Review

Genomic and neuroimaging approaches to bipolar disorder

Mojtaba Oraki Kohshour, Sergi Papiol, Christopher R. K. Ching and Thomas G. Schulze

Background

To date, besides genome-wide association studies, a variety of other genetic analyses (e.g. polygenic risk scores, whole-exome sequencing and whole-genome sequencing) have been conducted, and a large amount of data has been gathered for investigating the involvement of common, rare and very rare types of DNA sequence variants in bipolar disorder. Also, non-invasive neuroimaging methods can be used to quantify changes in brain structure and function in patients with bipolar disorder.

Aims

To provide a comprehensive assessment of genetic findings associated with bipolar disorder, based on the evaluation of different genomic approaches and neuroimaging studies.

Method

We conducted a PubMed search of all relevant literatures from the beginning to the present, by querying related search strings.

Results

ANK3, CACNA1C, SYNE1, ODZ4 and *TRANK1* are five genes that have been replicated as key gene candidates in bipolar disorder pathophysiology, through the investigated studies. The percentage of phenotypic variance explained by the identified variants is small (approximately 4.7%). Bipolar disorder polygenic risk scores are associated with other psychiatric phenotypes. The ENIGMA-BD studies show a replicable pattern of lower cortical thickness, altered white matter integrity and smaller subcortical volumes in bipolar disorder.

Conclusions

The low amount of explained phenotypic variance highlights the need for further large-scale investigations, especially among non-European populations, to achieve a more complete understanding of the genetic architecture of bipolar disorder and the missing heritability. Combining neuroimaging data with genetic data in large-scale studies might help researchers acquire a better knowledge of the engaged brain regions in bipolar disorder.

Bipolar disorder, a severe psychiatric condition

Bipolar disorder is a chronic, disabling and severe mental health condition that is defined by recurrent acute manic and depressive mood states, and affects >1% of the global population.[1] It is considered as one of the most heritable mental illnesses, with heritability estimates of approximately 80%, and is thought to be influenced by gene×environment interactions. Bipolar disorder ranks as the sixth most common cause of disability-adjusted life-years in human populations (9.9 million years worldwide), and is associated with a higher risk for severe conditions such as suicidality and social and cognitive functional disability.[2–4]

Psychiatric genetics, formed in the 20th century, is a branch of behavioural genetics that investigates the role of genetics in the onset and progression of mental illnesses. Attempts to use genomic data for prediction and risk stratification have been classically hampered by the highly polygenic and pleiotropic nature of psychiatric genetics.[5] Current models suggest that bipolar disorder is a heterogeneous and multifactorial disease, which probably involves several pathogenic pathways orchestrated by numerous common and rare genetic factors, environmental factors and epigenetic influences.[6] Over the past several decades, many genetic studies have been performed to uncover the genetic underpinnings of bipolar disorder. For complex disorders like bipolar disorder, association studies are expected to be more informative than linkage studies because of their higher statistical power to capture modest genetic effects. However, initial candidate gene approaches suffered a challenging lack of replicability of the results that raised questions regarding the actual utility of association studies in psychiatric genetics.[7]

Genome-wide association studies (GWASs), which were introduced as a powerful alternative to linkage and a hypothesis-free extension of candidate gene studies, have emerged as a revolutionary approach in research on complex genetic disorders. GWASs aim to reveal genotype-phenotype associations by testing the association between millions of common genetic variants distributed across the human genome with the phenotype of interest. GWASs have provided compelling evidence of the ability to unravel the genetic underpinnings of complex traits.[8–12] One important lesson from GWASs is that the cumulative number of loci to be discovered in GWASs is growing in line with the increasing sample size.[13] According to an estimate in 2013, >63 000 case–control samples would be required to identify the approximately 105 bipolar disorder risk loci discoverable by GWAS, and these would explain <6% of the inherited risk.[14] Recently, the decreasing cost of genotyping technology is encouraging the development of new, large GWASs through the formation of multinational research consortia.[13]

The first GWAS ever published, a case versus control association study including 92 788 gene-based single-nucleotide polymorphism (SNP) markers, was performed in Japan in 2002.[15] By genotyping 94 individuals of Japanese descent with myocardial infarction, and comparing the results with allelic frequencies

found in 658 healthy individuals at first, and then following the results in the expanded sample size (1133 cases and 1006 controls), Ozaki et al reported a candidate locus on chromosome 6p21 to be associated with susceptibility to myocardial infarction.[15] The first GWAS on bipolar disorder was published in 2007, and >20 GWASs to date have followed this initial study.[16] These studies have reported >120 lead SNPs ($P < 5 \times 10^{-8}$) contributing to bipolar disorder risk located on almost all human chromosomes, and have mapped the genes in their corresponding loci.[16,17] In GWASs, genotype imputation is used to expand the number of tested variants, complete fine-mapping, and permit meta-analysis with different arrays by harmonising data-sets, increasing the power of association. Imputation is an *in silico* method and can be performed throughout the entire genome, or in a more targeted area with the aim of predicting genotypes at the SNPs that are not genotyped directly in the samples.[18,19] Most of the top associated variants identified by GWASs are located in non-coding regions of the genome, suggesting that many of them might be involved in regulatory mechanisms of gene expression (expression quantitative trait loci) rather than directly affecting the amino acid sequence of proteins.[20]

Over the past decade, significant advances and cost reductions in sequencing technology have made it possible to sequence the whole exome and genome of a large number of patients with bipolar disorder and controls.[21] Based on the National Human Genome Research Institute data, in mid-2015, constructing a high-quality, whole human genome sequencing draft costed slightly over $4000. By the end of 2015, it had dropped to <$1500. Today, whole-exome and whole-genome sequences are frequently available for <$1000 on the commercial market. Whole-exome sequencing (WES) is less costly than whole-genome sequencing (WGS), but it has the disadvantage that it sequences only about 1% of the genome because it focuses on protein-coding regions (exons) to find variations that affect gene structures and functions.[22,23] The complementary approach, WGS, is a novel but still relatively costly technique that captures all allelic variation throughout the entire genome. It allows for the exploration of the non-coding genome, which can lead to a high level of understanding of developmental stages, brain regions, cell types and biological functions.[24] Nevertheless, sequencing the whole genome and exome in considerable numbers of patients with bipolar disorder may lead to the identification of new rare and very rare genetic variants associated with bipolar disorder, characterised by larger effect sizes than common variants.

In genetic association studies, which investigate many genetic variants to detect disease-associated ones, appropriate adjustment for multiple comparison is needed for distinguishing true association and avoiding a flood of false-positive cases.[25] The effectiveness of genetic association studies to discover causative genes of complex disorders requires a sample size with appropriate statistical power, which will be influenced significantly by disease prevalence, disease allele frequency, linkage disequilibrium, inheritance models and effect size of the genetic variants. The smallest sample size required to generate appropriate statistical power (e.g. 80% power) is known as an effective sample size. In large-scale association studies, the statistical power of 80% is widely used to eliminate false-negative associations and establish the most cost-effective sample size.[26]

The evidence collected over the past years indicates that bipolar disorder has an extremely polygenic architecture, with many genetic variants shaping the individual genetic vulnerability for this disorder. Polygenic risk scores (PRS), also referred to as polygenic scores, provide an estimate of this genetic vulnerability and represent a genetic tool with the potential to identify people who are at risk for bipolar disorder or for a worse course of the disease.[27]

Although the notion of endophenotypes has not yet proven to be as successful as expected for identifying genetic factors underlying the clinical heterogeneity of patients with bipolar disorder, it still holds interest in the context of large genomic surveys, particularly in the case of brain imaging. Magnetic resonance imaging (MRI) is a powerful, non-invasive method that allows researchers to survey brain morphology and function, to discover candidate endophenotypes and perhaps delineate pathways between genetic risk factors in bipolar disorder, alterations in the brain and the clinical manifestation of this disorder, including response to treatment.[28]

In this narrative review article, we provide an overview of genetic and neuroimaging studies on bipolar disorder by highlighting their significant results. The review includes GWASs and studies on WES and WGS that were identified in the PubMed database, as well as PRS and brain imaging sections. The definition of the key terms and abbreviations for genes are presented in the Supplementary Tables 1 and 2 available at https://doi.org/10.1192/bjo.2021.1082 (Glossary of Key Terms and Genes Abbreviation); however, readers who are unfamiliar with all of the terms, or who want to know more, are recommended to study the methodological details and technical points at,[29,30] and.[31]

GWASs

In GWASs that are carried out following a case–control study design, cases fulfil the clinical criteria for a given disorder but controls do not. Following genotyping of a large number of genetic variants encompassing the whole genome in all samples, each SNP will be tested for disease association by statistically comparing allele frequencies in the group of cases versus the group of controls (Fig. 1). A significant difference of a particular SNP or locus indicates that it may play an aetiological role. Here, a typical type 1 error threshold for genome-wide significance is often considered 5×10^{-8}, providing a multiple-testing correction threshold suitable in a genomic context that involves the analysis of several millions of genetic variants.[32]

We queried PubMed with the search string 'genome-wide association studies and bipolar disorder', and then searched for the original articles that reported on GWASs on data from patients with bipolar disorder and had significant results ($P < 5 \times 10^{-8}$). These studies are described in the following paragraphs. Although the first GWAS on bipolar disorder did not generate results that were significant at the specified threshold, we start our overview with a brief description of its findings.

The first GWAS on bipolar disorder was published in 2007 and included approximately 2000 British patients and 3000 controls. This GWAS did not find a genome-wide significant SNP for bipolar disorder and only identified rs420259 ($P = 6.3 \times 10^{-8}$) on chromosome 16p12 as the SNP with the strongest association. *PALB2*, *NDUFAB1* and *DCTN5* are located at this locus. *PALB2* plays a role in the stability of critical nuclear structures, *NDUFAB1* encodes a subunit of complex I of the mitochondrial respiratory chain and *DCTN5* encodes a protein that is involved in intracellular transport.[33]

In 2008, Baum et al published another of the first GWASs on bipolar disorder. The combined sample, which included a test sample (National Institute of Mental Health: 413 cases, 563 controls) and replication sample (Germany: 772 cases, 876 controls), comprised 1185 cases and 1439 controls. The group identified rs1012053 in the first intron of *DGKH* ($P = 1.5 \times 10^{-8}$; odds ratio 1.59) as an SNP with a strong significant association with bipolar disorder. *DGKH* is located within the bipolar disorder linkage region on chromosome 13q14 and encodes diacylglycerol kinase
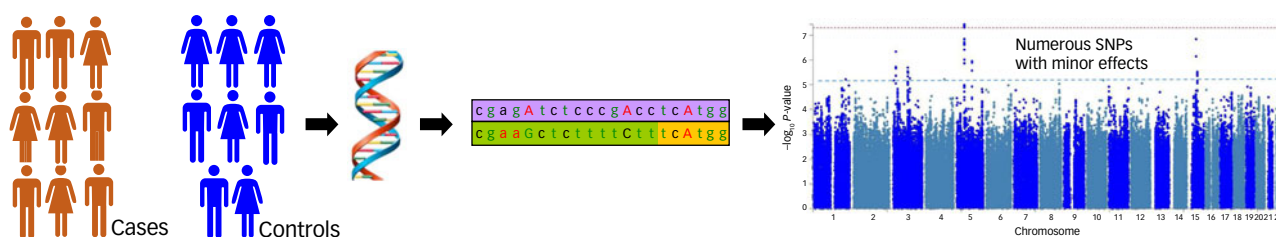
Fig. 1 Genome-Wide Association Studies (GWASs). SNP, single nucleotide polymorphism.

eta, an important protein in the lithium-sensitive phosphatidylino-sitol pathway.[10]

Also in 2008, Sklar et al reported on a GWAS on 1461 patients with bipolar disorder and 2008 controls. The authors did not detect SNPs that met the threshold for genome-wide significance, but haplotype analysis identified a haplotype in *MYO5B* on chromosome 18 that reached the genome-wide significance threshold (haplotypic $P = 2.04 \times 10^{-8}$; odds ratio 1.51). In this region, rs4939921 in the intronic region of *MYO5B* was the SNP with the strongest evidence for association. *MYO5B* is a large and brain-expressed gene that is involved in vesicle trafficking regulation, and encodes a multifunctional protein in dendritic spines and other brain structures.[34]

Also in 2008, Ferreira et al published a collaborative GWAS in which they combined their data-set (ED-DUB-STEP2: Included additional samples from the Systematic Treatment Enhancement Program for Bipolar Disorder study as well as cases and controls from the University of Edinburgh and Trinity College Dublin) with two other data-sets (from the Wellcome Trust Case-Control Consortium and Systematic Treatment Enhancement Program-University College London studies) to obtain a larger sample size (4387 cases, 6209 controls) to improve the power for identifying risk alleles of modest effect. The analysis detected the imputed SNP rs10994336 ($P = 9.1 \times 10^{-9}$; odds ratio 1.45) in *ANK3* on chromosome 10q21 as the strongest finding that surpassed the genome-wide significance threshold. The ANK3 protein was first found at the axonal initial segment and is involved in the regulation of voltage-gated sodium channels. The authors also reported on another variant in the third intron of the *CACNA1C* on chromosome 12p13 as the second strongest region of association (rs1006737; $P = 7.0 \times 10^{-8}$; odds ratio 1.181). They concluded that bipolar disorder pathogenesis may be caused by ion channelopathies.[35]

