dbBIP: a comprehensive bipolar disorder database for genetic research

Xiaoyan Li[1,†], Shunshuai Ma[1,†], Wenhui Yan[1], Yong Wu[2], Hui Kong[1], Mingshan Zhang[1], Xiongjian Luo[3,4] and Junfeng Xia ![ORCID][1,*]

[1]Key Laboratory of Intelligent Computing and Signal Processing of Ministry of Education and Information Materials and Intelligent Sensing Laboratory of Anhui Province, Institutes of Physical Science and Information Technology, Anhui University, 111 Jiulong Road, Shushan District, Hefei, Anhui 230601, China
[2]Affiliated Wuhan Mental Health Center, Tongji Medical College, Huazhong University of Science and Technology, 93 Youyi Road, Qiaokou District, Wuhan, Hubei 430030, China
[3]Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences & Yunnan Province, Kunming Institute of Zoology, Chinese Academy of Sciences, 32 Jiaochang East Road, Wuhua District, Kunming, Yunnan 650223, China
[4]Kunming College of Life Science, University of Chinese Academy of Sciences, 19 Qingsong Road, Panlong District, Kunming, Yunnan 650204, China

*Corresponding author: Fax: +86 551 63861872; Tel: +86-13966677630; Email: jfxia@ahu.edu.cn
†Xiaoyan Li and Shunshuai Ma should be considered joint first author.

Abstract

Bipolar disorder (BIP) is one of the most common hereditary psychiatric disorders worldwide. Elucidating the genetic basis of BIP will play a pivotal role in mechanistic delineation. Genome-wide association studies (GWAS) have successfully reported multiple susceptibility loci conferring BIP risk, thus providing insight into the effects of its underlying pathobiology. However, difficulties remain in the extrication of important and biologically relevant data from genetic discoveries related to psychiatric disorders such as BIP. There is an urgent need for an integrated and comprehensive online database with unified access to genetic and multi-omics data for in-depth data mining. Here, we developed the dbBIP, a database for BIP genetic research based on published data. The dbBIP consists of several modules, i.e.: (i) single nucleotide polymorphism (SNP) module, containing large-scale GWAS genetic summary statistics and functional annotation information relevant to risk variants; (ii) gene module, containing BIP-related candidate risk genes from various sources and (iii) analysis module, providing a simple and user-friendly interface to analyze one's own data. We also conducted extensive analyses, including functional SNP annotation, integration (including summary-data-based Mendelian randomization and transcriptome-wide association studies), co-expression, gene expression, tissue expression, protein–protein interaction and brain expression quantitative trait loci analyses, thus shedding light on the genetic causes of BIP. Finally, we developed a graphical browser with powerful search tools to facilitate data navigation and access. The dbBIP provides a comprehensive resource for BIP genetic research as well as an integrated analysis platform for researchers and can be accessed online at http://dbbip.xialab.info.

Database URL: http://dbbip.xialab.info

Introduction

Bipolar disorder (BIP) is a common and serious psychiatric condition that typically manifests with strong shifts in mood, vacillating between episodic bouts of mania and depression (1, 2). With a lifetime prevalence rate of 1–4% (3), BIP usually occurs in adolescence or early adulthood, but diagnosis is often established after the onset of symptoms with a common delay. In addition to mood fluctuations between mania/hypomania and depression, individuals with BIP frequently exhibit cognitive deficits and other psychiatric comorbidities (4). BIP patients also show higher rates of suicide and risks of related medical complications (such as osteoporosis, diabetes mellitus, metabolic syndromes and cardiovascular and endocrine disorders) compared to the general population (5, 6). Accordingly, BIP is associated with a huge economic burden.

For example, in the USA alone, the national economic burden of BIP exceeded 195 billion dollars in 2020, with 25% directly attributed to healthcare costs (7). Thus, due to its high morbidity, mortality and social and economic costs, BIP has become a major health problem globally.

At present, the etiology of BIP remains largely unknown. Evidence from multiple studies implicates both environmental and genetic factors in the initiation of BIP (8, 9). Twin studies suggest that the narrow-sense heritability of BIP is as high as 70% (10), indicating a critical role of genetic factors in this disorder. As such, genetic studies could help unravel the mechanisms underlying BIP and assist in the discovery of novel therapeutic targets. To reveal the genetic architecture underlying BIP, linkage (11, 12), association (13, 14) and meta-analysis studies (15) have identified many novel

BIP-associated susceptibility variants, genes and chromosomal regions (16), although the detection of credible susceptibility variants and genes has been limited due to the small sample size and low coverage of genetic markers. However, the wealth of data uncovered from genome-wide association studies (GWAS) provides a powerful approach to explore the genetic etiology of BIP (2). In 2007, the Wellcome Trust Case Control Consortium (17) reported the first BIP GWAS. Subsequently, multiple studies on different human populations have identified various genetic loci and genes associated with BIP (18–32). The most recent GWAS carried out by the Psychiatric Genomics Consortium (PGC) (41 917 cases and 371 549 controls) reported 64 genome-wide significant loci, including 33 newly discovered loci (33). In addition to the most common form of genetic variation in the human genome (i.e. SNPs), other structural genomic variants are reported to play prominent roles in BIP, e.g. copy number variations (CNV) and rare variants (34–36).

Despite considerable progress in identifying susceptibility variants for BIP, the extraction of useful and functional information from the massive amount of genetic data still poses a significant challenge (36). First, as with other common psychiatric diseases, most genetic risk variants identified from BIP GWAS are located in non-coding regions of the human genome (26, 31), and how they confer risk remains largely unknown. Second, while GWAS have succeeded in identifying BIP-associated genetic loci, only a handful have been resolved to individual genes (37) based on genetic characteristics. Thus, there are obvious difficulties in recognizing specific susceptibility genes underlying GWAS loci. Third, further integration of different multi-omics data sets [e.g. integrated analysis of genetic and gene expression data (33, 38)] is necessary to clarify specific susceptibility genes and provide experimental validation. Thus, to best utilize current genetic data, there is a pressing need for in-depth data collection and mining to help provide novel insights into the pathogenesis of BIP and develop more effective methods for early diagnosis and treatment.

In this study, we developed a comprehensive genetic database for BIP (dbBIP) to meet the increasing needs of data acquisition and analysis. The dbBIP database aims to provide researchers with comprehensive BIP genetic data based on data collection, data integration and functional analysis. The dbBIP sources come from the extensive and systematic integration and storage of diverse BIP studies and include data on genetic findings, SNP functional annotation, gene expression, brain expression quantitative trait loci (eQTL) and network-based proteins. The dbBIP not only offers a user-friendly interface to browse, search, analyze and visualize data but also provides meaningful information for further functional characterization of high-confidence candidate variants and genes.

Methods and materials
Genetic data
Many genetic studies have been carried out in the last decade to identify the genetic risk variants for BIP. The two largest BIP GWAS, i.e. PGC2 (31) and PGC3 (33), were included in the dbBIP database. The PGC2 data set from the PGC BIP Working Group, which includes 20 352 BIP cases and 31 358 healthy controls of European descent (14 countries in total), identified 30 genome-wide significant loci (including 20 newly

identified loci) associated with BIP (31). The more recent PGC3 GWAS data set from the PGC, which includes 41 917 BIP cases and 371 549 healthy controls, identified 64 genomic loci associated with BIP (33). We downloaded the SNP association results of PGC2 and PGC3 (https://www.med.unc.edu), with details regarding participant recruitment, sample preparation and statistical analysis reported in previous publications (31, 33). In total, summary data on 19 963 820 SNPs were downloaded.

