Application of bioinformatics tools for better interpretation of psychiatric GWAS through integration of biological information

A Cumulative Dissertation

Submitted to
Faculty of Psychology, University of Basel
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

by
Céline Sonja Reinbold, MSc
from Ehrendingen, Switzerland

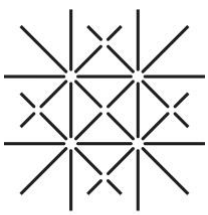
Basel, Switzerland
2018

Approved by the Faculty of Psychology
at the request of

Prof. Dr. rer. nat. Sven Cichon
Prof. Dr. med. Andreas Papassotiropoulos

Basel, _____

_____
Prof. Dr. Alexander Grob



Acknowledgments

This thesis symbolizes a key landmark in my life and I'm deeply grateful for each and every one who enabled me reaching this milestone, regardless of how bumpy the roads have been.

First and foremost, I want to thank Sven for his encouragement and guidance during my scientific adventures, our inspiring discussions and most of all the support during all stages of the life of a PhD student. Special thanks to Per and Stefan who always stood in my corner, loudly encouraged me to pursue my ambitions and never failed to raise spirit when needed. I want to thank Sascha for his gracious support during all circumstances and making me chuckle in the most awkward moments. You will always be my first and true work spouse. I am sincerely grateful to Angie who nourished my curiosity right from the start and enabled me, personally and professionally, to commence this part of my life in the first place. I also want to thank Andreas and Dominique who helped building a solid and extensive foundation for my scientific future and inspired me. I'm indebted to all the awesome scientists I met along the road whose expertise and meaningful insights helped me grow.

Very special thanks to the best Mom I could ever wish for. You supported me above and beyond in so many ways I cannot thank you enough! I am especially grateful for all the amazing people who became my family by affection whose unconditional support and love kept me going. I want to thank my fantastic friends for all the memorable moments, patience, inspiring discussion and support especially when time was short.

Last but not least I thank Odin for making my life brighter and fluffier.

Abstract

Psychiatric disorders are common, heritable, often chronic and devastating illnesses who heavily decline quality of life of the patients and their environment. The high genetic correlations across the disorders and their diagnostic criteria reinforce the growing discomfort with the current classification and boost the search for more refined measurements. Genome-wide association studies (GWAS) are a highly successful method for identifying common genetic risk variants underlying common disorders. In psychiatric disorders, the emerging picture suggests contribution from a large number of single-nucleotide polymorphisms (SNPs) of individually small effect sizes as well as rare copy number variants (CNVs) and rare variants discovered by next-generation sequencing. Most of these findings have emerged during the last years through large collaborative efforts which enabled powerful meta-analyses. Nevertheless, individual SNPs and CNVs seem to explain only a minor fraction of the heritable variance for psychiatric disorders. Therefore, the development and correct application of novel bioinformatics methods is necessary to cope with the limitations inherent to GWAS. Biology-informed methods already led to important advances with many discoveries of common, rare and de novo variants that are converging on specific pathways and biological mechanisms.

The studies described in this thesis aim to deepen our understanding of psychiatric disorders through the application of novel bioinformatics tools to existing GWAS data sets. We found evidence that schizophrenia-associated loci contribute to the development of bipolar disorder and that the overlapping SNPs converge in pathways previously reported in other psychiatric disorders. We revealed two genes and a pathway significantly associated with borderline personality disorder previously implicated in mental disorders and demonstrated the statistically significant genetic overlap with other psychiatric disorders. We identified two pathways suggesting an involvement of neurodevelopmental processes in the etiology of bipolar disorder. We found that common variants at nine previously reported BD-associated miRNAs do not strongly contribute to the differential responses to lithium treatment in BD. Taken together, these studies show that the application of biology-informed bioinformatic methods enhance the insights gained from GWAS and demonstrate the plethora of methods available nowadays. It is the hope that the progress in understanding the genetic architecture of psychiatric disorders will also help to improve the clinical classification and ultimately yield in better treatment options.