In 2010, Wang et al published a meta-analysis on 653 cases of bipolar disorder and 1034 controls, and 1172 cases of schizophrenia and 1379 controls. In their combined analysis of both disorders (1825 cases and 2413 controls), they detected two genome-wide significant SNPs that did not reach the threshold before in either of the two separate samples. One of the SNPs, rs11789399 ($P = 5.56 \times 10^{-9}$) on chromosome 9q33.1, is close to *ASTN2*, which is associated with schizophrenia, and the other, rs12201676 ($P = 3.88 \times 10^{-8}$) on chromosome 6q15, is near to *GABRR1* and *GABRR2*. ASTN2 is involved in promoting the migrations of neuronal precursors and modulating synaptic activity,[36] and GABRR1 and GABRR2 are two components of gamma-aminobutyric acid (GABA)-type alpha receptor complex that have important roles in many physiological functions in the brain.[37] The results of the meta-analysis and haplotype analyses of the detected loci represented further evidence of the overlap in genetic risk between schizophrenia and bipolar disorder.[38]

In 2011, Cichon et al published a GWAS and two follow-up replication studies. The best result in their GWAS on 682 patients with bipolar disorder and 1300 controls was a P-value of 3.42×10^{-6} and an odds ratio of 1.53 for rs1064395 in *NCAN* on chromosome 19p13.11. After combining this primary analysis with replication samples, this SNP reached genome-wide significance. For rs1064395, the first combined study (GWAS plus replication study 1: 2411 cases, 3613 controls) found a P-value of 3.02×10^{-8} and an odds ratio of 1.31, and the second combined study (GWAS plus replication studies 1 and 2: 8441 cases, 35 362 controls) found a P-value of 2.14×10^{-9} and an odds ratio of 1.17. The researchers concluded that the results support the idea that genetic variation in *NCAN*, which is highly expressed in the brain, is involved in the aetiology of bipolar disorder. Neurocan is an extracellular matrix glycoprotein that is thought to play a role in cell adhesion and migration. Laboratory mice showed localised expression of *NCAN* in the cortex and hippocampus of the brain, and unusual changes in these regions have been reported in patients with bipolar disorder.[9]

Also in 2011, Kerner et al presented a GWAS on 1000 White patients with bipolar disorder and 1034 controls. Most of the patients in this study had one or more comorbidities, such as substance misuse, obsessive–compulsive disorder, attention-deficit hyperactivity disorder, alcohol misuse and substance dependence. The study identified two genome-wide significant SNPs: rs1039002 and rs2727943. rs1039002 ($P = 1.7 \times 10^{-8}$) is a low-frequency variant on chromosome 6q27 that is associated with bipolar disorder in patients with psychosis and/or substance misuse in the absence of alcohol dependence. The gene with a known function that is nearest to this variant is *PDE10A*, which is highly expressed in the brain and plays a role in the elimination of cyclic adenosine monophosphate (cAMP) and cyclic guanosine monophosphate (cGMP); the inhibition of this gene seems to have therapeutic effects on psychotic symptoms. rs2727943 ($P = 3.3 \times 10^{-8}$) is on chromosome 3p26.3 and is associated with bipolar disorder in patients with alcohol dependence and other comorbidities. It is located in the intergenic region between *BIG-2* and *CNTN6*, and the encoded proteins may be involved in axon connections in the brain.[39]

As described in their 2011 publication, Sklar et al, as part of the Psychiatric GWAS Consortium Bipolar Disorder Working Group, performed a GWAS on primary samples (7481 bipolar disorder cases, 9250 controls), analysed the results in the replication sample (4493 independent cases, 42 542 independent controls) and performed a meta-analysis in the combined samples (11 974 cases, 51 792 controls). In the primary GWAS, rs10994397 ($P_{\text{genomic control}} = 7.1 \times 10^{-9}$; odds ratio 1.35) in *ANK3* on chromosome 10q21, and rs9371601 ($P_{\text{genomic control}} = 4.3 \times 10^{-8}$; odds ratio 1.15) in *SYNE1* on chromosome 6q25, were significant. Because of population substructure or cryptic relatedness, false associations might occur in case–control studies, increasing the variance of the trend test. The genomic control approach is used to calculate the variance inflation factor and adjust the test statistic. In the combined analysis (primary GWAS and replication samples),

these two variants did not reach the genome-wide significance threshold ($P = 1.52 \times 10^{-7}$ and $P = 6.14 \times 10^{-8}$, respectively), but four other SNPs did: rs4765913 ($P = 1.82 \times 10^{-9}$; odds ratio 1.14) in *CACNA1C* on chromosome 12; rs10896135 ($P = 2.77 \times 10^{-8}$; odds ratio 0.89) in a region of linkage disequilibrium (the non-random association of alleles at different loci in a given population) on chromosome 11, which contains 33 genes; rs12576775 ($P = 6.32 \times 10^{-9}$; odds ratio 0.88) in *ODZ4* on chromosome 11; and rs7296288 ($P = 8.6 \times 10^{-9}$; odds ratio 0.90) in a region of linkage disequilibrium of approximately 100 kb on chromosome 12, which contains seven genes. These results confirmed the evidence for an association of both *CACNA1C* and *ODZ4* with bipolar disorder. *CACNA1C* encodes an alpha-1 subunit of a voltage-dependent calcium channel, which is a key component in neuronal plasticity, and the protein encoded by *ODZ4* is one of the teneurins, a conserved family of cell-surface transmembrane proteins that plays a role in intercellular signalling, neuronal path-finding and proper connectivity throughout development.[40,41]

In a study reported in 2012, Green et al, in a combined analysis of their independent UK samples (1527 cases, 1579 controls) with the Psychiatric Genomics Consortium Bipolar Disorder Working Group (PGC-BD) data, found genome-wide significant evidence for the association of rs9371601 with bipolar disorder ($P = 2.9 \times 10^{-8}$; odds ratio 1.104). rs9371601 is located in intron 16 of *SYNE1*, which encodes nesprin-1 as a part of the complex that links the nucleoskeleton to the cytoskeleton. Some mutations in *SYNE1* cause neurodegenerative and neuromuscular dystrophy disorders, and increase risk for recurrent depression.[42]

In 2013, Chen et al published a GWAS meta-analysis of samples of European and Asian ancestry (6658 cases and 7155 controls). In phase 2 of their analysis, which tested the most significant findings from phase 1, the researchers examined an extended sample of 7773 cases and 9883 controls. They reported five significant SNPs ($P < 5 \times 10^{-8}$) for bipolar disorder in five regions; two of them had been reported before (rs7618915 within the 3p21 region and rs4948418 near *ANK3*), but three of them were novel: rs9834970 near *TRANK1*, rs2271893 near *LMAN2L* and rs4650608 near *PTGFR*. *TRANK1* on chromosome 3p22.2 encodes a *P*-loop containing nucleoside triphosphate hydrolase, which is associated with DNA/ATP binding or DNA repair; relevant cell line experiments suggest that the causal variant(s) in *TRANK1* result in loss of function. *LMAN2L* on chromosome 2 may be an ERGIC-53 (ER-Golgi intermediate compartment 53) regulator and play a role in the regulation of glycoprotein export from the endoplasmic reticulum. In addition, *PTGFR* on chromosome 1p31.1 is a member of the G-protein-coupled receptor (GPCR) family. All three of these genes have high expression in the brain.[14]

Also in 2013, Green et al published a meta-analysis of a combination of their independent sample data (1218 cases, 2913 controls) and a PGC-BD data-set (7481 bipolar disorder cases, 9250 controls). The analysis found four SNPs that surpassed the genome-wide significance threshold. Two of these, rs12576775 in *ODZ4* ($P = 6.20 \times 10^{-9}$; odds ratio 0.88) on chromosome 11 and rs4765913 in *CACNA1C* ($P = 9.78 \times 10^{-10}$; odds ratio 1.14) on chromosome 12, had been reported before; however, the other two, rs7296288 ($P = 8.97 \times 10^{-9}$; odds ratio 0.90) in the intergenic region between *RHEBL1* and *DHH* on chromosome 12q13.1 and rs3818253 ($P = 3.88 \times 10^{-8}$; odds ratio 1.16) in the intronic part of *TRPC4AP* on chromosome 20q11.2, represented novel findings.[43]

In 2014, Mühleisen et al reported on a GWAS on 9747 cases and 14 278 controls. The authors reported 56 genome-wide significant SNPs in five regions: *ANK3* on chromosome10q21.2 with 26 SNPs (top SNP: rs10994415-C, $P_{genomic\ control} = 6.88 \times 10^{-11}$; odds ratio 1.27), *TRANK1* on chromosome 3p22.2 with 14 SNPs (top SNP: rs6550435-G, $P_{genomic\ control} = 2.05 \times 10^{-8}$; odds ratio 1.13),

ODZ4 on chromosome 11q14.1 with ten SNPs (top SNP: rs12290811-A, $P_{genomic\ control} = 1.09 \times 10^{-9}$; odds ratio 1.19), *MIR2113-POU3F2* region on chromosome 6q16.1 with four SNPs (top SNP: rs12202969-A, $P_{genomic\ control} = 1.08 \times 10^{-8}$; odds ratio 1.12) and *ADCY2* on chromosome 5p15.31 with two SNPs (top SNP: rs17826816-G, $P_{genomic\ control} = 9.89 \times 10^{-9}$; odds ratio 1.14). This GWAS confirmed three previous loci (*ANK3*, *ODZ4* and *TRANK1*) and identified two novel loci (*ADCY2* and the region between *MIR2113* and *POU3F2*). ADCY2 is a cell-membrane-bound key enzyme in cAMP signalling and is involved in cAMP-dependent GPCR pathways. Although no specific gene can be considered for the region between *MIR2113* and *POU3F2* on chromosome 6q16.1, rs1906252 (one of the four SNPs in this region with genome-wide significance in this study) was found to be correlated with speed of information processing and cognitive phenotype in a GWAS on information processing speed.[44]

In 2016, in the context of a lithium pharmacogenomics study, Song et al published a GWAS that was based on Swedish and UK patients with bipolar disorder, and compared lithium responders with non-responders and lithium responders with healthy controls. The latter comparison, which analysed objective criteria (387 responders and 6684 controls), reported a variant with genome-wide significance, rs146727601 ($P = 1.22 \times 10^{-9}$; odds ratio 4.12) in *PTS* on chromosome 11q22.4. *PTS* encodes 6-pyruvoyltetrahy-dropterin synthase, which acts as a catalyst in the regulation of serotonin biosynthesis and nitric oxide synthase activity that, in turn, have been suggested as potential therapeutic targets for lithium. rs146727601, which is a two-base deletion, is also close to the *PLET1*. Song et al concluded that focusing on bipolar disorder subtypes could facilitate discovering the aetiology of bipolar disorder.[45,46]

Also in 2016, Hou et al published one of the largest GWASs, on 9784 patients with bipolar disorder and 30 471 controls. In their study, six autosomal loci surpassed the genome-wide significance threshold. Multiple SNPs in high linkage disequilibrium in several discovered loci were identified, but approximate conditional analysis with genome-wide complex trait analysis[47] confirmed that each of the six loci was consistent with just one specific signal. Of the six significant loci, four had been reported before, but two were novel. The most significant SNP in one of these new loci was rs2517959 ($P = 4.53 \times 10^{-9}$; odds ratio 1.13) on chromosome 17q12, which is in an intronic part of the *ERBB2*; this gene is expressed in the brain and encodes a member of the epidermal growth factor family of receptor tyrosine kinases that acts as a cell-surface receptor for neuregulins. The most significant SNP in the other novel locus was rs12553324 ($P = 5.87 \times 10^{-9}$; odds ratio 1.12), which is located in the intergenic region on chromosome 9p21.3, where an SNP (rs10965780) with moderate linkage disequilibrium ($r^2 = 0.6$) is related to the *ELAVL2*, which encodes a neuron-specific RNA binding protein for development of neurons. The four SNPs that had been previously reported were rs4236274 ($P = 8.49 \times 10^{-12}$; odds ratio 0.87) on chromosome 7p22.3, which was the most significant variant in the study and is located in an intron of *MAD1L1*; rs9834970 ($P = 4.83 \times 10^{-10}$; odds ratio 0.88) on chromosome 3p22.2, which was the second most significant finding in this study and is related to a region close to *TRANK1*; rs1054442 ($P = 1.20 \times 10^{-8}$; odds ratio 1.13) on chromosome 12q13.1 in *DDN*, which encodes a synapse-related cytoskeletal protein (Dendrin); and rs1487441 ($P = 2.58 \times 10^{-8}$; odds ratio 1.12) on chromosome 6q16.1, in an intergenic region between *MIR2113* and *POU3F2*. This study did not find any significant X-chromosome associations, which the authors suggested was because of the low roles of X-linked markers in the heritability of bipolar disorder. Hou et al also concluded that the highly polygenic background of bipolar disorder will require larger sample sizes to detect more involved variants.[48]