Although CNVs show robust associations with neurodevelopmental disorders (39–41), recent research indicates that they likely pose less risk to BIP than to autism and schizophrenia (41–45), and only a limited number of significant CNVs associated with BIP have been detected by genetic studies (46–48). Green *et al.* (34) analyzed the CNV status of 41 321 subjects (including 2591 BIP cases and 8842 healthy controls) and identified several independent CNV loci associated with BIP, including deletions at 3aq29, duplications at 1aq21.1 and duplications at 16p11.2. As such, we identified genes affected by the above-reported CNVs and deposited detailed CNV information into the dbBIP, including detection platform, number of cases and controls, CNV location and CNV-affected genes.

Both de novo and disruptive mutations are implicated in the pathogenesis of BIP (49, 50). The development of next-generation sequencing, especially whole exome/genome sequencing (WES/WGS), has enabled the detection of rare and de novo mutations at the exome and genome level (51, 52). To collect data on genetic mutations related to BIP, we searched the PubMed database for relevant articles published since 2015 using keywords 'whole exome sequencing and bipolar disorder' or 'whole genome sequencing and bipolar disorder'. All returned results were manually checked and examined, resulting in the collation of eight original WES/WGS studies with BIP results (http://dbbip.xialab.info/Exome_sequencing_publications).

Functional genomic data
Most BIP risk variants reside in non-coding genomic regions and lack functional interpretation. By systematically integrating BIP-related GWAS and functional genomics data, we can identify potential causal variation(s) at a given susceptibility locus. Here, we used three well-optimized functional annotation tools, i.e. Combined Annotation-Dependent Depletion (CADD) (53), Linear insight (LINSIGHT) (54) and RegulomeDB (55), to prioritize potential risk SNPs. CADD utilizes evolutionary conservation information and functional data from the ENCODE database (56) to rank variants that are likely to be pathogenic or deleterious. LINSIGHT predicts negative selection on non-coding regions as well as functional variants using genomic data via the generalized linear and probabilistic molecular evolution models. RegulomeDB uses high-throughput experimental data sets from ENCODE and other sources to identify potential regulatory variants. For CADD (scores of 0–99) and LINSIGHT (scores of 0–1), higher scores are indicative of functional SNPs, while for RegulomeDB (scores of 1–6), lower scores are indicative of functional SNPs. We downloaded all three annotated functional genomic data sets to annotate all variants identified by PGC3 GWAS (33), with all annotated results then integrated into the dbBIP. Detailed information on these annotation approaches can be found in previous publications (53–55).

As regulatory sequences typically reside in open chromatin, we hypothesized that BIP risk variants involved in neurodevelopmental processes may affect chromatin accessibility. Based on human-induced pluripotent stem cell (iPSC)-derived neurons, Zhang *et al.* (57) identified many allele-specific open chromatin (ASoC) variants. Thus, in view of the abundance of BIP GWAS variants in ASoC SNPs in glutamatergic (iN-Glut) samples (57), we collected functional genomic results from the iN-Glut samples and annotated the potentially functional SNPs that may affect chromatin accessibility during neurodevelopment in BIP. Data analysis methodologies are reported in Zhang *et al.* (57).

Regional association plots

We used LocusZoom (58) to visualize genomic regions of interest based on BIP summary statistics in PGC3 (33) (e.g. *P* values of genetic variants) downloaded from the PGC portal (https://www.med.unc.edu/pgc/download-results/), in full compliance with the PGC terms. Additional details are reported in previous study (58) and on the LocusZoom website (https://my.locuszoom.org/).

Integrative analysis data

As stated above, most GWAS-identified BIP genetic risk variants are located in non-coding regions, especially within putative regulatory elements, suggesting possible BIP risk via gene expression regulation (26). Both summary-data-based Mendelian randomization (SMR) (38) and FUSION/transcriptome-wide association studies (TWAS) (59) are powerful tools for the integration and identification of risk genes that may increase disease susceptibility under altered expression and thus were applied in the current study. For PGC2 (31), we carried out SMR and FUSION/TWAS analysis by integrating GWAS summary statistics and brain eQTL data from the CommonMind Consortium [CMC (60)], BrainSeq Consortium (second phase) [LIBD2-DLPFC (61)] and PsychENCODE Consortium (62) data sets. Details on the SMR and FUSION/TWAS integrative analysis procedures can be found in previous studies (63, 64). All integrative analysis results (including SMR and FUSION/TWAS) can be freely downloaded from the dbBIP database.

For PGC3 (33), the results reported in Mullins *et al.* (33) were included in our developed database. To prioritize candidate risk genes of BIP, Mullins *et al.* (33) integrated genetic associations from the PGC3 (33) GWAS with eQTL data from the PsychENCODE Consortium (62) (1321 brain samples) and eQTLGen Consortium (65) (31 684 whole blood samples) using SMR. They also integrated PGC3 GWAS summary statistics with brain eQTL summary data from the PsychENCODE Consortium (62) using the TWAS/FUSION method.

Differentially expressed genes

Increasing evidence supports the importance of gene expression dysregulation in BIP pathogenesis (66). To identify differentially expressed genes (DEGs) in BIP subjects and healthy controls, the PsychENCODE Consortium (62) established a large-scale gene expression data set using post-mortem dorsolateral prefrontal cortex (DLPFC) tissues of 144 BIP cases and 899 healthy controls (European ancestries) using RNA-sequencing (RNA-seq). This data set is regarded as one of the most representative of gene expression in BIP. Thus, we downloaded the gene mRNA expression values (http://resource.psychencode.org/) and included the data in the dbBIP. Detailed information on the study subjects is provided in the original publication (62). To visualize the data, we used the jQuery plug-in plotly (https://plot.ly/javascript/) to generate boxplots for the target genes in BIP cases and controls.

Protein–protein interaction data

Proteins are essential for biological processes, and those implicated in the same disease often show strong associations or interactions (67, 68). Protein–protein interaction (PPI) network analysis is an effective method to assess whether proteins encoded by BIP susceptibility genes show high interactions with other proteins. Li *et al.* (69) constructed a scored human InWeb_IM PPI network based on various indicators (e.g. reproducibility of interactive data) and quality control and integration of data, resulting in a more biologically relevant network than comparable resources. Currently, the InWeb_IM contains >500 000 interactions, aggregated from eight databases (i.e. DIP, BioGRID, BIND, IntAct, NetPath, MatrixDB, Reactome and WikiPathways) and covering 87% of human UniProt IDs (69). Thus, we downloaded the latest release of InWeb_IM PPI data, which can be visualized and queried in the dbBIP. For visualization, dbBIP can construct PPI plots based on interactions among input proteins, including force and circular display types.

Spatiotemporal expression pattern data

Exploring spatiotemporal variations in risk gene expression in the brain can help clarify the role of such genes in the pathogenesis of diseases such as BIP (70). To elucidate the potential role of candidate BIP susceptibility genes in the central nervous system, we used two independent expression data sets (71) to analyze spatiotemporal expression patterns. First, normalized RNA-seq data from BrainSpan (71) expression samples (http://www.brainspan.org/) of different aged subjects (8 post-conception weeks to 40 years, $N = 42$) were downloaded and transformed (63, 72). Expression data obtained from four prefrontal cortical regions of the brain (i.e. orbital, medial, dorsolateral and ventrolateral prefrontal cortices) were integrated into the dbBIP. The second expression data set was obtained from BrainCloud (73) (http://braincloud.jhmi.edu/) and contained gene expression values from the human prefrontal cortex of post-mortem brains across different ages as well as human mRNA expression data of 267 normal subjects analyzed using microarrays (73). Details on the BrainCloud data set are provided in the original publication (73). For query functions, the dbBIP supports user queries by gene symbols and generates a spatiotemporal expression plot for each queried gene.