Table of Contents

Table Index

Abbreviations

BD	Bipolar disorder
CNV	Copy number variation
DNA	Deoxyribonucleic acid
DSM	Diagnostic and Statistical Manual of Mental Disorders
FDR	False discovery rate
GO	Gene ontology
GSA	Gene-set analysis
GSA-SNP	Gene-set analysis for SNPs
GSEA	Gene-set enrichment analysis
GWAS	Genome-wide association study
h^2	Heritability
h^2_{SNP}	SNP-chip heritability
HWE	Hardy-Weinberg Equilibrium
INDEL	insertion or deletion
INRICH	Interval enrichment analysis
IPA	Ingenuity Pathway Analysis
iPSC	induced pluripotent stem cells
i-GSEA4GWASv2	improved gene set enrichment analysis for genome-wide association study version 2
kb	Kilo base
KEGG	Kyoto encyclopedia of genes and genomes
LD	Linkage disequilibrium
MAF	Minor allele frequency
MAGENTA	Meta-analysis gene-set enrichment of variant association
MAGMA	Multi-marker analysis of genomic annotation
MDD	Major depressive disorder
MHC	Major histocompatibility complex
miRNA	Micro ribonucleic acid
mRNA	Messenger ribonucleic acid
OR	Odds ratio
PCA	Principal component analysis
PRS	Polygenic risk score
QC	Quality control
Q-Q plot	Quantile-quantile plot
SCZ	Schizophrenia
SNP	Single nucleotide polymorphism
UCSC	University of California Santa Cruz
UTR	Untranslated region
VEGAS	Versatile gene-based association study
WES	Whole-exome sequencing
WGS	Whole-genome sequencing

1 Introduction

Mental disorders are among the disorders with the highest non-fatal burden (GBD 2016 Disease and Injury Incidence and Prevalence Collaborators, 2017) and have devastating effects on the quality of life of patients and their environment. This is aggravated by the fact that there is still no satisfactory drug treatment available (Breen et al., 2016). Psychiatric disorders are categorized as common disorders with lifetime prevalence estimates for anxiety disorders of 28.8% and 20.8% for mood disorders (Kessler et al., 2005). A major challenge in the evaluation of biological underpinnings of psychiatric disorders is their pronounced heterogeneity. Diagnoses are based on structured questionnaires, a combination of various symptoms and a minimal number and duration of these symptoms required to fulfill the criteria of a diagnosis (Breen et al., 2016; Papassotiropoulos & de Quervain, 2015). Different disease etiologies may result in the same clinical diagnosis. Despite these obstacles, substantial heritability (h^2) estimates have been reported ranging from 0.37 for major depressive disorder (MDD) to 0.75 for bipolar disorder (BD) up to 0.81 for schizophrenia (SCZ) (Sullivan, Daly, & O'Donovan, 2012). Genome-wide association studies (GWAS) have become the major tool in the unbiased investigation of common variants in common disorders since their underlying rationale is the “common disease, common variant” hypothesis, stating that common disorders are caused at least in part by variants shared by more than 1-5% of the population (Pritchard, 2001; Reich & Lander, 2001). Soon after the first successful studies of GWAS emerged, consortia were formed to rapidly increase sample sizes (Psychiatric GWAS Consortium Steering Committee, 2009). The public availability of the GWAS results from consortia accelerated research even further. However, the biological interpretation of GWAS results remains a key challenge since the function of many single-nucleotide polymorphisms (SNPs) is not well understood thereby the interpretation of the respective SNP is mostly based on the gene function of the corresponding gene. This may be a reasonable approach, but the mapping of SNPs to genes is not without a challenge as they may lie outside gene boundaries. Furthermore, even well-powered GWAS have only been able to explain a small portion of the phenotypic variance leaving plenty of missing heritability (Manolio et al., 2009; Visscher, 2008; Visscher et al., 2017). Despite the success of GWAS it has become abundantly clear that this method is just the beginning since testing for association of single loci is insufficient in dissecting the complex genetic architecture underlying