In 2017, Charney et al published a GWAS on 6447 cases of bipolar disorder and 12 639 controls from the International Cohort Collection for Bipolar Disorder. They found two regions that met the genome-wide significance threshold ($P < 5 \times 10^{-8}$): one was a two-base-pair deletion on chromosome 9 ($P = 2.48 \times 10^{-8}$; odds ratio 1.14), and the other was a locus on chromosome 10 that includes rs10884920 ($P = 1.20 \times 10^{-8}$; odds ratio 1.17). *ADD3* and *XPNPEP1* are the genes that are located in strong linkage disequilibrium with the reported locus on chromosome 10. A non-coding RNA sequence (LOC100505933) that is annotated as an anti-sense *ADD3* RNA is also located in this region. *ADD3* encodes a ubiquitously expressed cytoskeletal protein implicated in cerebral palsy that is responsible for assembling and disassembling actin filaments, and for the binding of actin to spectrin in the brain and other tissues. Aminopeptidase P is involved in the degradation and maturation of neuropeptides and mediates the proteolytic cleavage of the N-terminal amino acid in peptides with proline, such as oxytocin, corticotropin-releasing hormone, neuropeptide Y and substance P. The researchers also performed a meta-analysis of their data and the data from the meta-analysis by the PGC-BD, with a total of 13 902 cases and 19 279 controls, and found eight significant regions. The index SNPs from each independent locus and the related gene or genes in linkage disequilibrium ($R^2 > 0.25$) include rs56361249 ($P = 3.19 \times 10^{-10}$; odds ratio 1.12; *LMAN2L, FER1L5, CNNM4*) on chromosome 2, rs9834970 ($P = 1.59 \times 10^{-10}$; odds ratio 0.90; *TRANK1*) on chromosome 3, rs2302417 ($P = 2.75 \times 10^{-9}$; odds ratio 1.11; many genes) on chromosome 3, rs1203233 ($P = 4.46 \times 10^{-8}$; odds ratio 0.91; *SYNE1*) on chromosome 6, rs10994299 ($P = 1.28 \times 10^{-9}$; odds ratio 0.81; *ANK3*) on chromosome 10, rs10884920 ($P = 3.28 \times 10^{-8}$; odds ratio 1.12; *ADD3, LOC100505933, XPNPEP1*) on chromosome 10, rs4765913 ($P = 3.12 \times 10^{-9}$; odds ratio 1.13; *CACNA1C* family) on chromosome 12 and rs10459221 ($P = 2.33 \times 10^{-8}$; odds ratio 0.91; many genes) on chromosome 12. Regarding the heritability and genetic correlation of bipolar disorder subtypes, the authors reported a significant difference in the heritability: bipolar disorder type 1 SNP $h^2 = 0.35$ and bipolar disorder type 2 SNP $h^2 = 0.25$, with a genetic correlation between these subtypes of 0.78.[49]

In their 2018 publication, Ikeda et al described a GWAS on 2964 bipolar disorder cases and 61 887 controls from the Japanese population. They reported a significant locus on chromosome 11q12.2, a region known to contain regulatory genes *FADS1/2/3* for blood lipid traits and levels. Here, the top SNP was rs28456 ($P = 6.4 \times 10^{-9}$; odds ratio 1.18), which is located in *FADS2*. Subsequently, they performed a meta-analysis of the Japanese and PGC-BD data-sets, with a total of 10 445 cases and 71 137 controls. In this meta-analysis, again they found rs174576 ($P = 1.34 \times 10^{-10}$; odds ratio 1.13) as the most significant signal from the same region as identified in their first GWAS. According to the GTEx (Genotype-Tissue Expression) and The Brain eQTL Almanac data-sets, rs174576 is an expression quantitative trait loci of *FADS1* and *FADS2* in the cerebellum and temporal cortex, respectively. All quantitative trait loci SNPs for lipid traits (igh Density Lipoprotein cholesterol/Low Density Lipoprotein cholesterol/Triglyceride/Total-Cholesterol), $n-3/n-6$ polyunsaturated fatty acids) around the *FADS* region were in strong linkage disequilibrium with the index SNP (rs174576) and, because of this, the authors suggested the involvement of lipid abnormalities in pathophysiological mechanisms of bipolar disorder. This meta-analysis detected another novel association with bipolar disorder, *NFIX*, with the best signal from rs4926298 ($P = 5.8 \times 10^{-10}$; odds ratio 1.13), and also replicated three regions previously implicated in susceptibility to bipolar disorder, including *MAD1L1* (best: rs4332037; $P = 1.9 \times 10^{-9}$; odds ratio 1.17), *TRANK1* (best: rs9834970; $P = 2.1 \times 10^{-9}$; odds ratio 1.12) and *ODZ4* (best: rs12576775; $P = 3.3 \times 10^{-9}$; odds ratio 1.18). NFIX is a member of the nuclear factor one (NF1) family, and is involved in transcription and replication. Mutations in this gene can cause some syndromes associated with intellectual disability. *ODZ4* (also known as *TENM4*) encodes teneurin transmembrane protein 4, which plays a role in neuronal connectivity during development. *MAD1L1* encodes MAD1 mitotic arrest deficient-like 1, which is involved in regulating the cell cycle.[50]

In 2019, Stahl et al presented the findings from one of the largest GWASs in bipolar disorder so far, which included 20 352 cases and 31 358 controls of European descent (with a follow-up analysis of 822 variants with $P < 1 \times 10^{-4}$ in an additional sample of 9412 cases and 137 760 controls). In the combined GWAS plus follow-up analysis, the researchers identified 30 loci with genome-wide significance, 20 of which were novel. Twelve previously reported loci did not reach genome-wide significance in this GWAS meta-analysis. The strongest association signals were found for rs9834970 at the *TRANK1* locus ($P_{combined} = 5.7 \times 10^{-12}$; odds ratio 0.93) and rs2302417 at the *ITIH1* locus ($P_{combined} = 6.6 \times 10^{-11}$; odds ratio 0.93). Five other top significant signals out of the 30 loci are: rs11647445 at the *GRIN2A* locus on chromosome 16 ($P_{combined} = 1.1 \times 10^{-10}$; odds ratio 0.93), rs10744560 at the *CACNA1C* locus on chromosome 12 ($P_{combined} = 3.6 \times 10^{-10}$; odds ratio 1.076), rs201231874 at the *MRPS33* locus on chromosome 7 ($P_{combined} = 6.2 \times 10^{-10}$; odds ratio 0.92), rs61332983 at the intergenic locus (nearest gene: *PCGEM1*) on chromosome 2 ($P_{combined} = 7.9 \times 10^{-10}$; odds ratio 0.93) and rs12226877 at the *FADS2* locus on chromosome 11 ($P_{combined} = 9.9 \times 10^{-10}$; odds ratio 1.085). The identified loci include genes that encode ion channels and transporters (*CACNA1C, SCN2A* and *SLC4A1*), neurotransmitter receptors (*GRIN2A*) and synaptic components (*RIMS1* and *ANK3*). Regulation of insulin secretion and endocannabinoid signalling are two out of nine significant enriched gene-sets that resulted from pathway analysis. Stahl et al concluded that brain calcium channels and regulation of neurotransmitters play roles in the aetiology of bipolar disorder. In addition, they mentioned that discordance of results within and across GWASs could be a result of the intricate genetic architecture and phenotypic heterogeneity of bipolar disorder as a complex psychiatric disorder with a highly polygenic background.[51]

In 2021, Mullins et al reported on the largest GWAS to date, with 41 917 patients with bipolar disorder and 371 549 controls of European ancestry (57 bipolar disorder cohorts collected in Europe, North America and Australia). This GWAS meta-analysis detected 64 independent loci associated with bipolar disorder at genome-wide significance ($P < 5 \times 10^{-8}$), 33 of which were new. Among the 64 loci, the ten top associations were as follows: rs983 4970 ($P = 6.6 \times 10^{-19}$; odds ratio 1.087) located in *TRANK1* on chromosome 3, rs1487445 ($P = 1.5 \times 10^{-15}$; odds ratio 1.078) located in *POU3F2* on chromosome 6, rs11062170 ($P = 1.9 \times 10^{-15}$; odds ratio 1.081) located in *CACNA1C* on chromosome 12, rs13195402 ($P = 5.8 \times 10^{-15}$; odds ratio 1.146) located in *MHC* on chromosome 6, rs174592 ($P = 9.9 \times 10^{-14}$; odds ratio 1.074) located in *FADS2* on chromosome 11, rs113779084 ($P = 1.4 \times 10^{-13}$; odds ratio 1.079) located in *THSD7A* on chromosome 7, rs2336147 ($P = 3.6 \times 10^{-13}$; odds ratio 1.070) located in *ITIH1* on chromosome 3, rs10994415 ($P = 1.1 \times 10^{-11}$; odds ratio 1.125) located in *ANK3* on chromosome 10, rs2273738 ($P = 1.6 \times 10^{-11}$; odds ratio 1.096) located in *ADD3* on chromosome 10 and rs62489493 ($P = 2.6 \times 10^{-11}$; odds ratio 1.094) located in *miR124-1* on chromosome 8. The genes encoding targets for antipsychotics, calcium channel blockers, anti-epileptics, and anaesthetics were shown to have significant signal enrichment.[17]

Polygenicity and polygenic scores

One of the main lessons from the aforementioned GWASs is the extensive polygenicity that characterises the genetic architecture of bipolar disorder. This fact implies that the genetic factors that shape the individual vulnerability to bipolar disorder (and most human complex traits) involve hundreds to thousands of common genetic variants that are frequently found in humans, and each contribute to vulnerability with a discrete effect size.

Nowadays, the individual polygenic risk load in each individual can be estimated by using PRS. In general, PRS are calculated as a sum of an individual's phenotype-associated alleles (i.e. in the present context, alleles associated with bipolar disorder) weighted by the effect size of each variant. The effect sizes are derived from large GWASs. PRS provide a quantitative measure of the risk burden in each individual.

One of the main advantages of PRS are their usefulness as genetic tools for genetic analyses, because they summarise the global genomic risk rather than focusing on specific variants. The sample size required to evaluate the role of polygenic risk on a certain trait is substantially smaller than the one needed to quantify the genetic contribution of each genetic variant contributing to such PRS (discovery GWAS).[29] Thus, they hold great potential as a risk predictor for preventive approaches or for stratifying patients according to the likelihood that they will respond better to a given clinical intervention.[27]

PRS have shown fairly good replicability across independent samples if the discovery GWAS was well powered. In the particular case of bipolar disorder, leave-one-out analyses showed that the association of bipolar disorder-PRS with risk of bipolar disorder is consistent in each independent sample.[17] The same study showed that in European samples of patients with bipolar disorder, bipolar disorder PRS explained approximately 4.57% of phenotypic variance in bipolar disorder on the liability scale. Moreover, although they explained less phenotypic variance, bipolar disorder PRS were also significantly associated with bipolar disorder risk in people of African and East Asian ancestries.

Given the pleiotropic effect of most variants (i.e. their simultaneous effect on the expression of several traits and/or disorders), PRS can also help to unravel genetic similarities between various traits and disorders. The first large GWAS on schizophrenia published in 2009 already showed that patients with bipolar disorder have a higher genetic load for schizophrenia (schizophrenia PRS) than the general population, a finding that aligned with previous clinical and epidemiological evidence of a large overlap between both disorders.[52]

Among the ever-increasing findings on PRS–phenotype associations, several results are particularly interesting for our understanding of the genetic architecture of bipolar disorder. First, patients with bipolar disorder not only have a higher genetic risk for bipolar disorder than healthy controls, but they also have an increased genetic risk load for schizophrenia (schizophrenia PRS) and major depression (major depressive disorder PRS).[53] Second, a comparison of bipolar disorder types 1 and 2 showed that these two groups do not have a similar profile of polygenic risk, at least regarding the risk of depression.[54] Third, bipolar disorder PRS are associated not only with other psychiatric phenotypes, but also with several clinical features and potential biomarkers in different psychiatric conditions.[55] Finally, certain clinically relevant features in bipolar disorder, such as response to lithium treatment or age at onset, are not associated with bipolar disorder PRS but are associated with other psychiatric PRS.[4,56,57]

WES

WES analysis on bipolar disorder samples can reveal the contribution of rare variants to the genetic architecture of this heritable complex disorder.[58] WES is an economical approach for looking into the genetic variations that are most likely to have an effect on gene structure and function. Although it may be advantageous to have a more confined data-set for clinical applications, one of the key drawbacks of WES is that it only sequences roughly 1% of the genome. WES cannot check large portions of the genome that we do not already recognise as protein-coding regions, and so nucleotides in other regions will be ignored. As a result, for example, if a high-impact variant is in a distal regulatory element or in 5′ and 3′ untranslated regions, it will go unnoticed.[22,59] We queried PubMed with the search string 'whole exome sequencing and bipolar disorder' to identify studies that performed WES in bipolar disorder samples.

