniello A, Cerezo M, Gil L, Groza T, Güneş O, Hall P, Hayhurst J, Ibrahim A, Ji Y, John S, Lewis E, MacArthur JAL, McMahon A, Osumi-Sutherland D, Panoutsopoulou K, Pendlington Z, Ramachandran S, Stefancsik R, Stewart J, Whetzel P, Wilson R, Hindorff L, Cunningham F, Lambert SA, Inouye M, Parkinson H, Harris LW. The NHGRI-EBI GWAS catalog: knowledgebase and deposition resource. Nucleic Acids Res. 2023;51:D977–85. https://doi.org/10.1093/nar/gkac1010.
20. Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, McMahon A, Morales J, Mountjoy E, Sollis E, Suveges D, Vrousgou O, Whetzel PL, Amode R, Guillen JA, Riat HS, Trevanion SJ, Hall P, Junkins H, Flicek P, Burdett T, Hindorff LA, Cunningham F, Parkinson H. The NHGRI-EBI GWAS catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Res. 2019;47:D1005–12. https://doi.org/10.1093/nar/gky1120.
21. Sanai N, Nguyen T, Ihrie RA, Mirzadeh Z, Tsai H-H, Wong M, Gupta N, Berger MS, Huang E, Garcia-Verdugo J-M, Rowitch DH. Alvarez-Buylla, corridors of migrating neurons in the human brain and their decline during infancy. Nature. 2011;478:382–6. https://doi.org/10.1038/nature10487.
22. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic Acids Res. 2019;47:D886–94. https://doi.org/10.1093/nar/gky1016.
23. Keenan AB, Torre D, Lachmann A, Leong AK, Wojciechowicz ML, Utti V, Jagodnik KM, Kropiwnicki E, Wang Z, Ma'ayan A. ChEA3: transcription factor enrichment analysis by orthogonal omics integration. Nucleic Acids Res. 2019;47:W212–24. https://doi.org/10.1093/nar/gkz446.
24. Dougherty JD, Schmidt EF, Nakajima M, Heintz N. Analytical approaches to RNA profiling data for the identification of genes enriched in specific cells. Nucleic Acids Res. 2010;38:4218–30. https://doi.org/10.1093/nar/gkq130.
25. Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, Kheradpour P, Zhang Z, Wang J, Ziller MJ, Amin V, Whitaker JW, Schultz MD, Ward LD, Sarkar A, Quon G, Sandstrom RS, Eaton ML, Wu Y-C, Pfenning AR, Wang X, Claussnitzer M, Liu Y, Coarfa C, Harris RA, Shoresh N, Epstein CB, Gjoneska E, Leung D, Xie W, Hawkins RD, Lister R, Hong C, Gascard P, Mungall AJ, Moore R, Chuah E, Tam A, Canfield TK, Hansen RS, Kaul R, Sabo PJ, Bansal MS, Carles A, Dixon JR, Farh K-H, Feizi S, Karlic R, Kim A-R, Kulkarni A, Li D, Lowdon R, Elliott G, Mercer TR, Neph SJ, Onuchic V, Polak P, Rajagopal N, Ray P, Sallari RC, Siebenthall KT, Sinnott-Armstrong NA, Stevens M, Thurman RE, Wu J, Zhang B, Zhou X, Beaudet AE, Boyer LA, De Jager PL, Farnham PJ, Fisher SJ, Haussler D, Jones SJM, Li W, Marra MA, McManus MT, Sunyaev S, Thomson JA, Tlsty TD, Tsai L-H, Wang W, Waterland RA, Zhang MQ, Chadwick LH, Bernstein BE, Costello JF, Ecker JR, Hirst M, Meissner A, Milosavljevic A, Ren B, J.A., Stamatoyannopoulos. T. Wang, M. Kellis, Integrative analysis of 111 reference human epigenomes, Nature 518 (2015) 317–330. https://doi.org/10.1038/nature14248
26. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26:841–2. https://doi.org/10.1093/bioinformatics/btq033.
27. Doan RN, Bae B-I, Cubelos B, Chang C, Hossain AA, Al-Saad S, Mukaddes NM, Oner O, Al-Saffar M, Balkhy S, Gascon GG, Nieto M, Walsh CA. Mutations in human accelerated regions disrupt cognition and Social Behavior. Cell. 2016;167:341–e35412. https://doi.org/10.1016/j.cell.2016.08.071.
28. Aibar S, González-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, Rambow F, Marine J-C, Geurts P, Aerts J, van den Oord J, Atak ZK, Wouters J, Aerts S. SCENIC: single-cell regulatory network inference and clustering. Nat Methods. 2017;14:1083–6. https://doi.org/10.1038/nmeth.4463.
29. Garcia-Alonso L, Holland CH, Ibrahim MM, Turei D, Saez-Rodriguez J. Benchmark and integration of resources for the estimation of human transcription factor activities. Genome Res. 2019;29:1363–75. https://doi.org/10.1101/gr.240663.118.