Tissue expression analysis data

Here, gene expression data obtained from the Genotype-Tissue Expression (GTEx) project (v8p release) (74) were used to investigate expression levels of various target genes across separate human tissues (especially the brain). Gene expression levels were quantified using RNA-seq, and relevant gene expression data from 53 human tissues were downloaded from the GTEx (http://gtexportal.org/) (74), which also provides information on collection, extraction, expression and

processing procedures (74). The dbBIP can generate histogram plots to visualize tissue expression analysis of the above data.

QTL data

Various non-coding genetic variants play essential roles in BIP (2, 75). These genetic associations potentially signal variants that impact gene expression by affecting RNA transcription, splicing and stability (74). Given the importance of eQTL data in investigating candidate risk variants and their effects, we explored different eQTL data sets to identify genes in the brain that may show transcriptional level changes under the influence of the BIP susceptibility variants. This resulted in the inclusion of three large-scale brain eQTL studies in our database [i.e. CMC (60), O'Brien *et al.* (76) and PsychENCODE Consortium (62) eQTL data sets]).

The CMC eQTL data set (60) contains 1 817 945 significant cis-eQTL (at a $P < 1 \times 10^{-3}$) from the DLPFC tissue of 467 human subjects of European descent. This data set was downloaded from the website (https://www.synapse.org/CMC) for inclusion in the dbBIP. The eQTL data set reported in O'Brien *et al.* (76) ($N = 120$) contains eQTL results derived from pre-natal post-mortem brains collected at the second trimester. Significant eQTL results ($P < 0.05$) were downloaded and included in the dbBIP. The large brain PsychENCODE Consortium eQTL data set (62) ($N = 1695$), which explores the relationship between genetic variation and gene expression in the human brain (especially DLPFC tissues), was directly downloaded from PsychENCODE (http://resource.psychencode.org). In diseased brains, high-effect isoform changes are considered to be highly reflective of genetic risk (62). Therefore, we also downloaded and deposited transcript QTL (tQTL) data from CMC (60) and PsychENCODE (62)) in the database.

Co-expression data

Recent studies suggest that BIP susceptibility genes are significantly co-expressed in time- and tissue-specific patterns. Thus, co-expression analysis could help to identify high-priority candidate genes for BIP from genetic studies. To identify whether BIP susceptibility genes are co-expressed in the human brain, we performed gene co-expression analyses using the RNA-seq expression data of BIP cases ($N = 144$) from the PsychENCODE Consortium (62). Pearson correlation coefficients were then calculated, as described previously (77), and those genes with correlation coefficients ≥ 0.8 were retained in the dbBIP.

Prioritization of BIP risk genes

To integrate evidence derived from different BIP studies and identify and prioritize potential candidate risk genes for BIP, we used the arbitrary cumulative scoring method developed by Ayalew *et al.* (78). This approach supposes that BIP-associated genes may be discovered in independent studies and are thus scored as promising BIP candidate genes based on (I) GWAS-identified genes (31, 33), (ii) CNV-disrupted genes (34), (iii) WES/WGS-identified genes (49, 50, 79–89), (iv) SMR integrative analysis [genetic associations from large-scale PGC3 (33) and PsychENCODE brain eQTL data (62)], (v) TWAS integrative analysis (genetic summary statistics from PGC3 (33) and PsychENCODE brain eQTL data (62)), (VI) DEGs (62) and (VII) brain expression results from The Human Protein Atlas (90) (genes are 'expressed' when FPKM expression >5). Consequently, each analysis contributes one point to the identified gene via polyevidence scoring (78), and the final gene score is calculated based on a cumulative scoring strategy, with high scores suggesting multiple analysis evidence supporting the gene as a BIP susceptibility gene.

Results

Database summary

Based on comprehensive genetic and multi-omics data collection and re-analysis, we systematically integrated related data and results in the dbBIP (http://dbbip.xialab.info) (Table 1). Thus, the dbBIP not only contains genetic susceptibility variants (i.e. SNPs) and potential risk genes for BIP but also provides in-depth analysis results, including SNP functional annotation, integrative analysis and DEG analysis. Specifically, the dbBIP integrates powerful online analysis tools and allows advanced users to easily customize and extend analysis, e.g. static LocusZoom, QTL (including eQTL and tQTL), PPI, co-expression, spatiotemporal expression and tissue expression analyses (Figure 1). All data generated or analyzed in this study are freely available to view and download at the dbBIP website (http://dbbip.xialab.info/Download).

We established a common MySQL relational database to store all dbBIP information, which runs on an Ubuntu 14.10 LTS operating system. A user-friendly web platform for browsing and searching was implemented using PhpMyAdmin and JavaScript, powered by Bootstrap (responsive and mobile-first front end is a web interface based on a free and open-source CSS framework).

The dbBIP database provides users with a powerful search engine and a user-friendly web interface to access, browse and download different data types and connections. Users simply need to enter query items, with the input format and content easily located on each dbBIP query page. In addition to 'Quick Search' via keyword, dbBIP presents an 'Advanced Search' function for genes to allow users to combine queries for a detailed overview of gene results. Most returned results in the dbBIP are output as tables. Therefore, the DataTables plug-in (https://datatables.net/) was added to the database to allow advanced users to search and manipulate (show/hide/reorder) table columns. The database provides a detailed explanation of the returned results by each query, including the original data source and definition of each table column.

Key dbBIP modules

Currently, the dbBIP contains three main modules, i.e. SNP, Gene and Analysis modules (Figure 1 and Table 1). The SNP module contains three separate tabs: 'PGC2 GWAS', 'PGC3 GWAS' and 'Functional SNPs'. The PGC2 and PGC3 tabs allow powerful functional searches and GWAS SNP queries and provide various statistics, including SNP position, odds ratio, *P* value and annotation information (i.e. CADD, LINSIGHT and RegulomeDB). The 'Functional SNPs' tab currently contains 1985 GWAS risk SNPs that affect chromatin accessibility during neurodevelopment in BIP (based on integration of Zhang *et al.* 57) data (i.e. ASoC is associated with functional disease variants) and PGC3 GWAS variants.