In 2013, Chen et al published a study that used a new hybrid likelihood model (burden or mutation position test) to examine WES data from 191 cases of bipolar disorder and 107 controls, and identified seven gene-sets related to bipolar disorder. The MAPK signalling pathway (defined by the KEGG Pathway Database) was the most significant gene-set and reached trend-level significance after multiple testing correction (nominal $P = 0.0065$; false discovery rate 0.095). This highly conserved pathway included 267 genes such as *DUSP6, CACNA1C* and the *phospholipase A2* family of genes.[60]

In 2016, Kataoka et al published a trio-based WES study on bipolar disorder that examined 237 samples (79 affected probands and their unaffected parents), and identified 71 *de novo* point mutations in 70 genes and one *de novo* copy number variation (a deletion located upstream of the 3q29 deletion syndrome locus with a reported association with bipolar disorder) in 79 probands with bipolar disorder. Each proband had zero to four mutations, and 42 probands had at least one mutation. The researchers reported that *EHD1, KLF4, KMT2C, MACF1, UNC13B* and *XPO4* were detected as being intolerant to a functional variation hit by *de novo* loss-of-function mutations from WES, and can be considered as bipolar disorder susceptibility gene candidates. Gene ontology enrichment analysis with the Database for Annotation, Visualization and Integrated Discovery (version 6.7) resulted in no significant gene ontology term after correction for multiple testing.[61]

In 2017, Lescai et al presented the results of WES of an isolated population (Faroese) that included 28 patients with bipolar disorder and 214 controls, and then tried to replicate the results in a British sample (2025 cases and 1358 controls). In the discovery step, they reported 17 exome-wide significant variants in 16 genes in the single-variant analysis and three significant genes in the gene-based statistics. *NOS1* (rs79487279 as a mis-sense variant) and *NCL* (gene-based test) were present in the British replication samples, but *PITPNM2* (12:123489064_C as a mis-sense variant) and *PIK3C2A* were absent. Both *PITPNM2* and *PIK3C2A* are components of the phosphatidylinositol pathway, which has been linked to psychiatric disorders. The only statistically significant variable after Bonferroni correction ($P = 0.002$, $P_{corrected} = 0.032$) was rs79487279 in *NOS1*, which encodes neuronal nitric oxide synthase. This enzyme catalyses the production of nitric oxide, which, as a neurotransmitter, has interactions with the dopaminergic and serotonergic systems.[62]

Also in 2017, Monson et al published the findings of WES for approximately 19 000 genes of 1018 patients with bipolar disorder (387 people with a history of suicide attempt and 631 people without such a history), but they did not find any individual suicide-associated significant variant or gene at study-wide significance.[22]

In a study published in 2018, Cruceanu et al sequenced whole exomes of 186 patients with bipolar disorder or related individuals from 40 multiplex families and approximately 800 controls. The authors postulated the causality of deleterious rare variants in

disease transmission across generations. By following a variant filtration approach and focusing on potential pathogenic variants, they identified rare variants in many genes and brain-related pathways. *GPCR* family genes, which are involved in regulating brain functions and are also considered as drug targets, were enriched in the identified rare variants among most of the studied families (in 38 out 40 families, with a mean of five variants per family). *CRHR2, HTR1B, GRM1, GRM4* and *DRD5* are some of the *GPCR* genes with flagged mutations. The identified non-sense mutation (R384X: arginine to stop codon) in *CRHR2* (which plays a role in stress response) removes the last 28 amino acids, resulting in impaired complex formation, with regulators leading to increased signal transduction and effects on several downstream cellular functions. In *GRM1*, which is involved in synaptic plasticity-related learning and memory, a mis-sense mutation (exon6:c.T1524G, p. D508E) occurred that had no effect on receptor function, but may have had an effect on the kinetics of GPCR activation.[63]

In 2018, Salvoro et al published WES results on 17 patients (including six cases of bipolar disorder) from a familial sample (including 90 cases of schizophrenia, 50 cases of bipolar disorder and 21 cases of schizophrenia) in combination with identity-by-descent (IBD) mapping data to identify risk loci in a related closed population. They identified a total of 13 330 rare (minor allele frequency [MAF] <0.01) and 3903 novel variants. After mapping and phasing into the total list (IBD$_{tot}$) haplotypes and checking for sharing in whole sample, 1393 variants (1052 rare and 341 novel) were identified as IBD$_{risk}$ haplotypes. Pathway and gene-set enrichment analysis indicated the involvement of extracellular matrix interactions, axon guidance and synaptic functions (development of neuronal connectivity) in the aetiology of and susceptibility to bipolar disorder.[64]

In 2018, Zhang et al reported the results of WES and linkage analysis of six related patients with bipolar disorder. They identified six variants in five candidate genes (*GRID1, CDHR1, DYDC2, GHITM* and *MINPP1*) with potential effects on the onset and development of bipolar disorder. Among these six variants, rs2306265 (p. V529I) in *GRID1* and rs45584033 (p.P812S) in *CDHR1* were flagged as damaging (by Sorting Intolerant From Tolerant [SIFT], a program that predicts effects of amino acid changes on protein function) or probably damaging (by Polyphen-2, or Polymorphism Phenotyping version 2, a program for predicting the possible impact of amino acid substitutions on stability and function of proteins). *GRID1* maps to chromosome 10 and encodes a subunit of glutamate receptor channels, which are involved in synaptic plasticity and transmission in the central nervous system. *CDHR1* is located also on chromosome 10 and encodes a photoreceptor-specific cadherin protein that is involved in outer segment disc morphogenesis.[23]

In 2018, Toma et al published findings of WES of 117 members (72 with bipolar disorder and 45 unaffected by bipolar disorder) of 15 European-origin multiplex and extended families from Australia. In addition, they performed linkage analysis and haplotype reconstruction on WES-derived data, and copy number variation analysis to enrich, validate and improve the outputs as much as possible. They identified 532 potentially aetiologic and 541 likely neutral variants. Enrichment analysis results showed that the potentially aetiologic variants were overrepresented in postsynaptic density genes, which are extremely sequence conserved and include scaffolding proteins. The likely neutral variants were not enriched for any gene-sets. The authors discovered that the number of brain-expressed likely gene-disruptive variants was negatively associated with the age at onset, but they did not observe any differences in the total number of *de novo* variants between affected and unaffected offspring (32 sequenced parent–child trios: 22 parents with bipolar disorder and ten unaffected offspring). Combining rare variants of all classes with linkage evidence for each family resulted in a protein-truncating mutation in the X-linked gene *IRS4*, as a plausible new candidate for bipolar disorder. *PCDHA* gene cluster, which is expressed in serotonergic neurons, and *PRODH* which encodes a nuclear-encoded mitochondrial protein, both of which act as neuromodulators of dopaminergic synaptic transmission, are two of the genes with several rare recurrent variations identified in this study. Toma et al concluded that cumulative rare pathogenic variants in several postsynaptic density genes could disrupt synaptic homeostasis, with consequences in a variety of psychiatric disorders, such as bipolar disorder.[58]

As described in their 2018 publication, Ganna et al investigated the effect of rare protein-truncating variants (PTVs) on 13 quantitative traits and ten diseases, including bipolar disorder, by analysing WES data from 100 296 individuals (1201 with bipolar disorder) from a combination of cohort and case–control disease studies. The authors focused on PTVs that occur in a set of 3172 PTV-intolerant genes, and showed an association with early-onset neurodevelopmental and psychiatric disorders (for bipolar disorder: $P = 8 \times 10^{-4}$; odds ratio 1.2). They reported that genes intolerant to PTVs are well-suited to capture the effect of rare to ultra-rare PTVs on the cognitive, behavioural and developmental spectra.[65]

In 2018, Maaser et al reported the results of WES of 15 cases of bipolar disorder from two large Cuban families. Four people from family 1 (pedigree 1) and 11 from family 2 (pedigrees 2–4) were selected. The analysis detected 17 rare and non-synonymous frameshift-causing variants (14 mis-sense and 3 small insertions and deletions) in 17 genes that were shared by all patients with bipolar disorder within the related pedigree. Of 17 recognised genes harbouring these variants, according to the GTEx database, 13 genes (*SERPING1, TMEM220, MYH7, THYN1, ATR, CSNK1G3, ZNF433, OLFML2B, CAPN2, RCCD1, DNAH7, SELENOO* and *COL3A1)* are expressed in the brain. The variant rs141075266 in *SERPING1* (p.L349F) was considered as the most interesting variant in this study. *SERPING1*, which was previously reported as a risk gene for schizophrenia with genome-wide significance, encodes the serpin peptidase inhibitor clade G member 1, which plays an important role in the regulation of complement activation.[66]

Also in 2018, Husson et al published results of WES of 92 patients with bipolar 1 disorder (positive family history, 63% early onset) and 1051 controls of French ancestry. The authors were in search of rare predicted damaging variants (MAF < 0.01), but no gene in this study met the exome-wide *P*-value threshold (2.5×10^{-6}) or the false discovery rate threshold (10%); however, they reported 13 genes (as top hits for rare variants) with $P < 10^{-3}$ and a high odds ratio (range: 3–23.7) that are expressed in the brain. Among these genes, *PLCXD3, ARHGAP9* and *TCF7L1* have been reported to have potential roles in the therapeutic effect of lithium treatment.[67]

In 2019, Han et al described their WES study in 53 patients with bipolar disorder and 82 healthy controls. In the study, they also applied structural MRI data to find variants associated with cortical grey matter thickness and the integrity of white matter. They detected 122 bipolar disorder-related genes (such as *KMT2C, AHNAK* and *CDH23*) and 27 recurrent copy number alteration regions. In their paper, they mentioned the probable role of rs4639425 in *KMT2C*, which plays a role in chromatin remodelling via regulation of histone H3 lysine 4 (H3K4) methylation, in microstructural alterations of white matter tracts of the brain.[68]

In 2019, Ganesh et al presented the results of WES of 32 patients with severe mental illness (26 with bipolar disorder, four with schizophrenia, one with schizophrenia-like psychosis and one with schizoaffective disorder) from eight multiplex families and 33 Indian healthy controls. This study detected 42 rare, non-

synonymous deleterious variants in 42 genes (approximately five per pedigree) in the form of a familiar private mutational profile. Twelve of the variants were novel, 20 were previously reported to be involved in different neuropsychiatric syndromes and nine had been reported in Mendelian syndromes with early-onset neurodevelopmental features. Two of the newly identified variants were c. A565G (p.S189G) and c.A4496C (p.Q1499P; segregating with bipolar disorder in a pedigree) in GRM1 and NIPBL, respectively. GRM1, which is involved in spinocerebellar ataxia, encodes metabotropic glutamate receptor 1, and pathogenic mutation in this gene causes early age at disease onset and associated intellectual disability. NIPBL encodes cohesin loading factor, which plays a role in cortical neuronal migration and causes Cornelia de Lange syndrome. MDN1, which plays a role in ribosomal large subunit assembly, and WDFY4, which acts as an integral component of membrane, are two reported bipolar disorder-related genes that were identified in this study as harbouring two detected rare variants (rs148868949 and rs571808731, respectively).[69]

Also in 2019, Szatkiewicz et al reported on WES of 25 patients with bipolar disorder (12 with bipolar disorder type 1 and 13 with bipolar disorder type 2) from a dense multiplex Swedish pedigree. First, they identified 129 651 variants (1.4% of which were novel) and, after some annotation filtering such as selecting variants with unknown or highest allele frequency; predicted as high impact; or those located in brain-expressed genes, 102 variants remained; however, they did not detect any shared strong predictions for more than half of the cases.[70]

In 2020, Anjanappa et al published the results of WES of 11 patients with bipolar disorder from a four-generation Indian family and, based on the linkage investigations, focused on two genomic intervals on chromosome 1 and 6. They reported two rare variants: a rare, non-synonymous, heterozygous variant that was present in seven patients (rs776739426; p.D821Y) in KANK4 on chromosome 1p31.1–13.2, and a variant that was present in ten patients (rs910586667; in the 5′ untranslated region) in CAP2 on chromosome 6p22.3–22.2. KANK4 encodes a protein with KN motif that has potential nuclear localisation and export signal. CAP2 encodes one of the cyclase-associated proteins, which play a role in the regulation of actin dynamics. The authors suggested that KANK4 and CAP2, which are also expressed in the brain, alone or together are involved in bipolar disorder in the investigated family.[71]

In their 2020 publication, Forstner et al described their findings from WES of 81 patients with bipolar disorder from 23 Spanish and four German multigenerational, multiply affected families (three members from each family). The analysis considered only variants shared by all three investigated family members. The researchers identified 378 rare, non-synonymous and potentially functional variants in 368 genes. Eight of these genes (ADGB, DCAF5, NCKAP5, PKHD1L1, AOAH, CAND2, DIDO1 and RGS12) were shared by at least two independent families. All of the genes except for PKHD1L1 and ADGB are expressed in the brain, and five of them (ADGB, DCAF5, NCKAP5, PKHD1L1 and RGS12) had complete or nearly full penetrance in the extended segregation analysis. NCKAP5 and RGS12 had already been suggested by candidate gene studies as being associated with bipolar disorder. Although the role of NCKAP5, which is located on chromosome 2q21.2, is unknown, RGS12 is a member of the GTPase activating protein family mapping to chromosome 4p16.3, and is involved in regulation of GPCR signalling. The 368 genes identified in this study include 19 overlapping genes, e.g. ANK3, MACF1, MYO10 and SYNE1 (cytoskeleton-associated genes), with the previously reported genes from bipolar disorder GWASs and sequencing studies, highlighting the potential role of cytoskeletal pathways in the aetiology of bipolar disorder.[72]

Also in 2020, Jacobs et al presented the results of WES of five family members (two monozygotic twins with bipolar disorder and an unaffected brother and parents) that identified seven variants in six genes mostly involved in GABA/glutamatergic pathways. These six genes and their identified variants were as follows: NF1 on chromosome 17q11.2, a non-synonymous single nucleotide variation (nsSNV; c.3200A > T; p.Asp1067-Val) and a frameshift mutation owing to a deletion (c.3198CAGAG/Del; p.Asp1067fs); GGA3 on chromosome 17q25.1, a stop-loss mutation (c.672T > C); DISC1 on chromosome 1q42.2, an nsSNV (c.1694G > T; p.Cys565Phe); NMUR2 on chromosome 5q33.1, an nsSNV (c.617T > A; p.Val206Asp); HIP1R on chromosome 12q24.31, an nsSNV (c.1495G > A; Val499Met); and BOD1 on chromosome 5q35.2, an nsSNV (c.6C > T).[73]

Also in 2020, Courtois et al published results of WES of 34 patients with bipolar disorder from eight multiplex families of French ancestry that included at least two people with bipolar disorder. After filtration of variants, they detected a total of 34 rare deleterious mutations in 34 genes (29 of which are expressed in neurons) that are involved in neuronal development, morphogenesis and regulation of neuronal gene expression. Among these genes, SMARCC2 (involved in transcription), HELLS (involved in genome-wide methylation), UPF2 (involved in proper development of synapse architecture and synaptic vesicle efficacy) and WDR37 (unknown function) showed a higher mutation rate in patients with bipolar disorder than in the Exome Aggregation Consortium (ExAC) population after correction for multiple testing. An analysis of the clinical profile of patients with these four genes revealed that mutated SMARCC2 caused a trend to more manic episodes, whereas mutated HELLS increased the chances of mixed episodes.[74]

In 2020, Engelbrecht et al described the findings of WES of eight Afrikaner family members (five members with bipolar disorder and three unaffected members) and identified two rare mis-sense variants in FAM71B and SLC26A9 that were shared by all five affected members, but were not present in any of the unaffected members. The authors mentioned that these variants were located in genes involved in signal transduction and synaptic transmission.[75]