psychiatric disorders. Gene and gene-set based methods can be seen as complementary follow-up approaches since they focus on the joint effect of SNPs. Typical gene-set analysis (GSA) approaches aggregate SNPs to genes and then aggregate them to sets of genes based on shared properties. These properties can be based on biological or functional characteristics and can be retrieved from databases or created by the researcher. The application of GSA to GWAS data has proven to be a valuable approach since it addresses several limitations characteristic to GWAS. GSA enables the interpretation of the joint effect of SNPs with moderate effects on the basis of prior biological or functional knowledge which is not possible when only GWAS are computed. Due to the polygenic nature of psychiatric disorders the accumulation of these variants will empower the detection of genetic risk factors (The Network Pathway Analysis Subgroup of the Psychiatric Genomics Consortium, 2015). A bulk of GSA methods have been proposed over the years but no gold-standard has been defined yet resulting in various challenges like absence of comparability, reproducibility and reliability (de Leeuw, Neale, Heskes, & Posthuma, 2016; Mooney & Wilmot, 2015; Ramanan, Shen, Moore, & Saykin, 2012; L. Wang, Jia, Wolfinger, Chen, & Zhao, 2011). Thanks to GSA, not only the biology-based interpretation of GWAS results but also the detection of novel variants and genes associated with the disorder is possible. Moreover, the easy accessibility of GWAS led to the development of many more methods e.g. methods aiming to improve GWAS algorithms by optimizing limitations inherent to GWAS (de Leeuw, Mooij, Heskes, & Posthuma, 2015; Loh et al., 2015; Svishcheva, Axenovich, Belonogova, van Duijn, & Aulchenko, 2012), methods focusing on fine-mapping and deducing causalities (Benner et al., 2016; Bowden, Davey Smith, & Burgess, 2015), or methods estimating and partitioning genetic variance (Bulik-Sullivan, Finucane, et al., 2015; Bulik-Sullivan, Loh, et al., 2015; Finucane et al., 2015; J. J. Lee, McGue, Iacono, & Chow, 2018). Other methods also enable the investigation of genetic correlation between disorders or the predictions of disease risks (Purcell et al., 2009). The application of these bioinformatics methods enabled insights into disease-related biological processes by identifying hundreds of low-frequency and common variants that contribute to psychiatric disorders and revealed the genetic overlap between disorders (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013; Visscher et al., 2017). But despite all the new insights and technical advances in the field, the underlying pathological mechanisms

of psychiatric disorders still remain elusive and drug discovery is halting (Fibiger, 2012).

This thesis aims to contribute to the field of psychiatric genomics by leveraging (publicly) available data-sets and bioinformatics tools primarily focusing on the biologically driven computational analysis of bipolar disorder from different angles. To achieve this, various gene and gene-set based methods were applied to different cohorts enabling the discovery of new susceptibility genes and gene-sets.

This thesis is based on the following four publications. The letters indicate my contributions to each publication and are listed after each reference: **A** - Designed the experiment or contributed to the design; **B** - Performed the experiment; **C** - Analyzed the data or contributed to the analysis; **D** - Wrote the paper or contributed to paper writing; * - these authors contributed equally.

- Forstner, A. J.*, Hecker, J.*, Hofmann, A., Maaser, A., Reinbold, C. S., Mühleisen, T. W., . . . Nöthen, M. M. (2017). Identification of shared risk loci and pathways for bipolar disorder and schizophrenia. PLoS ONE, 12(2), e0171595. doi:10.1371/journal.pone.0171595 **(C-D)**
- Witt, S. H.*, Streit, F.*, Jungkunz, M., Frank, J., Awasthi, S., Reinbold, C. S., . . . Rietschel, M. (2017). Genome-wide association study of borderline personality disorder reveals genetic overlap with bipolar disorder, major depression and schizophrenia. Transl Psychiatry, 7(6), e1155. doi:10.1038/tp.2017.115 **(C-D)**
- Mühleisen, T. W.*, Reinbold, C. S.*, Forstner, A. J., Abramova, L. I., Alda, M., Babadjanova, G., . . . Cichon, S. (2017). Gene set enrichment analysis and expression pattern exploration implicate an involvement of neurodevelopmental processes in bipolar disorder. Journal of Affective Disorders. doi:10.1016/j.jad.2017.11.068 **(A-D)**
- Reinbold, C. S.*, Forstner, A. J.*, Hecker, J., Fullerton, J. M., Hoffmann, P., Hou, L., . . . Cichon, S. (2018). Analysis of the influence of microRNAs in Lithium Response in Bipolar Disorder. Accepted for publication in Frontiers in Psychiatry. doi:10.3389/fpsyt.2018.00207 **(A-D)**