The Gene module consists of seven tabs and six different levels of data: (I) Genes prioritized from multiple sources
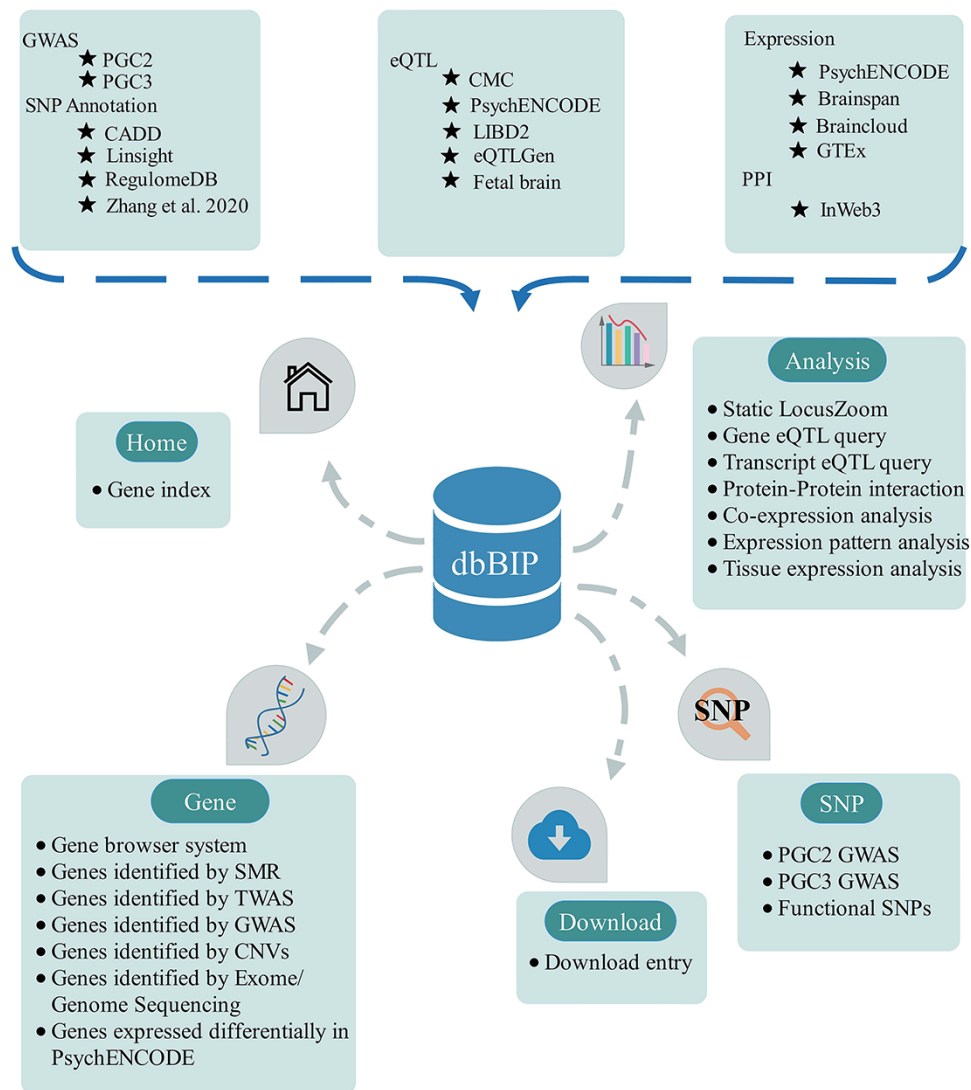
Figure 1. Overview of database content and construction. The dbBIP contains genetic data and analytical tools with browse, search, download and visualize functions.

Table 1. Data description of SNP, gene and analysis modules

Module	Entry	Data set	Tissue	Reference
SNP	PGC2 GWAS	PGC2	Blood	(31)
	PGC3 GWAS	PGC3	Blood	(33)
	Functional SNPs	Zhang *et al.* (2020)	iPSC-derived neurons	(57)
Gene	Genes identified by SMR	CMC, LIBD2-DLPFC, PsychENCODE and eQTLGen	Brain and Blood	(60–62, 65)
	Genes identified by TWAS	CMC, LIBD2-DLPFC and PsychENCODE	Brain	(60–62)
	Genes identified by GWAS	PGC2 and PGC3	Blood	(31, 33)
	Genes identified by CNVs	Green *et al.* (2016)	Blood	(34)
	Genes identified by exome sequencing	Literature	Blood	(49–50, 79–84)
	Genes expressed differentially in PsychENCODE	Gandal *et al.* (2018)	Brain	(62)
Analysis	Static LocusZoom	PGC3	Blood	(33)
	Gene eQTL query	CMC, fetal brain and PsychENCODE	Brain	(60, 76, 62)
	Transcript eQTL query	CMC and PsychENCODE	Brain	(60, 62)
	PPI	Li *et al.* (2016)	Human tissue	(69)
	Co-expression analysis	Gandal *et al.* (2018)	Brain	(62)
	Expression pattern analysis	Brainspan and BrainCloud	Brain	(71, 73)
	Tissue expression analysis	GTEx	Human tissue	(74)

of data; (II) Genes detected from integrative analysis of two GWAS [i.e. PGC2 (31) and PGC3 (33)] and four eQTL data sets [i.e. CMC (60), LIBD2-DLPFC (61), PsychENCODE (62) and eQTLGen (65)] using the SMR approach; (III) Genes detected from integrative analysis of two GWAS [i.e. PGC2 (31) and PGC3 (33)] and four eQTL data sets [i.e. CMC (60), LIBD2-DLPFC (61), PsychENCODE (62) and eQTL-Gen (65)] using the TWAS/FUSION approach; (IV) Genes detected in two large-scale BIP GWAS data sets; (V) Genes influenced by CNVs (based on large-scale CNV research). Fifteen CNVs were included and annotated, with information on CNV location, genes affected by CNV and CNV detection platform provided; (VI) Genes identified by WES/WGS based on 13 studies (http://dbbip.xialab.info/Exome_sequencing_publications) and (VII) DEGs based on RNA-seq data between DLPFC BIP patients ($N = 144$) and control subjects ($N = 899$).