In 2021, Umehara et al published the findings of WES of two patients with bipolar disorder from a Japanese family with a high frequency of bipolar disorder and recurrent major depressive disorder. In these two cases and four additional patients with bipolar disorder, two individuals of unknown status and ten healthy controls from this family, genome-wide linkage analysis were conducted before WES to restrict the linkage regions. Then, after performing different filtering steps on detected variants, they finally observed a rare heterozygous mutation in exon 31 of DOCK5 (c.3170A > G, p.E1057G). DOCK5 is involved in producing an atypical Rho guanine nucleotide exchange factor for Rac and/or Cdc42 GTPases. It is also an essential factor in transmitting glycogen synthase kinase 3 beta signals in mast cells in allergic reactions, and could play a role in microglial activation.[76]

Also in 2021, Toma et al described the results of WES and linkage analysis of 117 members (50 patients with bipolar disorder) from 15 Australian families with a high density of mood disorders. After quality control and family gene-based segregation test, three brain-expressed genes containing rare mis-sense variants were identified on chromosome 10: ANK3, NRBF2 and PCDH15. For bipolar disorder, 10q11.23-q21.3 is an important locus in that it explains part of the related genetic risk. By using combined annotation dependent depletion and deleterious annotation of genetic variants using neural networks scoring tools to assay the protein-disrupting or -altering pathogenicity of detected rare variants in these genes, the researchers revealed the high rankings of these variants in the human genome (top 1 to 0.01% of the deleterious variants). PCDH15 is involved in neuronal differentiation and synapse

formation, and *NRBF2* plays a role in autophagy that is essential for neuronal survival and function.[77]

Looking at the detected variants and related genes and loci throughout WES in different studies, some well-known candidate pathways and genes, such as MAPK signalling, phosphatidylinositol, cytoskeletal and GABA/glutamatergic pathways, as well as the GPCR family and some other genes, are identified. Most of these pathways and genes are responsible for several important neuronal-related properties such as development, morphogenesis, differentiation, survival, connectivity and axon guidance; regulation of gene expression; extracellular matrix interactions; signal transduction; synapse formation, plasticity and homeostasis; and synaptic transmission. Cumulative pathogenic variations in these genes, which can disrupt the important neuronal cellular and molecular processes, may put individuals at risk for mental disorders like bipolar disorder.[23,58,60,62–64,72–75]

WGS

By scanning the complete genome (coding and non-coding regions), WGS has the potential to discover every allelic variant, and especially rare variants. Additional risk loci for neuropsychiatric disorders may be discovered through further analysis of the entire genome and within regulatory regions near candidate genes. However, because sample size is one of the most critical criterion in psychiatric genetic investigations, WGS remains expensive and technically challenging for very large sample sizes.[22,24] Despite such economical constrains, a very recent UK Biobank release already includes WGS data of >150 000 participants.[78] To obtain the results of WGS studies on bipolar disorder, we queried PubMed with the search string 'whole-genome sequencing and bipolar disorder'.

As described in a 2014 publication, Fiorentino et al sequenced the whole genome of 99 patients with bipolar disorder. After filtering of the detected variants, ten SNPs remained for *CACNA1C* and three SNPs for *ANK3*. Then, the identified filtered SNPs were genotyped in 1510 cases with bipolar disorder and 1095 controls. The two variants ultimately identified in this study as being associated with bipolar disorder were rs79398153 in intron 3 of *CACNA1C* ($P = 0.015$), which is located in an Encyclopedia of DNA Elements (ENCODE)-defined region for H3K4me1, H3K27ac and H3Kme3; and rs139972937 in exon 34 of *ANK3* ($P = 0.042$), which causes a non-conservative amino acid change from asparagine to serine (N2643S). Because these two novel variants are not in linkage disequilibrium with the previous bipolar disorder-associated variants in *ANK3* and *CACNA1C*, and are independent, the authors pointed out that if these variants can alter the expression and function of *ANK3* and *CACNA1C*, they could explain the 'missing heritability'.[79]

Also in 2014, Georgi et al published a study that combined the data from WGS of 50 members (23 affected individuals with bipolar disorder type 1 or 2) with high-density SNP-array genotypes and microsatellite genotypes of 388 members of a large Old Order Amish pedigree (a genetic isolate of European ancestry currently residing in North America). WGS detected a total of 11.1 million variants (within the 97% called genome fraction), and approximately 3.79 million SNPs and insertions and deletions in each individual, 20 000 of which were located within exons. The group searched for exonic mis-sense variants that were common in the 23 patients with bipolar disorder and rare variants (MAF < 2%) in 1000 Genomes Project data, the exome variant server database, and a set of 54 HapMap WGS. The authors found none of the resulting 514 damaging/deleterious variants (predicted by both Polyphen-2 and SIFT) were both present in all bipolar disorder samples and

absent or rare in the 1000 Genomes Project data. Many of the detected non-synonymous exonic, putative damaging variants are common in European populations (in *SNTG2, TTC15, POLN* and *EVC*), and some of them are rare or novel (in *LAMB1, MKL2, ZNF732, ZAN* and *ANKS4B*). Also, *ANK3, ODZ2, ODZ4* and *ITIH3* were identified to have rare and putative damaging variants. Georgi et al concluded that there is no evidence from their study for implication of specific set of risk loci or common pathways for bipolar disorder.[80]

In 2016, Ross et al described their WGS of three patients with bipolar disorder type 1 from a multiplex pedigree of European ancestry with seven members with bipolar disorder type 1. The authors reported 115 heterozygous rare variants in 113 genes shared by all three patients. Among all of these genes, the authors focused more on *CACNA1D* (ranked 24th genome-wide under the dominant mode of inheritance). In this gene, one of the detected variants is G-to-C nucleotide transversion, which causes substitution of proline for alanine at position 1751 (A1751P). Co-segregation of this variant with another four people with bipolar disorder type 1 in the pedigree was confirmed by Sanger sequencing. Also, the researchers detected a transition mutation in this gene in one patient that was a conversion of C to T, causing substitution of tryptophan for arginine (R1771W). Because of the role of *CACNA1D* in calcium signalling, the authors concluded that changes in this sequence may indirectly lead to bipolar disorder.[81]

Also in 2016, Fiorentino et al published their WGS data on 99 patients with bipolar disorder type 1 in whom they searched for variations in potential miR-708 binding sites by using Targetscan 6.2 to find bipolar disorder-associated variants. They reported four non-reference (GRCh37/hg19) and non-aetiological allelic variants in *APOC3* (rs5128), *GALNT13* (rs707082), *ST6GALNAC4* (rs1043026) and *SLC22A23* (rs5873874). *APOC3* and *SLC22A23* have been reported to be involved in response to antipsychotic treatments.[82]

In their 2017 publication, Bouwkamp et al presented WGS findings in a patient from a family with a high incidence of affective psychosis (bipolar disorder type 1 or schizoaffective disorder, bipolar type). By using WGS, they discovered a breakpoint of a balanced translocation (detected by karyotyping and fluorescence *in situ* hybridisation) between chromosomes 6 and 15. The breakpoint is located in intron 11 of *PNLDC1* on the long arms of chromosome 6 and intron 1 of *BCL2L10* on chromosome 15. *PNLDC1* encodes a protein containing an RNaseH-like domain and RNAse_CAF1 domain, and *BCL2L10* encodes an intracellular membrane-associated BCL2 family protein that is localised in mitochondria of brain cells.[83]

Also in 2017, Steinberg et al reported the results of WGS of one patient with bipolar disorder and six patients with schizophrenia and schizoaffective disorder in an Icelandic family with enriched psychosis burden. c.2377G > T ($P = 2.2 \times 10^{-4}$), which is located in the exon of *RBM12* on chromosome 20q11, was found in all of the patients. The protein that results from this non-sense mutation is a truncated version without RNA-recognition motif. By using PRS, the authors found that patients with psychosis with c.2377G > T had higher scores for bipolar disorder than non-psychotic carriers.[84]

Also in 2017, Sharp et al published an imputation analysis on a WGS sample of 99 patients with bipolar affective disorder that detected eight significant bipolar disorder-associated SNPs in the *CPG2* region of *SYNE1* on chromosome 6q25.2; *CPG2* is a brain-specific splice variant of *SYNE1*.[85]

In 2019, Szatkiewicz et al described the results of WGS of ten patients with bipolar disorder from a large pedigree in Northern Sweden, but they did not identify any high-impact variants.[70]

In 2021, Goes et al presented their findings from WGS of 97 trios of Ashkenazi Jewish descent without first-degree family

history of bipolar disorder, and reported 6882 *de novo* variants, including 107 variants within protein-coding genes. Twenty loss-of-function and 77 mis-sense damaging *de novo* variants were identified by combining the findings on exon-related variations with the findings of a previously published study (exome analysis) on trios. The detected variants were significantly enriched in genes related to postsynaptic density. The phosphoinositides pathway, which is hypothesised to participate in the mechanism of action of lithium, is one of the pathways for the identified variants and involved genes.[86]

In 2020, Nakajima et al published the results of WGS of a proband from a family in which bipolar disorder co-segregated with a Mendelian kidney disorder. WGS detected a damaging mis-sense mutation in *NTRK1* on chromosome 1q22 that encodes Tropomyosin-related kinase A (TrkA). This mutation causes a substitution of glutamate to lysine at the 492 residue of TrkA protein (E492 K). TrKA participates in cholinergic neurotransmission, and this mutation might potentially interfere in its function, leading to an increased risk for bipolar disorder.[87]

Also in 2020, Truvé et al reported on the findings of WGS of six unrelated patients with bipolar disorder and three healthy controls that extracted a list of 105 candidate SNPs for genotyping. After genotyping of these 105 SNPs in 47 unrelated cases and 44 controls and by focusing on the rarest genetic variants, *NEK3*, *ANK3* and *TUBB1* were introduced as the genes with the most associated variants. *NEK3*, which harboured the most associated SNPs in this study, encodes a protein kinase, which plays a role in neuronal morphogenesis and microtubule polarity. *TUBB1* encodes beta-1 chain of tubulin as the major constituent of the microtubules. Normal microtubules play roles in axon structure and transport, and respond to cellular and environmental stress; the mutant and unregulated microtubules cannot be effective players in stressful conditions, and may result in the triggering of mood episodes in patients with bipolar disorder.[20]

Finally, in 2020, Sul et al published high-coverage WGS of 449 individuals from 26 Colombian and Costa Rican determined pedigrees for bipolar disorder type 1, to detect associated rare and common variants, but they did not identified rare variants co-segregating in these pedigrees.[88]

Like WES, the majority of sample sizes in WGS studies are also small (<100), and this, coupled with the expenses, might be a disadvantage of such investigations. In the WGS researches in our survey, some rare and very rare novel variants in different genes have been reported. These variants are more linked with or play roles in the phosphoinositides pathway, calcium signalling, cholinergic neurotransmission, mitochondrial-related events, response to antipsychotic treatments, neuronal morphogenesis and microtubule polarity.[20,81–83,86,87] Furthermore, several WGS studies found no high-impact rare variants or revealed no evidence of implication of specific set of risk loci or common pathways that may be used to explain the 'missing heritability' of bipolar disorder.[70,80,88]

MRI neuroimaging

MRI is a widely used tool for studying human brain structure and function, with the goal of establishing links, or endophenotypes, between behavioural symptoms in mental disorders and underlying pathophysiology and genomic risk. Over the past several decades, researchers have found many associations between MRI-derived brain markers and bipolar disorder, with a growing body of evidence showing that bipolar disorder is associated with alterations in the neuroanatomical structures of emotion and reward processing.[89]

MRI is a versatile tool that is capable of measuring different features of brain structure and function. Studies of grey matter have reported lower volumes in subcortical structures, such as the thalamus, amygdala and hippocampus,[90,91] and abnormalities in temporal, frontal and insular cortices.[92–94] Studies with specialised diffusion-weighted MRI sequences, which are used to quantify white matter integrity, have found disrupted widespread alterations in bipolar disorder, including in fronto-limbic connections.[89] Furthermore, functional MRI, which takes advantage of the magnetic properties of oxygenated/deoxygenated blood flow to infer neural activity, has shown alterations in frontal, temporal and striatal networks that regulate emotion and reward regulation in bipolar disorder.[89] However, variable and sometimes conflicting results are not uncommon across neuroimaging studies of bipolar disorder. MRI studies can be expensive to perform, and many include <100 participants. In addition, patient heterogeneity (e.g. age at onset, duration/severity of illness and treatment regimen), coupled with the relatively subtle effect of bipolar disorder on MRI-derived brain measures, represents a significant challenge to improving the replication and reproducibility of neuroimaging findings in bipolar disorder.