30. Chen M, Zhou X. VIPER: variability-preserving imputation for accurate gene expression recovery in single-cell RNA sequencing studies. Genome Biol. 2018;19:196. https://doi.org/10.1186/s13059-018-1575-1.
31. Lin GN, Guo S, Tan X, Wang W, Qian W, Song W, Wang J, Yu S, Wang Z, Cui D, Wang H. PsyMuKB: an integrative De Novo variant knowledge base for Developmental disorders. Genomics Proteom Bioinf. 2019;17:453–64. https://doi.org/10.1016/j.gpb.2019.10.002.
32. Melland H, Bumbak F, Kolesnik-Taylor A, Ng-Cordell E, John A, Constantinou P, Joss S, Larsen M, Fagerberg C, Laulund LW, Thies J, Emslie F, Willemsen M, Kleefstra T, Pfundt R, Barrick R, Chang R, Loong L, Alfadhel M, van der Smagt J, Nizon M, Kurian MA, Scott DJ, Ziarek JJ, Gordon SL, Baker K. Expanding the genotype and phenotype spectrum of SYT1-associated neurodevelopmental disorder. Genet Sci. 2022;24:880–93. https://doi.org/10.1016/j.gim.2021.12.002.
33. Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, Chen X, Taipale J, Hughes TR, Weirauch MT. Hum Transcription Factors Cell. 2018;172:650–65. https://doi.org/10.1016/j.cell.2018.01.029.
34. Chen J, Yen A, Florian CP, Dougherty JD. MYT1L in the making: emerging insights on functions of a neurodevelopmental disorder gene. Transl Psychiatry. 2022;12:292. https://doi.org/10.1038/s41398-022-02058-x.
35. Weigel B, Tegethoff JF, Grieder SD, Lim B, Nagarajan B, Liu Y-C, Truberg J, Papageorgiou D, Adrian-Segarra JM, Schmidt LK, Kaspar J, Poisel E, Heinzelmann E, Saraswat M, Christ M, Arnold C, Ibarra IL, Campos J, Krijgsveld J, Monyer H, Zaugg JB, Acuna C, Mall M. MYT1L haploinsufficiency in human neurons and mice causes autism-associated phenotypes that can be reversed by genetic and pharmacologic intervention. Mol Psychiatry. 2023;28:2122–35. https://doi.org/10.1038/s41380-023-01959-7.
36. Wijnen IGM, Veenstra-Knol HE, Vansenne F, Gerkes EH, de Koning T, Vos YJ, Tijssen MAJ, Sival D, Darin N, Vanhoutte EK, Oosterloo M, Pennings M, van de Warrenburg BP, Kamsteeg E-J. De novo variants in CAMTA1 cause a syndrome variably associated with spasticity, ataxia, and intellectual disability. Eur J Hum Genet. 2020;28:763–9. https://doi.org/10.1038/s41431-020-0600-5.
37. Dehaene-Lambertz G, Spelke ES. The infancy of the human brain. Neuron. 2015;88:93–109. https://doi.org/10.1016/j.neuron.2015.09.026.
38. Silbereis JC, Pochareddy S, Zhu Y, Li M, Sestan N. The Cellular and Molecular landscapes of the developing Human Central Nervous System. Neuron. 2016;89:248–68. https://doi.org/10.1016/j.neuron.2015.12.008.
39. Luby J, Allen N, Estabrook R, Pine DS, Rogers C, Krogh-Jespersen S, Norton ES, Wakschlag L. Mapping infant neurodevelopmental precursors of mental disorders: how synthetic cohorts & computational approaches can be used to enhance prediction of early childhood psychopathology. Behav Res Ther. 2019;123:103484. https://doi.org/10.1016/j.brat.2019.103484.

40. Izett E, Rooney R, Prescott SL, De Palma M, McDevitt M. Prevention of Mental Health Difficulties for Children Aged 0–3 years: a review. Front Psychol. 2020;11:500361. https://doi.org/10.3389/fpsyg.2020.500361.
41. Fiori LM, Turecki G. Investigating epigenetic consequences of early-life adversity: some methodological considerations. Eur J Psychotraumatol. 2016;7:31593. https://doi.org/10.3402/ejpt.v7.31593.
42. Turecki G, Ota VK, Belangero SI, Jackowski A, Kaufman J. Early life adversity, genomic plasticity, and psychopathology. Lancet Psychiatry. 2014;1:461–6. https://doi.org/10.1016/S2215-0366(14)00022-4.
43. Gandal MJ, Leppa V, Won H, Parikshak NN, Geschwind DH. The road to precision psychiatry: translating genetics into disease mechanisms. Nat Neurosci. 2016;19:1397–407. https://doi.org/10.1038/nn.4409.

Publisher's note