In the Analysis module, we compiled LocusZoom, eQTL, tQTL, PPI, co-expression, spatiotemporal expression and tissue expression data. This module provides a user-friendly and powerful interface to query and analyze one's own data in the dbBIP. LocusZoom allows users to search and draw regional associations of interest. Users can also query the eQTL and tQTL results included in the dbBIP. The spatiotemporal expression pattern tab allows users to investigate if target genes are preferentially expressed in specific brain regions and/or at specific developmental stages. We embedded the BrainSpan and BrainCloud data sets in the dbBIP. The PPI tab provides a one-click test to discern potential PPIs among queried proteins. Based on the above calculated gene score, we utilized ECharts.js (https://echarts.baidu.com) to color code each queried gene in the PPI network. The co-expression tab prioritizes BIP candidate genes from large-scale transcriptome study and allows users to explore if BIP susceptibility
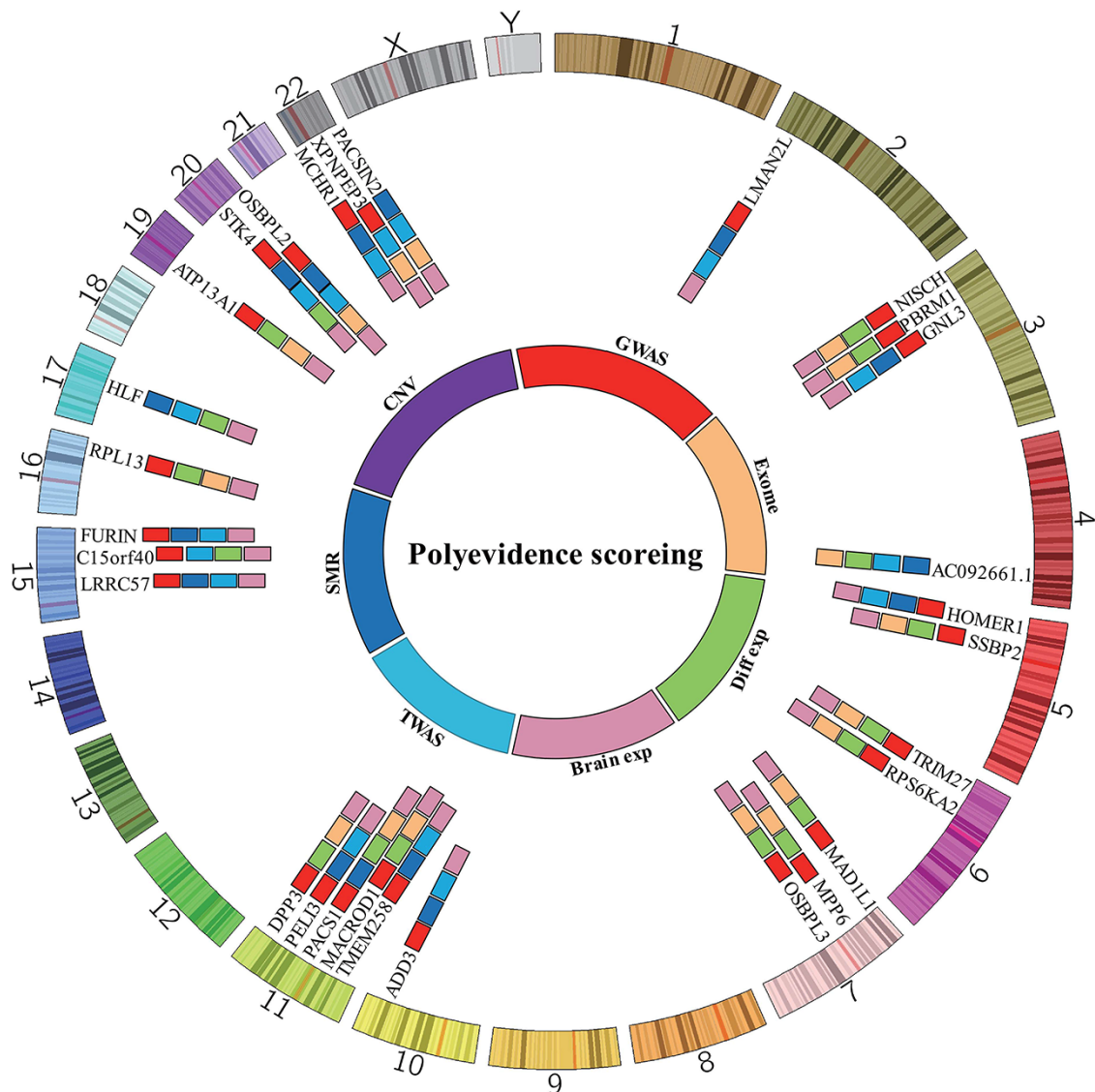

Figure 2. Top candidate causal genes identified in this study. By integrating prediction results from different methods, 29 high-confidence causal genes were identified. OSBPL2, STK4 and PACS1 had the highest scores and thus represent the most promising causal genes for BIP.

Table 2. Significant pathways of genes with a polyevidence score of 3 and above

Category	Pathway[a]	*P* value	*P* adj
GOTERM_CC_DIRECT	Membrane	1.28E–06	3.63E–04
GOTERM_MF_DIRECT	Protein binding	1.08E–06	4.42E–04
GOTERM_CC_DIRECT	Dendrite	1.02E–04	1.45E–02
GOTERM_CC_DIRECT	Cytosol	1.81E–04	1.59E–02
GOTERM_CC_DIRECT	Neuronal cell body	2.24E–04	1.59E–02

[a]The table shows significant pathways identified by DAVID that are enriched among genes that have a polyevidence score of 3 and above. *P* adj values represent *P* values corrected by the Benjamini–Hochberg procedure in DAVID.

genes are co-expressed in specific brain subregions. The tissue expression analysis tab allows users to explore BIP risk gene expression in distinct human tissues.

Prioritized BIP genes and enriched pathways

As well as offering users an easy-to-use online search and analysis tool, the database also prioritizes susceptibility genes of BIP risk via cumulative scores to help researchers select the most promising candidates for functional investigations. Overall, 29 prioritized candidate risk genes (score of 3 or greater) for BIP were identified by polyevidence scoring (Figure 2). Three potential risk genes, i.e. OSBPL2, STK4 and PACS1, showed the highest scores in the prioritized task, suggesting that they may represent prospective BIP susceptibility genes. OSBPL2 is located on chromosome 20aq13 and has been implicated by genome-wide significant association with nearby SNPs in BIP genetic study (33). Furthermore, diseases associated with OSBPL2 include deafness, and previous studies report a prevalence of BIP in deaf and hard-of-hearing outpatients (91, 92), with deaf youth potentially more vulnerable to BIP. STK4 is located in chromosome 20aq13.12, and genetic variants located in or near STK4 showed genome-wide significant association with BIP (31, 32). STK4-related pathways include the MAPK signaling pathway (93), and we noticed that pathways involved in the genetic predisposition to BIP include the MAPK signaling pathway as well (94). PACS1 is located in chromosome 2p13.1, and genetic variants in PACS1 showed genome-wide significant association with BIP in PGC2 (31) and PGC3 (33). Interestingly, Chen *et al.* found that overexpression of PACS1 reduced the density of dendritic spines, revealing the potential biological mechanisms of this gene in BIP (95). In addition, we used the Database for Annotation, Visualization and Integration Discovery (DAVID) (96, 97) for pathway analysis of genes of interest (i.e. polyevidence score ≥3), which were found to be significantly enriched in membrane, dendrite and neuronal cell body-related pathways (Table 2).

Discussion

BIP is a common and severe psychiatric disorder marked by episodic disturbances in mood, cognition and behavior. Both genetic and environmental risk factors participate in the pathogenesis of BIP, although its high heritability (up to 70%) points to genetic factors playing a primary role in its occurrence. Research on BIP genetic architecture has shown significant progress in recent years, and over 60 significant

susceptibility loci have been successfully identified. Despite this, BIP etiology remains poorly understood. Thus, there is a pressing need to systematically integrate multiple layers of data from diverse sources, such as genetic, gene expression, PPI, co-expression and eQTL data, to extract meaningful biological information for BIP genetic studies. Hence, to fill this gap, we developed a web-based platform that integrates multi-omics resources from different BIP studies. This is the first BIP genetic database that focuses on interpreting genetic data from GWAS based on multi-omics data and integrated analysis.

To date, only the BIP genetic database (BDgene) (98) is available for BIP research. Compared with the BDgene database, the dbBIP has several advantages. First, the BIP susceptibility genes included in the BDgene database were mainly based on small sample linkage and genetic association studies, and the database has not been updated since 2016. Second, the dbBIP offers a powerful analysis module for advanced users to perform customized analysis, including LocusZoom, eQTL, tQTL, PPI, co-expression, temporal and spatial expression pattern and tissue expression analyses. Third, the priority of candidate risk genes in the dbBIP was determined via in-depth data integration based on multi-omics data. Accordingly, the top prioritized genes (as good candidates) provide positive preliminary results, which deserve further functional characterization. Fourth, the dbBIP provides a one-stop searching resource for genes and offers comprehensive information collection from the above three modules. Lastly, as new technologies and analysis methods are rapidly evolving and novel BIP susceptibility variants and genes are constantly being identified, the dbBIP database will be updated to incorporate recent findings, thus providing a valuable and up-to-date resource for the BIP research community.