One effort to improve the replication, reproducibility and generalisability of neuroimaging makers across bipolar disorder (and many other brain disorders) is the Enhancing Neuro Imaging Genetics through Meta Analysis Consortium (ENIGMA). Founded in 2009 to support GWASs on MRI-derived brain measures, ENIGMA now consists of >2000 scientists who collaborate across 50 working groups to study a range of brain illnesses and genetic risk factors by implementing standardising processing and analysis methods for large-scale, multisite neuroimaging studies (http://enigma.ini.usc.edu/).[95]

The ENIGMA Bipolar Disorder Working Group (ENIGMA-BD; http://enigma.ini.usc.edu/ongoing/enigma-bipolar-working-group) is a growing collaboration of 150 investigators across 55 international research institutions who are working to pool data and resources to enable the largest neuroimaging studies of bipolar disorder ever conducted to be performed.[28] By applying standardised processing and quality control protocols to existing, independently collected MRI study samples from around the world, ENIGMA-BD aims to build more ecologically valid samples and revive old data-sets to identify new, more generalisable bipolar disorder biomarkers. The standardisation of brain measures across large, multisite MRI samples boosts the power to detect subtle bipolar disorder-specific brain alterations and makes transdiagnostic studies possible across the ENIGMA psychiatric working group samples (e.g. major depression, schizophrenia and obsessive–compulsive disorder). ENIGMA-BD has published the largest studies of cortical and subcortical brain structure,[96,97] white matter integrity,[98] hippocampal subfields,[99] bipolar disorder familial risk[100,101] and machine learning for bipolar disorder classification.[102]

The ENIGMA-BD studies show a replicable pattern of lower cortical thickness, altered white matter integrity and smaller subcortical volumes in bipolar disorder. Common medications such as lithium show normalising associations with brain structure, whereas anticonvulsants appear to show the opposite effect and, in some cases, are associated with thinner or smaller brain structures. Importantly, and in line with prior reports, bipolar disorder appears to be associated with small-to-moderate effect sizes on MRI-derived brain markers.[28]

Recent work has linked ENIGMA-standardised brain MRI phenotypes to underlying neurobiological mechanisms and genetic factors. One study showed that case–control brain differences across ENIGMA psychiatric working groups (including ENIGMA-BD) partially mirrored known genetic correlations between disorders.[103] Another study showed that a convergent

pattern of underlying cell-specific gene expression could be used to partially explain patterns of common cortical thickness alterations across psychiatric illnesses.[104] ENIGMA-BD and the PGC-BD are working together to study how PRS derived from large-scale Psychiatric Genomics Consortium GWASs modulate brain structure and function in bipolar disorder. Finally, ENIGMA GWASs are mapping genetic correlations between variants that modulate brain structure and those that modulate psychiatric and cognitive traits, and results show that MRI-derived brain measures represent an endophenotype on the causal pathway between the genome and cognition.[105]

Replicated gene candidates

Among the various GWASs and WES and WGS studies that were performed to understand the genetic architecture of bipolar disorder, five genes – *ANK3, CACNA1C, SYNE1, ODZ4* and *TRANK1* – have been remarkably replicated in a significant number of these studies in different populations (Table 1), and can most likely be considered as reliable bipolar disorder-related gene candidates.

One of the most important identified genes is *ANK3*, which belongs to the ankyrin repeat domain containing (*ANKRD*) gene group. Ankyrin 3, or ankyrin G, is a protein-coding gene located on chromosome 10q21.2 that encodes ankyrin G protein. Results from the GTEx data-set (www.gtexportal.org) indicate that, in the brain, ankyrin G protein is expressed more in the cerebellar hemisphere and cerebellum, is positioned mainly at the nodes of Ranvier and the axon initial segment, and plays roles in stabilising sodium channels and potential propagation. Some *ANK3* risk variants have major effects on sustained attention and cognitive function, and polymorphisms in this gene were reported to be linked to an increased risk of mental illnesses, including bipolar disorder.[106] Gene ontology annotations (www.genecards.org) related to this gene include 'structural constituent of cytoskeleton' and 'protein binding, bridging'.[107]

Another outstanding detected gene is *CACNA1C*, which belongs to the calcium voltage-gated channel alpha-1 subunit gene group. *CACNA1C* is located on chromosome 12p13.33, encodes a protein that has an important role in regulating calcium influx in cells and is essential for normal brain development and plasticity. Gene ontology annotations related to this gene, which is a well-known candidate risk gene for a variety of neuropsychiatric disorders, including bipolar disorder, include 'enzyme binding' and 'ion channel activity'.[107,108]

SYNE1 is another replicated gene in bipolar disorder genetic research and belongs to the spectrin repeat containing nuclear envelope family gene group. The GTEx data-set indicates that, in the brain, *SYNE1*, which is located on chromosome 6q25.2, is expressed more in the cerebellar hemisphere and cerebellum. Its related gene ontology annotation is 'nucleotide binding'.[107] Human *SYNE1* has 145 exons and encodes multiple proteins, such as CPG2, a brain-specific protein found mainly in excitatory postsynaptic regions, where it influences synaptic plasticity and function.[109] A significant decrease in the expression level of this gene in the post-mortem brains of patients with bipolar disorder has been reported.[110]

ODZ4 (*TENM4*) is another important identified gene related to bipolar disorder and belongs to the tenascin family (teneurin subfamily). *ODZ4*, which is located on chromosome 11q14.1, encodes a transmembrane protein that is expressed predominantly in neurons, and is involved in neural development and regulating the establishment of proper connectivity within the nervous system.[111] According to the GTEx data-set, in the brain it is expressed more in the frontal cortex, nucleus accumbens, cortex

and hypothalamus; its related gene ontology annotation is 'protein homodimerization activity'.[107]

TRANK1 (*LBA1*), a replicated gene in GWASs on bipolar disorder, is located on chromosome 3p22.2. According to the GTEx data-set, it is more expressed in the cerebellum and cerebellar hemisphere. Valproic acid, a mood stabiliser, was found to have a significant effect on the expression of *TRANK1*, a finding that introduced it as a potential target for bipolar disorder treatment. Gene ontology annotations related to *TRANK1* include 'hydrolase activity'. Some studies reported that *TRANK1* dysregulation has the potential to disrupt the expression of genes involved in neuronal development and differentiation, and in synaptic plasticity.[112]

Future research perspective

In the previous decade, a large number of genetic studies on patients with bipolar disorder of various diverse ancestries have been conducted around the world, and the rate at which new discoveries are being made is rapidly growing. Most of the GWASs of bipolar disorder are still primarily in European ancestry samples, and we need to strive for more diverse cohorts to improve our generalisability in non-European cohorts. More than 78% of the reported GWAS participants are European, indicating a dearth of internationally representative samples. As a result, most PRS perform significantly worse in other groups, particularly in people of African heritage. To fully realise the global potential of genetics research and precision medicine, it is becoming more vital to invest in policies and processes that promote the involvement of a wide range of study participants. The use of cross-ancestry meta-analysis findings and sample diversity helps us to get a better knowledge of genetic architecture and the relevance of genetic research to all populations, as well as improving overall prediction.[113] Indeed, technological advancements coupled with consortium efforts that combine data and resources from many research centres have led to the development of more effective analytical methodologies for prioritising major phenotype-associated variants, and investigating new molecular pathways implicated in this complex disorder. On the other hand, defining better strategies for considering molecular and environmental bases through gene×environment interactions could lead to better understanding of brain health and disorders, in particular psychiatric disorders.

Increasing the sample size of GWASs over the past decade has helped to shed more light on the genetics of bipolar disorder. However, because of the high level of phenotypic and genotypic heterogeneity of this disorder, even the most recent GWAS, which was the largest one to date and included approximately 42 000 bipolar disorder cases and identified 64 associated genomic loci, found fewer loci with genome-wide significance than a similar study on schizophrenia, and still the proportion of phenotypic variance explained by these discovered variants is low (approximately 4.7%); this result shows the need for more studies with large sample sizes able to obtain a better comprehensive picture of the genetic architecture of bipolar disorder.[17] Likewise, well-powered GWASs in bipolar disorder can improve bipolar disorder PRS as a reliable genetic risk-predictor tool for preventive and prognostic approaches through providing well-replicated results across independent samples.

WES and WGS can contribute more substantially to the decoding of the major genetic foundations of neuropsychiatric disorders, including other complex genetic disorders, through improvement of technology and analytical and quality control methods.

Although neuroimaging research faces some challenges, such as patient heterogeneity and the relatively subtle effect of bipolar disorder on MRI-derived brain measures, integration of neuroimaging

Table 1 Replicated gene candidates for bipolar disorder identified through various genomic approaches

Gene	GWASs		WES studies		WGS studies	
	Author (year)	Sample size	Author (year)	Sample size	Author (year)	Sample size
ANK3	Ferreira et al (2008)[35]	4387 cases 6209 controls	Forstner et al (2020)[72]	81 cases	Fiorentino et al (2014)[79]	99 cases
	Psychiatric GWAS Consortium Bipolar Disorder Working Group (2011)[a40]	7481 cases 9250 controls				
	Chen et al (2013)[14]	7773 cases 9883 controls			Georgi et al (2014)[80]	50 family members, including 23 cases
	Mühleisen et al (2014)[44]	9747 cases 14 278 controls	Toma et al (2021)[77]	117 members from 15 families, including 50 cases		
	Charney et al (2017)[49]	13 902 cases 19 279 controls				
	Stahl et al (2019)[51]	20 352 cases 31 358 controls			Truvé et al (2020)[20]	6 cases, 3 controls
	Mullins et al (2021)[17]	41 917 cases 371 549 controls				
CACNA1C	Ferreira et al (2008)[35]	4387 cases 6209 controls	Chen et al (2013)[60]	191 cases 107 controls	Fiorentino et al (2014)[79]	99 cases
	Psychiatric GWAS Consortium Bipolar Disorder Working Group (2011)[b40]	11 974 cases 51 792 controls				
	Green et al (2013)[43]	8699 cases 12 163 controls				
	Charney et al (2017)[49]	13 902 cases 19 279 controls				
	Stahl et al (2019)[51]	20 352 cases 31 358 controls				
	Mullins et al (2021)[17]	41 917 cases 371 549 controls				
SYNE1	Green et al (2013)[42]	1527 cases, 1579 controls	Forstner et al (2020)[72]	81 cases	Sharp et al (2017)[85]	99 cases
	Psychiatric GWAS Consortium Bipolar Disorder Working Group (2011)[a40]	7481 cases, 9250 controls				
	Charney et al (2017)[49]	13 902 cases 19 279 controls				
	Mullins et al (2021)[17]	41 917 cases 371 549 controls				
ODZ4	Psychiatric GWAS Consortium Bipolar Disorder Working Group (2011)[b40]	11 974 cases 51 792 controls			Georgi et al (2014)[80]	50 family members, including 23 cases
	Green et al (2013)[43]	8699 cases 12 163 controls				
	Mühleisen et al (2014)[44]	9747 cases 14 278 controls				
	Ikeda et al (2018)[50]	10 445 cases 71 137 controls				
	Mullins et al (2021)[17]	41 917 cases 371 549 controls				
TRANK1	Chen et al (2013)[14]	7773 cases 9883 controls				
	Mühleisen et al (2014)[44]	9747 cases 14 278 controls				
	Ikeda et al (2018)[50]	10 445 cases 71 137 controls				
	Charney et al (2017)[49]	13 902 cases 19 279 controls				
	Hou et al (2016)[48]	9784 cases 30 471 controls				
	Stahl et al (2019)[51]	20 352 cases 31 358 controls				
	Mullins et al (2021)[17]	41 917 cases 371 549 controls				

GWAS, genome-wide association study; WES, whole-exome sequencing; WGS, whole-genome sequencing.
a. Primary GWAS.
b. Combined analysis.

data with genomic data in large-scale studies can help to obtain a greater knowledge of the brain regions that mediate the link between genetic risk and behavioural manifestations of bipolar disorder.[28] It will be advantageous to develop more genomic approaches to different complex disorders such as bipolar disorder in the near future, which will hopefully uncover the precise role of the genetic risk loci, especially the regulatory and non-coding elements, in the presentation of these disorders. By developing more advanced DNA analysis technologies and making them more accurate (to avoid ambiguous variant interpretations) and accessible at lower costs, we can achieve high-resolution insight into bipolar disorder genetic risk, and pave the way for further major discoveries that account for the missing heritability of bipolar disorder.[24]

To summarise, one of the most comprehensively replicated outcomes of the genomic approach in bipolar disorder is the high heterogeneity and polygenic background of the disorder. Although a large range of diverse genetic studies have been performed on bipolar disorder, the precise genetic underpinnings of the disorder are still unknown and need to be further explored. Some gene candidates have been identified more often in large genetic studies on bipolar disorder, indicating that they might be potential risk genes for bipolar disorder. The top five gene candidates discovered more through all investigated GWASs and WES and WGS studies are *ANK3*, *CACNA1C*, *SYNE1*, *ODZ4* and *TRANK1*. In the field of psychiatric genetics, where the environment has strong effects on the form and development of a disorder, it is critical to pay more attention to the issue of gene×environment interactions and sample heterogeneity. Bipolar disorder PRS in European patients explained approximately 4.57% of the phenotypic variances, and there is some shared risk loci between bipolar disorder and schizophrenia and major depressive disorder. Large-scale collaborative neuroimaging studies show a replicable pattern of lower cortical thickness, altered white matter integrity and smaller subcortical volumes in bipolar disorder. Common medications such as lithium show normalising associations with brain structure, whereas anticonvulsants appear to show the opposite effect and, in some cases, are associated with thinner or smaller brain structures. Importantly, and in line with prior reports, bipolar disorder appears to be associated with small-to-moderate effect sizes on MRI-derived brain markers. Replicating the findings on potential predictors in independent, well-defined large cohorts may ultimately lead to the discovery of most reliable biomarkers for this disorder.

Mojtaba Oraki Kohshour (ORCID), PhD, Institute of Psychiatric Phenomics and Genomics, University Hospital LMU Munich, Germany; and Department of Immunology, School of Medicine, Ahvaz Jundishapur University of Medical Sciences, Iran; **Sergi Papiol** (ORCID), PhD, Institute of Psychiatric Phenomics and Genomics, University Hospital LMU Munich, Germany; and Department of Psychiatry and Psychotherapy, University Hospital LMU Munich, Germany; **Christopher R. K. Ching** (ORCID), PhD, Imaging Genetics Center, Mark and Mary Stevens Neuroimaging and Informatics Institute, Keck School of Medicine, University of Southern California, USA; **Thomas G. Schulze** (ORCID), MD, Institute of Psychiatric Phenomics and Genomics, University Hospital LMU Munich, Germany; and Department of Psychiatry and Behavioral Sciences, SUNY Upstate Medical University, USA

Correspondence: Thomas G. Schulze. Email: thomas.schulze@med.uni-muenchen.de

First received 24 Sep 2021, final revision 1 Dec 2021, accepted 10 Dec 2021

Declaration of interest

T.G.S. is a member of the *BJPsych* Editorial Board and did not take part in the review or decision-making process for this paper. All other authors report no conflicts of interest.