This study also has several limitations. First, a straightforward and arbitrary scoring algorithm using various data (e.g. genetic studies, integrative analysis and gene expression researches) was selected to prioritize promising BIP candidate risk genes. However, simple and operational scoring systems can miss potential overlap between the source data used for scoring (e.g. genes identified by integrative analysis also contain information from genetics and gene expression studies), which may affect score credibility. Second, we treated all evidence from different origins equally. However, GWAS tend to provide more reliable candidate risk genes than CNV studies and should be given greater weight. Multiple algorithms can be developed and used in the future to address these limitations.

Our newly developed dbBIP database offers a wealth of resources for translating genetic results and elucidating the molecular and pathogenic mechanisms underlying the occurrence and development of BIP. The database includes all recently available BIP-related data (including genetic and multi-omics data), thus allowing researchers the opportunity to access and analyze BIP susceptibility genes under a unified online tool. Therefore, the dbBIP provides a practical and convenient platform for BIP research from a genetics perspective.

Acknowledgements

available. We thank the participants and investigators of the Working Group of the Psychiatric Genomics Consortium, the GTEx Project and the PsychENCODE Consortium for generating and providing summary statistics and thereby making this work possible.

Funding

Conflict of interest

The authors declare that they have no conflicts of interest.

Author contributions

J.F.X., X.Y.L. and X.J.L. conceived and devised the study. X.Y.L. performed most of the bioinformatics analyses, including SNP functional annotation, SMR and TWAS analysis, differential expression analysis, QTL analysis, spatiotemporal expression pattern analysis and tissue expression analysis. S.S.M. undertook website construction and conducted co-expression pattern analysis. Y.W. collected PPI data. X.Y.L., S.S.M., W.H.Y. and H.K. carried out literature searching, screening and data collection. J.F.X., X.J.L., X.Y.L., S.S.M., W.H.Y., Y.W., H.K. and M.S.Z. performed data generation, analysis and interpretation of the results. X.Y.L. drafted the first version of the manuscript. J.F.X. and X.J.L. supervised the project and direction. All authors provided critical feedback and approved the final version of the manuscript.

Data availability

All data relevant to this study are included in the article or available online in the dbBIP. The data generated in this study are also available from the corresponding author upon reasonable request.

References

1. Vieta,E., Berk,M., Schulze,T.G. *et al.* (2018) Bipolar disorders. *Nat. Rev. Dis. Primers*, **4**, 18008.
2. Carvalho,A.F., Firth,J. and Vieta,E. (2020) Bipolar disorder. *N. Engl. J. Med.*, **383**, 58–66.
3. Merikangas,K.R., Jin,R., He,J.P. *et al.* (2011) Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. *Arch. Gen. Psychiatry*, **68**, 241–251.
4. Millan,M.J., Agid,Y., Brune,M. *et al.* (2012) Cognitive dysfunction in psychiatric disorders: characteristics, causes and the quest for improved therapy. *Nat. Rev. Drug Discov.*, **11**, 141–168.
5. Correll,C.U., Solmi,M., Veronese,N. *et al.* (2017) Prevalence, incidence and mortality from cardiovascular disease in patients with pooled and specific severe mental illness: a large-scale meta-analysis of 3,211,768 patients and 113,383,368 controls. *World Psychiatry*, **16**, 163–180.
6. Plans,L., Barrot,C., Nieto,E. *et al.* (2019) Association between completed suicide and bipolar disorder: a systematic review of the literature. *J. Affect. Disord.*, **242**, 111–122.
7. Bessonova,L., Ogden,K., Doane,M.J. *et al.* (2020) The economic burden of bipolar disorder in the united states: a systematic literature review. *Clinicoecon Outcomes Res.*, **12**, 481–497.
8. Craddock,N. and Jones,I. (2001) Molecular genetics of bipolar disorder. *Br J Psychiatry*, **178**, S128–S133.
9. Rowland,T.A. and Marwaha,S. (2018) Epidemiology and risk factors for bipolar disorder. *Ther. Adv. Psychopharmacol.*, **8**, 251–269.
10. Gordovez,F.J.A. and McMahon,F.J. (2020) The genetics of bipolar disorder. *Mol. Psychiatry*, **25**, 544–559.
11. Kassem,L., Lopez,V., Hedeker,D. *et al.* (2006) Familiality of polarity at illness onset in bipolar affective disorder. *Am. J. Psychiatry*, **163**, 1754–1759.
12. Grover,D., Verma,R., Goes,F.S. *et al.* (2009) Family-based association of YWHAH in psychotic bipolar disorder. *Am. J. Med. Genet. B Neuropsychiatr. Genet.*, **150B**, 977–983.
13. Craddock,N., Dave,S. and Greening,J. (2001) Association studies of bipolar disorder. *Bipolar Disord.*, **3**, 284–298.
14. Heiden,A., Schussler,P., Itzlinger,U. *et al.* (2000) Association studies of candidate genes in bipolar disorders. *Neuropsychobiology*, **42Suppl 1**, 18–21.
15. Seifuddin,F., Mahon,P.B., Judy,J. *et al.* (2012) Meta-analysis of genetic association studies on bipolar disorder. *Am. J. Med. Genet. B Neuropsychiatr. Genet.*, **159B**, 508–518.
16. Risch,N. and Merikangas,K. (1996) The future of genetic studies of complex human diseases. *Science*, **273**, 1516–1517.
17. Wellcome Trust Case Control, C. (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*, **447**, 661–678.
18. Psychiatric,G.C.B.D.W.G. (2011) Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat. Genet.*, **43**, 977–983.
19. Baum,A.E., Akula,N., Cabanero,M. *et al.* (2008) A genome-wide association study implicates diacylglycerol kinase eta (DGKH) and several other genes in the etiology of bipolar disorder. *Mol. Psychiatry*, **13**, 197–207.
20. Charney,A.W., Ruderfer,D.M., Stahl,E.A. *et al.* (2017) Evidence for genetic heterogeneity between clinical subtypes of bipolar disorder. *Transl. Psychiatry*, **7**, e993.
21. Chen,D.T., Jiang,X., Akula,N. *et al.* (2013) Genome-wide association study meta-analysis of European and Asian-ancestry samples identifies three novel loci associated with bipolar disorder. *Mol. Psychiatry*, **18**, 195–205.
22. Cichon,S., Muhleisen,T.W., Degenhardt,F.A. *et al.* (2011) Genome-wide association study identifies genetic variation in neurocan as a susceptibility factor for bipolar disorder. *Am. J. Hum. Genet.*, **88**, 372–381.
23. Ferreira,M.A., O'Donovan,M.C., Meng,Y.A. *et al.* (2008) Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. *Nat. Genet.*, **40**, 1056–1058.
24. Green,E.K., Grozeva,D., Forty,L. *et al.* (2013) Association at SYNE1 in both bipolar disorder and recurrent major depression. *Mol. Psychiatry*, **18**, 614–617.
25. Green,E.K., Hamshere,M., Forty,L. *et al.* (2013) Replication of bipolar disorder susceptibility alleles and identification of two novel genome-wide significant associations in a new bipolar disorder case-control sample. *Mol. Psychiatry*, **18**, 1302–1307.
26. Hou,L., Bergen,S.E., Akula,N. *et al.* (2016) Genome-wide association study of 40,000 individuals identifies two novel loci associated with bipolar disorder. *Hum. Mol. Genet.*, **25**, 3383–3394.
27. Ligthart,S., Vaez,A., Vosa,U. *et al.* (2018) Genome analyses of >200,000 individuals identify 58 Loci for chronic inflammation and highlight pathways that link inflammation and complex disorders. *Am. J. Hum. Genet.*, **103**, 691–706.
28. Schulze,T.G., Detera-Wadleigh,S.D., Akula,N. *et al.* (2009) Two variants in Ankyrin 3 (ANK3) are independent genetic risk factors for bipolar disorder. *Mol. Psychiatry*, **14**, 487–491.
29. Scott,L.J., Muglia,P., Kong,X.Q. *et al.* (2009) Genome-wide association and meta-analysis of bipolar disorder in individuals of European ancestry. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 7501–7506.

30. Smith,E.N., Bloss,C.S., Badner,J.A. *et al.* (2009) Genome-wide association study of bipolar disorder in European American and African American individuals. *Mol. Psychiatry*, **14**, 755–763.

31. Stahl,E.A., Breen,G., Forstner,A.J. *et al.* (2019) Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat. Genet.*, **51**, 793–803.

32. Li,H.J., Zhang,C., Hui,L. *et al.* (2021) Novel risk loci associated with genetic risk for bipolar disorder among Han Chinese individuals: a genome-wide association study and meta-analysis. *JAMA Psychiatry*, **78**, 320–330.

33. Mullins,N., Forstner,A.J., O'Connell,K.S. *et al.* (2021) Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. *Nat. Genet.*, **53**, 817–829.

34. Green,E.K., Rees,E., Walters,J.T. *et al.* (2016) Copy number variation in bipolar disorder. *Mol. Psychiatry*, **21**, 89–93.

35. Malhotra,D., McCarthy,S., Michaelson,J.J. *et al.* (2011) High frequencies of de novo CNVs in bipolar disorder and schizophrenia. *Neuron*, **72**, 951–963.

36. Zhang,C., Xiao,X., Li,T. *et al.* (2021) Translational genomics and beyond in bipolar disorder. *Mol. Psychiatry*, **26**, 186–202.

37. Akula,N., Marenco,S., Johnson,K. *et al.* (2021) Deep transcriptome sequencing of subgenual anterior cingulate cortex reveals cross-diagnostic and diagnosis-specific RNA expression changes in major psychiatric disorders. *Neuropsychopharmacology*, **46**, 1364–1372.

38. Zhu,Z., Zhang,F., Hu,H. *et al.* (2016) Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.*, **48**, 481–487.

39. Kirov,G., Rees,E., Walters,J.T. *et al.* (2014) The penetrance of copy number variations for schizophrenia and developmental delay. *Biol. Psychiatry*, **75**, 378–385.

40. Leppa,V.M., Kravitz,S.N., Martin,C.L. *et al.* (2016) Rare inherited and De Novo CNVs reveal complex contributions to ASD risk in multiplex families. *Am. J. Hum. Genet.*, **99**, 540–554.

41. Pinto,D., Pagnamenta,A.T., Klei,L. *et al.* (2010) Functional impact of global rare copy number variation in autism spectrum disorders. *Nature*, **466**, 368–372.

42. Grozeva,D., Kirov,G., Ivanov,D. *et al.* (2010) Rare copy number variants: a point of rarity in genetic risk for bipolar disorder and schizophrenia. *Arch. Gen. Psychiatry*, **67**, 318–327.

43. Moreno-De-Luca,D., Sanders,S.J., Willsey,A.J. *et al.* (2013) Using large clinical data sets to infer pathogenicity for rare copy number variants in autism cohorts. *Mol. Psychiatry*, **18**, 1090–1095.

44. Charney,A.W., Stahl,E.A., Green,E.K. *et al.* (2019) Contribution of rare copy number variants to bipolar disorder risk is limited to schizoaffective cases. *Biol. Psychiatry*, **86**, 110–119.

45. Bergen,S.E., O'Dushlaine,C.T., Ripke,S. *et al.* (2012) Genome-wide association study in a Swedish population yields support for greater CNV and MHC involvement in schizophrenia compared with bipolar disorder. *Mol. Psychiatry*, **17**, 880–886.

46. Priebe,L., Degenhardt,F.A., Herms,S. *et al.* (2012) Genome-wide survey implicates the influence of copy number variants (CNVs) in the development of early-onset bipolar disorder. *Mol. Psychiatry*, **17**, 421–432.

47. Chen,J., Calhoun,V.D., Perrone-Bizzozero,N.I. *et al.* (2016) A pilot study on commonality and specificity of copy number variants in schizophrenia and bipolar disorder. *Transl Psychiatry*, **6**, e824.

48. Georgieva,L., Rees,E., Moran,J.L. *et al.* (2014) De novo CNVs in bipolar affective disorder and schizophrenia. *Hum. Mol. Genet.*, **23**, 6677–6683.

49. Jia,X., Goes,F.S., Locke,A.E. *et al.* (2021) Investigating rare pathogenic/likely pathogenic exonic variation in bipolar disorder. *Mol. Psychiatry*, **26**.

50. Kataoka,M., Matoba,N., Sawada,T. *et al.* (2016) Exome sequencing for bipolar disorder points to roles of de novo loss-of-function and protein-altering mutations. *Mol. Psychiatry*, **21**, 885–893.

51. Wynn,J., Martinez,J., Duong,J. *et al.* (2015) Association of researcher characteristics with views on return of incidental findings from genomic research. *J. Genet. Couns.*, **24**, 833–841.

52. Wang,W., Corominas,R. and Lin,G.N. (2019) De novo mutations from whole exome sequencing in neurodevelopmental and psychiatric disorders: from discovery to application. *Front. Genet*, **10**, 258.

53. Kircher,M., Witten,D.M., Jain,P. *et al.* (2014) A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.*, **46**, 310–315.

54. Huang,Y.F., Gulko,B. and Siepel,A. (2017) Fast, scalable prediction of deleterious noncoding variants from functional and population genomic data. *Nat. Genet.*, **49**, 618–624.

55. Boyle,A.P., Hong,E.L., Hariharan,M. *et al.* (2012) Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res.*, **22**, 1790–1797.

56. Consortium,E.P. (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature*, **489**, 57–74.

57. Zhang,S., Zhang,H., Zhou,Y. *et al.* (2020) Allele-specific open chromatin in human iPSC neurons elucidates functional disease variants. *Science*, **369**, 561–565.

58. Boughton,A.P., Welch,R.P., Flickinger,M. *et al.* (2021) Locus-Zoom.js: interactive and embeddable visualization of genetic association study results. *Bioinformatics*, **37**: 3017–3018.

59. Gusev,A., Ko,A., Shi,H. *et al.* (2016) Integrative approaches for large-scale transcriptome-wide association studies. *Nat. Genet.*, **48**, 245–252.

60. Fromer,M., Roussos,P., Sieberts,S.K. *et al.* (2016) Gene expression elucidates functional impact of polygenic risk for schizophrenia. *Nat. Neurosci.*, **19**, 1442–1453.

61. Collado-Torres,L., Burke,E.E., Peterson,A. *et al.* (2019) Regional heterogeneity in gene expression, regulation, and coherence in the frontal cortex and hippocampus across development and schizophrenia. *Neuron*, **103**, 203–216 e8.

62. Gandal,M.J., Zhang,P., Hadjimichael,E. *et al.* (2018) Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. *Science*, **362**.