References

1 Vieta E, Berk M, Schulze TG, Carvalho AF, Suppes T, Calabrese JR, et al. Bipolar disorders. *Nat Rev Dis Primers* 2018; **4**: 18008.

2 Leboyer M, Kupfer DJ. Bipolar disorder: new perspectives in health care and prevention. *J Clin Psychiatry* 2010; **71**(12): 1689–95.

3 Jain, R. Managing bipolar disorder from urgent situations to maintenance therapy. *J Clin Psychiatry* 2007; **9**(5): 367–80.

4 Amare AT, Schubert KO, Hou L, Clark SR, Papiol S, Cearns M, et al. Association of polygenic score for major depression with response to lithium in patients with bipolar disorder. *Mol Psychiatry* 2021; **26**(6): 2457–70.

5 Rees E, Owen MJ. Translating insights from neuropsychiatric genetics and genomics for precision psychiatry. *Genome Med* 2020; **12**: 43.

6 Ludwig B, Dwivedi Y. Dissecting bipolar disorder complexity through epigenomic approach. *Mol Psychiatry* 2016; **21**(11): 1490–8.

7 Tabor HK, Risch NJ, Myers RM. Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nat Rev Genet* 2002; **3**(5): 391–7.

8 Tam V, Patel N, Turcotte M, Bossé Y, Paré G, Meyre D. Benefits and limitations of genome-wide association studies. *Nat Rev Genet* 2019; **20**(8): 467–84.

9 Cichon S, Mühleisen TW, Degenhardt FA, Mattheisen M, Miró X, Strohmaier J, et al. Genome-wide association study identifies genetic variation in neurocan as a susceptibility factor for bipolar disorder. *Am J Hum Genet* 2011; **88**(3): 372–81.

10 Baum AE, Akula N, Cabanero M, Cardona I, Corona W, Klemens B, et al. A genome-wide association study implicates diacylglycerol kinase eta (DGKH) and several other genes in the etiology of bipolar disorder. *Mol Psychiatry* 2008; **13**(2): 197–207.

11 Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science* 1996; **273**(5281): 1516–7.

12 Witte JS. Genome-wide association studies and beyond. *Annu Rev Public Health* 2010; **31**: 9–20. 4 p following 20.

13 Park J-H, Wacholder S, Gail MH, Peters U, Jacobs KB, Chanock SJ, et al. Estimation of effect size distribution from genome-wide association studies and implications for future discoveries. *Nat Genet* 2010; **42**(7): 570–5.

14 Chen DT, Jiang X, Akula N, Shugart YY, Wendland JR, Steele CJM, et al. Genome-wide association study meta-analysis of European and Asian-ancestry samples identifies three novel loci associated with bipolar disorder. *Mol Psychiatry* 2013; **18**(2): 195–205.

15 Ozaki K, Ohnishi Y, Iida A, Sekine A, Yamada R, Tsunoda T, et al. Functional SNPs in the lymphotoxin-alpha gene that are associated with susceptibility to myocardial infarction. *Nat Genet* 2002; **32**(4): 650–4.

16 Gordovez FJA, McMahon FJ. The genetics of bipolar disorder. *Mol Psychiatry* 2020; **25**(3): 544–59.

17 Mullins N, Forstner AJ, O'Connell KS, Coombes B, Coleman JRI, Qiao Z, et al. Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. *Nat Genet* 2021; **53**(6): 817–29.

18 Marchini J, Howie B. Genotype imputation for genome-wide association studies. *Nat Rev Genet* 2010; **11**(7): 499–511.

19 Porcu E, Sanna S, Fuchsberger C, Fritsche LG. Genotype imputation in genome-wide association studies. *Curr Protoc Hum Genet* 2013; **Chapter 1** (suppl 78): Unit 1.25.1–1.25.14. (2013)

Supplementary material

Supplementary material is available online at https://doi.org/10.1192/bjo.2021.1082.

Data availability

Data availability is not applicable to this article as no new data were created or analysed in this review.

20 Truvé K, Parris TZ, Vizlin-Hodzic D, Salmela S, Berger E, Ågren H, et al. Identification of candidate genetic variants and altered protein expression in neural stem and mature neural cells support altered microtubule function to be an essential component in bipolar disorder. *Transl Psychiatry* 2020; **10**: 390.

21 Kato T. Whole genome/exome sequencing in mood and psychotic disorders. *Psychiatry Clin Neurosci* 2015; **69**(2): 65–76.

22 Monson ET, Pirooznia M, Parla J, Kramer M, Goes FS, Gaine ME, et al. Assessment of whole-exome sequence data in attempted suicide within a bipolar disorder cohort. *Mol Neuropsychiatry* 2017; **3**(1): 1–11.

23 Zhang T, Hou L, Chen DT, McMahon FJ, Wang J-C, Rice JP. Exome sequencing of a large family identifies potential candidate genes contributing risk to bipolar disorder. *Gene* 2018; **645**: 119–23.

24 Sanders SJ, Neale BM, Huang H, Werling DM, An J-Y, Dong S, et al. Whole genome sequencing in psychiatric disorders: the WGSPD Consortium. *Nat Neurosci* 2017; **20**(12): 1661–8.

25 Yi N, Xu S, Lou X-Y, Mallick H. Multiple comparisons in genetic association studies: a hierarchical modeling approach. *Stat Appl Genet Mol Biol* 2014; **13**(1): 35–48.

26 Hong EP, Park JW. Sample size and statistical power calculation in genetic association studies. *Genomics Inform* 2012; **10**(2): 117–22.

27 Eeltink E, van der Horst MZ, Zinkstok JR, Aalfs CM, Luykx JJ. Polygenic risk scores for genetic counseling in psychiatry: lessons learned from other fields of medicine. *Neurosci Biobehav Rev* 2021; **121**: 119–27.

28 Ching CRK, Hibar DP, Gurholt TP, Nunes A, Thomopoulos SI, Abé C, et al. What we learn about bipolar disorder from large-scale neuroimaging: findings and future directions from the ENIGMA bipolar disorder working group. *Hum Brain Mapp* 2022; **43**(1): 56–82.

29 Uffelmann E, Huang QQ, Munung NS, de Vries J, Okada Y, Martin AR, et al. Genome-wide association studies. *Nat Rev Methods Primers* 2021; **1**: 59.

30 Ku C-S, Naidoo N, Pawitan Y. Revisiting Mendelian disorders through exome sequencing. *Hum Genet* 2011; **129**(4): 351–70.

31 Ng PC, Kirkness EF. Whole genome sequencing. *Methods Mol Biol* 2010; **628**: 215–26.

32 Sullivan PF. The Psychiatric GWAS Consortium: big science comes to psychiatry. *Neuron* 2010; **68**(2): 182–6.

33 The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007; **447**(7145): 661–78.

34 Sklar P, Smoller JW, Fan J, Ferreira MAR, Perlis RH, Chambert K, et al. Whole-genome association study of bipolar disorder. *Mol Psychiatry* 2008; **13**(6): 558–69.

35 Ferreira MAR, O'Donovan MC, Meng YA, Jones IR, Ruderfer DM, Jones L, et al. Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. *Nat Genet* 2008; **40**(9): 1056–8.

36 Bauleo A, Montesanto A, Pace V, Brando R, De Stefano L, Puntorieri D, et al. Rare copy number variants in ASTN2 gene in patients with neurodevelopmental disorders. *Psychiatr Genet* 2021; **31**(6): 239–45.

37 Jiménez-Jiménez FJ, Esguevillas G, Alonso-Navarro H, Zurdo M, Turpín-Fenoll L, Millán-Pascual J, et al. Gamma-aminobutyric acid (GABA) receptors genes polymorphisms and risk for restless legs syndrome. *Pharmacogenomics J* 2018; **18**(4): 565–77.

38 Wang K-S, Liu X-F, Aragam N. A genome-wide meta-analysis identifies novel loci associated with schizophrenia and bipolar disorder. *Schizophr Res* 2010; **124**(1–3): 192–9.

39 Kerner B, Lambert CG, Muthén BO. Genome-wide association study in bipolar patients stratified by co-morbidity. *PLoS One* 2011; **6**(12): e28477.

40 Psychiatric GWAS Consortium Bipolar Disorder Working Group. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat Genet* 2011; **43**(10): 977–83.

41 Zheng G, Freidlin B, Li Z, Gastwirth JL. Genomic control for association studies under various genetic models. *Biometrics* 2005; **61**(1): 186–92.

42 Green EK, Grozeva D, Forty L, Gordon-Smith K, Russell E, Farmer A, et al. Association at SYNE1 in both bipolar disorder and recurrent major depression. *Mol Psychiatry* 2013; **18**(5): 614–7.

43 Green EK, Hamshere M, Forty L, Gordon-Smith K, Fraser C, Russell E, et al. 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* 2013; **18**(12): 1302–7.

44 Mühleisen TW, Leber M, Schulze TG, Strohmaier J, Degenhardt F, Treutlein J, et al. Genome-wide association study reveals two new risk loci for bipolar disorder. *Nat Commun* 2014; **5**: 3339.

45 Song J, Bergen SE, Di Florio A, Karlsson R, Charney A, Ruderfer DM, et al. Genome-wide association study identifies SESTD1 as a novel risk gene for lithium-responsive bipolar disorder. *Mol Psychiatry* 2016; **21**(9): 1290–7.

46 Song J, Bergen SE, Di Florio A, Karlsson R, Charney A, Ruderfer DM, et al. Genome-wide association study identifies SESTD1 as a novel risk gene for lithium-responsive bipolar disorder. *Mol Psychiatry* 2017; **22**(8): 1223.

47 Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* 2011; **88**(1): 76–82.

48 Hou L, Bergen SE, Akula N, Song J, Hultman CM, Landén M, et al. Genome-wide association study of 40,000 individuals identifies two novel loci associated with bipolar disorder. *Hum Mol Genet* 2016; **25**(15): 3383–94.

49 Charney AW, Ruderfer DM, Stahl EA, Moran JL, Chambert K, Belliveau RA, et al. Evidence for genetic heterogeneity between clinical subtypes of bipolar disorder. *Transl Psychiatry* 2017; **7**: e993.

50 Ikeda M, Takahashi A, Kamatani Y, Okahisa Y, Kunugi H, Mori N, et al. A genome-wide association study identifies two novel susceptibility loci and trans population polygenicity associated with bipolar disorder. *Mol Psychiatry* 2018; **23**(3): 639–47.

51 Stahl EA, Breen G, Forstner AJ, McQuillin A, Ripke S, Trubetskoy V, et al. Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat Genet* 2019; **51**(5): 793–803.

52 Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009; **460**(7256): 748–52.

53 Andlauer TFM, Guzman-Parra J, Streit F, Strohmaier J, González MJ, Gil Flores S, et al. Bipolar multiplex families have an increased burden of common risk variants for psychiatric disorders. *Mol Psychiatry* 2021; **26**(4): 1286–98.

54 Guzman-Parra J, Streit F, Forstner AJ, Strohmaier J, González MJ, Gil Flores S, et al. Clinical and genetic differences between bipolar disorder type 1 and 2 in multiplex families. *Transl Psychiatry* 2021; **11**: 31.

55 Mistry S, Harrison JR, Smith DJ, Escott-Price V, Zammit S. The use of polygenic risk scores to identify phenotypes associated with genetic risk of bipolar disorder and depression: a systematic review. *J Affect Disord* 2018; **234**: 148–55.

56 Hou L, Heilbronner U, Degenhardt F, Adli M, Akiyama K, Akula N, et al. Genetic variants associated with response to lithium treatment in bipolar disorder: a genome-wide association study. *Lancet* 2016; **387**(10023): 1085–93.

57 Amare AT, Schubert KO, Hou L, Clark SR, Papiol S, Heilbronner U, et al. Association of polygenic score for schizophrenia and HLA antigen and inflammation genes with response to lithium in bipolar affective disorder: a genome-wide association study. *JAMA Psychiatry* 2018; **75**(1): 65–74.

58 Toma C, Shaw AD, Allcock RJN, Heath A, Pierce KD, Mitchell PB, et al. An examination of multiple classes of rare variants in extended families with bipolar disorder. *Transl Psychiatry* 2018; **8**: 65.

59 Biesecker LG, Shianna KV, Mullikin JC. Exome sequencing: the expert view. *Genome Biol* 2011; **12**(9): 128.

60 Chen Y-C, Carter H, Parla J, Kramer M, Goes FS, Pirooznia M, et al. A hybrid likelihood model for sequence-based disease association studies. *PLoS Genet* 2013; **9**(1): e1003224.

61 Kataoka M, Matoba N, Sawada T, Kazuno A-A, Ishiwata M, Fujii K, et al. Exome sequencing for bipolar disorder points to roles of de novo loss-of-function and protein-altering mutations. *Mol Psychiatry* 2016; **21**(7): 885–93.

62 Lescai F, Als TD, Li Q, Nyegaard M, Andorsdottir G, Biskopstø M, et al. Whole-exome sequencing of individuals from an isolated population implicates rare risk variants in bipolar disorder. *Transl Psychiatry* 2017; **7**: e1034.