63. Yang,C.P., Li,X., Wu,Y. *et al.* (2018) Comprehensive integrative analyses identify GLT8D1 and CSNK2B as schizophrenia risk genes. *Nat. Commun.*, **9**, 838.

64. Li,X., Su,X., Liu,J. *et al.* (2021) Transcriptome-wide association study identifies new susceptibility genes and pathways for depression. *Transl Psychiatry*, **11**, 306.

65. Võsa,U., Claringbould,A., Westra,H.-J. *et al.* (2018) Unraveling the polygenic architecture of complex traits using blood eQTL metaanalysis. *bioRxiv*, 447367.

66. Chen,H., Wang,N., Zhao,X. *et al.* (2013) Gene expression alterations in bipolar disorder postmortem brains. *Bipolar Disord.*, **15**, 177–187.

67. Huttlin,E.L., Ting,L., Bruckner,R.J. *et al.* (2015) The BioPlex network: a systematic exploration of the human interactome. *Cell*, **162**, 425–440.

68. Hein,M.Y., Hubner,N.C., Poser,I. *et al.* (2015) A human interactome in three quantitative dimensions organized by stoichiometries and abundances. *Cell*, **163**, 712–723.

69. Li,T., Wernersson,R., Hansen,R.B. *et al.* (2017) A scored human protein-protein interaction network to catalyze genomic interpretation. *Nat. Methods*, **14**, 61–64.

70. Ziats,M.N. and Rennert,O.M. (2014) Identification of differentially expressed microRNAs across the developing human brain. *Mol. Psychiatry*, **19**, 848–852.

71. Kang,H.J., Kawasawa,Y.I., Cheng,F. *et al.* (2011) Spatio-temporal transcriptome of the human brain. *Nature*, **478**, 483–489.

72. Gulsuner,S., Walsh,T., Watts,A.C. *et al.* (2013) Spatial and temporal mapping of de novo mutations in schizophrenia to a fetal prefrontal cortical network. *Cell*, **154**, 518–529.

73. Colantuoni,C., Lipska,B.K., Ye,T. *et al.* (2011) Temporal dynamics and genetic control of transcription in the human prefrontal cortex. *Nature*, **478**, 519–523.

74. Consortium,G.T., Laboratory,D.A., Coordinating Center -Analysis Working, G. *et al.* (2017) Genetic effects on gene expression across human tissues. *Nature*, **550**, 204–213.

75. Barr,C.L. and Misener,V.L. (2016) Decoding the non-coding genome: elucidating genetic risk outside the coding genome. *Genes Brain Behav.*, **15**, 187–204.

76. O'Brien,H.E., Hannon,E., Hill,M.J. *et al.* (2018) Expression quantitative trait loci in the developing human brain and their enrichment in neuropsychiatric disorders. *Genome Biol.*, **19**, 194.

77. Bin,Y., Zhang,W., Tang,W. *et al.* (2020) Prediction of neuropeptides from sequence information using ensemble classifier and hybrid features. *J. Proteome Res.*, **19**, 3732–3740.

78. Ayalew,M., Le-Niculescu,H., Levey,D.F. *et al.* (2012) Convergent functional genomics of schizophrenia: from comprehensive understanding to genetic risk prediction. *Mol. Psychiatry*, **17**, 887–905.

79. Ament,S.A., Szelinger,S., Glusman,G. *et al.* (2015) Rare variants in neuronal excitability genes influence risk for bipolar disorder. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, 3576–3581.

80. Lescai,F., Als,T.D., Li,Q. *et al.* (2017) Whole-exome sequencing of individuals from an isolated population implicates rare risk variants in bipolar disorder. *Transl. Psychiatry*, **7**, e1034.

81. Toma,C., Shaw,A.D., Allcock,R.J.N. *et al.* (2018) An examination of multiple classes of rare variants in extended families with bipolar disorder. *Transl. Psychiatry*, **8**, 65.

82. Forstner,A.J., Fischer,S.B., Schenk,L.M. *et al.* (2020) Whole-exome sequencing of 81 individuals from 27 multiply affected bipolar disorder families. *Transl. Psychiatry*, **10**, 57.

83. Goes,F.S., Pirooznia,M., Tehan,M. *et al.* (2019) De novo variation in bipolar disorder. *Mol. Psychiatry*, **26**, 4127–4136.

84. Maaser,A., Forstner,A.J., Strohmaier,J. *et al.* (2018) Exome sequencing in large, multiplex bipolar disorder families from Cuba. *PLoS One*, **13**, e0205895.

85. Palmer,D.S., Howrigan,D.P., Chapman,S.B. *et al.* (2021) Exome sequencing in bipolar disorder reveals shared risk gene AKAP11 with schizophrenia. *medRxiv*. 2021.03.09.21252930.

86. Husson,T., Duboc,J.B., Quenez,O. *et al.* (2018) Identification of potential genetic risk factors for bipolar disorder by whole-exome sequencing. *Transl. Psychiatry*, **8**, 268.

87. Zhang,T., Hou,L., Chen,D.T. *et al.* (2018) Exome sequencing of a large family identifies potential candidate genes contributing risk to bipolar disorder. *Gene*, **645**, 119–123.

88. Goes,F.S., Pirooznia,M., Parla,J.S. *et al.* (2016) Exome sequencing of familial bipolar disorder. *JAMA Psychiatry*, **73**, 590–597.

89. Toma,C., Shaw,A.D., Overs,B.J. *et al.* (2020) De novo gene variants and familial bipolar disorder. *JAMA Netw. Open*, **3**, e203382.

90. Fagerberg,L., Hallstrom,B.M., Oksvold,P. *et al.* (2014) Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol. Cell Proteomics*, **13**, 397–406.

91. Diaz,D.R., Landsberger,S.A., Povlinski,J. *et al.* (2013) Psychiatric disorder prevalence among deaf and hard-of-hearing outpatients. *Compr. Psychiatry*, **54**, 991–995.

92. Landsberger,S.A., Diaz,D.R., Spring,N.Z. *et al.* (2014) Psychiatric diagnoses and psychosocial needs of outpatient deaf children and adolescents. *Child Psychiatry Hum. Dev.*, **45**, 42–51.

93. Liu,H., Wang,L.E., Liu,Z. *et al.* (2013) Association between functional polymorphisms in genes involved in the MAPK signaling pathways and cutaneous melanoma risk. *Carcinogenesis*, **34**, 885–892.

94. Do Prado,C.H., Rizzo,L.B., Wieck,A. *et al.* (2013) Reduced regulatory T cells are associated with higher levels of Th1/TH17 cytokines and activated MAPK in type 1 bipolar disorder. *Psychoneuroendocrinology*, **38**, 667–676.

95. Chen,R., Yang,Z., Liu,J. *et al.* (2022) Functional genomic analysis delineates regulatory mechanisms of GWAS-identified bipolar disorder risk variants. *Genome Med.*, **14**, 53.

96. Huang,D.W., Sherman,B.T., Tan,Q. *et al.* (2007) DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Res.*, **35**, W169–W175.

97. Huang,D.W., Sherman,B.T., Tan,Q. *et al.* (2007) The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biol.*, **8**, R183.

98. Chang,S.H., Gao,L., Li,Z. *et al.* (2013) BDgene: a genetic database for bipolar disorder and its overlap with schizophrenia and major depressive disorder. *Biol. Psychiatry*, **74**, 727–733.