63 Cruceanu C, Schmouth J-F, Torres-Platas SG, Lopez JP, Ambalavanan A, Darcq E, et al. Rare susceptibility variants for bipolar disorder suggest a role for G protein-coupled receptors. *Mol Psychiatry* 2018; **23**(10): 2050–6.

64 Salvoro C, Bortoluzzi S, Coppe A, Valle G, Feltrin E, Mostacciuolo ML, et al. Rare risk variants identification by identity-by-descent mapping and whole-exome sequencing implicates neuronal development pathways in schizophrenia and bipolar disorder. *Mol Neurobiol* 2018; **55**(9): 7366–76.

65 Ganna A, Satterstrom FK, Zekavat SM, Das I, Kurki MI, Churchhouse C, et al. Quantifying the impact of rare and ultra-rare coding variation across the phenotypic spectrum. *Am J Hum Genet* 2018; **102**(6): 1204–11.

66 Maaser A, Forstner AJ, Strohmaier J, Hecker J, Ludwig KU, Sivalingam S, et al. Exome sequencing in large, multiplex bipolar disorder families from Cuba. *PLoS One* 2018; **13**(10): e0205895.

67 Husson T, Duboc J-B, Quenez O, Charbonnier C, Rotharmel M, Cuenca M, et al. Identification of potential genetic risk factors for bipolar disorder by whole-exome sequencing. *Transl Psychiatry* 2018; **8**: 268.

68 Han M-R, Han K-M, Kim A, Kang W, Kang Y, Kang J, et al. Whole-exome sequencing identifies variants associated with structural MRI markers in patients with bipolar disorders. *J Affect Disord* 2019; **249**: 159–68.

69 Ganesh S, Ahmed P H, Nadella RK, More RP, Seshadri M, Viswanath B, et al. Exome sequencing in families with severe mental illness identifies novel and rare variants in genes implicated in Mendelian neuropsychiatric syndromes. *Psychiatry Clin Neurosci* 2019; **73**(1): 11–9.

70 Szatkiewicz J, Crowley JJ, Adolfsson AN, Åberg KA, Alaerts M, Genovese G, et al. The genomics of major psychiatric disorders in a large pedigree from Northern Sweden. *Transl Psychiatry* 2019; **9**: 60.

71 Anjanappa RM, Nayak S, Moily NS, Manduva V, Nadella RK, Viswanath B, et al. A linkage and exome study implicates rare variants of KANK4 and CAP2 in bipolar disorder in a multiplex family. *Bipolar Disord* 2020; **22**(1): 70–8.

72 Forstner AJ, Fischer SB, Schenk LM, Strohmaier J, Maaser-Hecker A, Reinbold CS, et al. Whole-exome sequencing of 81 individuals from 27 multiply affected bipolar disorder families. *Transl Psychiatry* 2020; **10**: 57.

73 Jacobs A, Hagin M, Shugol M, Shomron N, Pillar N, Fananas L, et al. The black sheep of the family- whole-exome sequencing in family of lithium response discordant bipolar monozygotic twins. *Eur Neuropsychopharmacol* 2020; 34: 19-27.

74 Courtois E, Schmid M, Wajsbrot O, Barau C, Le Corvoisier P, Aouizerate B, et al. Contribution of common and rare damaging variants in familial forms of bipolar disorder and phenotypic outcome. *Transl Psychiatry* 2020; **10**: 124.

75 Engelbrecht H-R, Dalvie S, Agenbag G, Stein DJ, Ramesar RS. Whole-exome sequencing in an Afrikaner family with bipolar disorder. *J Affect Disord* 2020; **276**: 69–75.

76 Umehara H, Nakamura M, Nagai M, Kato Y, Ueno S-I, Sano A. Positional cloning and comprehensive mutation analysis of a Japanese family with lithium-responsive bipolar disorder identifies a novel DOCK5 mutation. *J Hum Genet* 2021; **66**(3): 243–9.

77 Toma C, Shaw AD, Heath A, Pierce KD, Mitchell PB, Schofield PR, et al. A linkage and exome study of multiplex families with bipolar disorder implicates rare coding variants of ANK3 and additional rare alleles at 10q11-q21. *J Psychiatry Neurosci* 2021; **46**(2): E247–57.

78 Halldorsson BV, Eggertsson HP, Moore KHS, Hauswedell H, Eiriksson O, Ulfarsson MO, et al. The sequences of 150,119 genomes in the UK Biobank. *BioRxiv* [Preprint] 2021. Available from: https://www.biorxiv.org/content/10.1101/2021.11.16.468246v1 [cited 20 Nov 2021].

79 Fiorentino A, O'Brien NL, Locke DP, McQuillin A, Jarram A, Anjorin A, et al. Analysis of ANK3 and CACNA1C variants identified in bipolar disorder whole genome sequence data. *Bipolar Disord* 2014; **16**(6): 583–91.

80 Georgi B, Craig D, Kember RL, Liu W, Lindquist I, Nasser S, et al. Genomic view of bipolar disorder revealed by whole genome sequencing in a genetic isolate. *PLoS Genet* 2014; **10**(3): e1004229.

81 Ross J, Gedvilaite E, Badner JA, Erdman C, Baird L, Matsunami N, et al. A rare variant in CACNA1D segregates with 7 bipolar I disorder cases in a large pedigree. *Mol Neuropsychiatry* 2016; **2**(3): 145–50.

82 Fiorentino A, O'Brien NL, Sharp SI, Curtis D, Bass NJ, McQuillin A. Genetic variation in the miR-708 gene and its binding targets in bipolar disorder. *Bipolar Disord* 2016; **18**(8): 650–6.

83 Bouwkamp CG, Kievit AJA, Olgiati S, Breedveld GJ, Coesmans M, Bonifati V, et al. A balanced translocation disrupting BCL2L10 and PNLDC1 segregates with affective psychosis. *Am J Med Genet B Neuropsychiatr Genet* 2017; **174**(3): 214–9.

84 Steinberg S, Gudmundsdottir S, Sveinbjornsson G, Suvisaari J, Paunio T, Torniainen-Holm M, et al. Truncating mutations in RBM12 are associated with psychosis. *Nat Genet* 2017; **49**(8): 1251–4.

85 Sharp SI, Lange J, Kandaswamy R, Daher M, Anjorin A, Bass NJ, et al. Identification of rare nonsynonymous variants in SYNE1/CPG2 in bipolar affective disorder. *Psychiatr Genet* 2017; **27**(3): 81–8.

86 Goes FS, Pirooznia M, Tehan M, Zandi PP, McGrath J, Wolyniec P, et al. De novo variation in bipolar disorder. *Mol Psychiatry* 2021; 26: 4127–36.

87 Nakajima K, Miranda A, Craig DW, Shekhtman T, Kmoch S, Bleyer A, et al. Ntrk1 mutation co-segregating with bipolar disorder and inherited kidney disease in a multiplex family causes defects in neuronal growth and depression-like behavior in mice. *Transl Psychiatry* 2020; **10**: 407.

88 Sul JH, Service SK, Huang AY, Ramensky V, Hwang S-G, Teshiba TM, et al. Contribution of common and rare variants to bipolar disorder susceptibility in extended pedigrees from population isolates. *Transl Psychiatry* 2020; **10**: 74.

89 Phillips ML, Swartz HA. A critical appraisal of neuroimaging studies of bipolar disorder: toward a new conceptualization of underlying neural circuitry and a road map for future research. *Am J Psychiatry* 2014; **171**(8): 829–43.

90 Hajek T, Kopecek M, Kozeny J, Gunde E, Alda M, Höschl C. Amygdala volumes in mood disorders–meta-analysis of magnetic resonance volumetry studies. *J Affect Disord* 2009; **115**(3): 395–410.

91 Rimol LM, Hartberg CB, Nesvåg R, Fennema-Notestine C, Hagler DJJ, Pung CJ, et al. Cortical thickness and subcortical volumes in schizophrenia and bipolar disorder. *Biol Psychiatry* 2010; **68**(1): 41–50.

92 Hajek T, Kopecek M, Höschl C, Alda M. Smaller hippocampal volumes in patients with bipolar disorder are masked by exposure to lithium: a meta-analysis. *J Psychiatry Neurosci* 2012; **37**(5): 333–43.

93 Rimol LM, Nesvåg R, Hagler DJJ, Bergmann O, Fennema-Notestine C, Hartberg CB, et al. Cortical volume, surface area, and thickness in schizophrenia and bipolar disorder. *Biol Psychiatry* 2012; **71**(6): 552–60.

94 Hanford LC, Nazarov A, Hall GB, Sassi RB. Cortical thickness in bipolar disorder: a systematic review. *Bipolar Disord* 2016; **18**(1): 4–18.

95 Thompson PM, Jahanshad N, Ching CRK, Salminen LE, Thomopoulos SI, Bright J, et al. ENIGMA and global neuroscience: a decade of large-scale studies of the brain in health and disease across more than 40 countries. *Transl Psychiatry* 2020; **10**: 100.

96 Hibar DP, Westlye LT, van Erp TGM, Rasmussen J, Leonardo CD, Faskowitz J, et al. Subcortical volumetric abnormalities in bipolar disorder. *Mol Psychiatry* 2016; **21**(12): 1710–6.

97 Hibar DP, Westlye LT, Doan NT, Jahanshad N, Cheung JW, Ching CRK, et al. Cortical abnormalities in bipolar disorder: an MRI analysis of 6503 individuals from the ENIGMA Bipolar Disorder Working Group. *Mol Psychiatry* 2018; **23**(4): 932–42.

98 Favre P, Pauling M, Stout J, Hozer F, Sarrazin S, Abé C, et al. Widespread white matter microstructural abnormalities in bipolar disorder: evidence from mega- and meta-analyses across 3033 individuals. *Neuropsychopharmacology* 2019; **44**(13): 2285–93.

99 Haukvik UK, Gurholt TP, Nerland S, Elvsåshagen T, Akudjedu TN, Alda M, et al. In vivo hippocampal subfield volumes in bipolar disorder-a mega-analysis from The Enhancing Neuro Imaging Genetics through Meta-Analysis Bipolar Disorder Working Group. *Hum Brain Mapp* 2022; **43**(1): 385–98.

100 de Zwarte SMC, Brouwer RM, Agartz I, Alda M, Aleman A, Alpert KI, et al. The association between familial risk and brain abnormalities is disease specific: an ENIGMA-relatives study of schizophrenia and bipolar disorder. *Biol Psychiatry* 2019; **86**(7): 545–56.

101 de Zwarte SMC, Brouwer RM, Agartz I, Alda M, Alonso-Lana S, Bearden CE, et al. Intelligence, educational attainment, and brain structure in those at familial high-risk for schizophrenia or bipolar disorder. *Hum Brain Mapp* 2022; **43**(1): 414–30.

102 Nunes A, Schnack HG, Ching CRK, Agartz I, Akudjedu TN, Alda M, et al. Using structural MRI to identify bipolar disorders - 13 site machine learning study in 3020 individuals from the ENIGMA bipolar disorders working group. *Mol Psychiatry* 2020; **25**(9): 2130–43.

103 Radonjić NV, Hess JL, Rovira P, Andreassen O, Buitelaar JK, Ching CRK, et al. Structural brain imaging studies offer clues about the effects of the shared genetic etiology among neuropsychiatric disorders. *Mol Psychiatry* 2021; **26**(6): 2101–10.

104 Patel Y, Parker N, Shin J, Howard D, French L, Thomopoulos SI, et al. Virtual histology of cortical thickness and shared neurobiology in 6 psychiatric disorders. *JAMA Psychiatry* 2021; **78**(1): 47–63.

105 Grasby KL, Jahanshad N, Painter JN, Colodro-Conde L, Bralten J, Hibar DP, et al. The genetic architecture of the human cerebral cortex. *Science* 2020; **367**(6484): eaay6690.

106 Iqbal Z, Vandeweyer G, van der Voet M, Waryah AM, Zahoor MY, Besseling JA, et al. Homozygous and heterozygous disruptions of ANK3: at the crossroads of neurodevelopmental and psychiatric disorders. *Hum Mol Genet* 2013; **22**(10): 1960–70.

107 Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, et al. The GeneCards suite: from gene data mining to disease genome sequence analyses. *Curr Protoc Bioinformatics* 2016; **54**: 1.30.1–33.

108 Kabir ZD, Martínez-Rivera A, Rajadhyaksha AM. From gene to behavior: L-type calcium channel mechanisms underlying neuropsychiatric symptoms. *Neurotherapeutics* 2017; **14**(3): 588–613.

109 Li W, Yang Y, Luo B, Zhang Y, Song X, Li M, et al. Association of SYNE1 locus with bipolar disorder in Chinese population. *Hereditas* 2019; **156**: 19.

110 Rathje M, Waxman H, Benoit M, Tammineni P, Leu C, Loebrich S, et al. Genetic variants in the bipolar disorder risk locus SYNE1 that affect CPG2 expression and protein function. *Mol Psychiatry* 2021; **26**(2): 508–23.

111 Hor H, Francescatto L, Bartesaghi L, Ortega-Cubero S, Kousi M, Lorenzo-Betancor O, et al. Missense mutations in TENM4, a regulator of axon guidance and central myelination, cause essential tremor. *Hum Mol Genet* 2015; **24**(20): 5677–86.

112 Chang H, Cai X, Yang Z-H, Xiao X, Li M. Regulation of TRANK1 by GSK-3 in the brain: unexpected interactions. *Mol Psychiatry* [Epub ahead of print] 30 Apr 2021. Available from: https://doi.org/10.1038/s41380-021-01102-2.

113 Peterson RE, Kuchenbaecker K, Walters RK, Chen C-Y, Popejoy AB, Periyasamy S, et al. Genome-wide association studies in ancestrally diverse populations: opportunities, methods, pitfalls, and recommendations. *Cell* 2019; **179**(3): 589–603.