2 Theoretical Background

2.1 Psychiatric disorders

Mental disorders may be very broadly summarized by problems that people experience with their mind and their mood. Their devastating impact on the quality of life of patients, their relatives and the society as a whole has been reported widely (GBD 2016 Disease and Injury Incidence and Prevalence Collaborators, 2017). One of the most frequently used standardized manuals to diagnose psychiatric disorders is the Diagnostic and Statistical Manual of Mental Disorders (DSM) of the American Psychiatric Association now with its fifth edition released (American Psychiatric Association, 2013b). Therein, the diagnosis of a disorders is based on checklists with disorders defined by the presence of a minimal number of symptoms, a required duration and the associated distress or disability, resulting in clinical heterogeneity within disorders (Breen et al., 2016; Papassotiropoulos & de Quervain, 2015). As defined by the American Psychiatric Association (2013a), bipolar disorder is a chronic mental disease characterized by recurrent episodes of depression and mania or hypomania with a mean age at onset of the first episode with approximately 18 years. Additionally, patients frequently suffer from co-occurring mental disorders. Approximately three-fourths of BD patients suffer from any anxiety disorder and more than half from a substance use disorder (Merikangas et al., 2011). This is aggravated by the fact that the lifetime risk of committing suicide in people suffering from bipolar disorder has been estimated to be at least 15 times that of the general population (Marangell et al., 2006). Epidemiological studies revealed unsettling lifetime prevalence estimates of 46.6% of participants suffering from at least one of the DSM-IV disorders assessed (Kessler et al., 2005) with still no satisfactory drug treatment available for any disorder (Breen et al., 2016).

The familial aggregation for most of the major psychiatric conditions has been reported since the very beginning of the systematic investigation of psychiatric disorders (Kendler & Eaves, 2005). Moreover, early family studies already suggested that multiple psychiatric disorders cluster within affected families leading to the assumption that heritable factors within and across disorders must exist (Kendler et al., 2011).

However successful the recent investigation of genetic contributions to psychiatric disorders has been, it has also raised concerns such as the categorical definition of disorders since genetic analysis clearly suggest a more continuous relation between disorders and health (Larsson, Anckarsater, Råstam, Chang, & Lichtenstein, 2011; Robinson et al., 2016).

Even though the application of the analysis methods within this thesis was mainly to gain insights into the biological underpinnings of BD, it is important to note that the methods and strategies can easily be applied to any other psychiatric disorders or genetically complex traits in general.

2.2 Genetic architecture of psychiatric disorders

Genetic architecture refers to the broad-sense phenotypic heritability since it can be summarized as all the characteristics of genetic variation contributing to the heritable phenotypic variability (Mackay, 2001). More specifically, this refers to the nature and number of genetic variants contributing to a disease, their population frequencies and effect sizes and their interactions with each other and the environment (Gratten, Wray, Keller, & Visscher, 2014). The exposure of the genetic architecture of a complex disorder is elementary when aiming to fully understand its cause of disease. An addition to the traditional concept of heritability is the so-called SNP-chip heritability (h^2_{SNP}) referring to the proportion of variance explained by all variants assayed by GWAS arrays (Wray et al., 2014). Estimates suggest that one-third to two-thirds of heritability of complex disorders can be explained by common and imputed SNPs (Manolio et al., 2009; Visscher et al., 2017; Yang et al., 2013). It is important to keep in mind that the human genome not just varies between people on the level of single nucleotides but also on a structural level, including copy number variations (CNVs), insertions or deletions (INDELs) and translocations (Alkan, Coe, & Eichler, 2011; Sudmant et al., 2015). Another distinction is based on the minor allele frequency (MAF) of a variant where common, low-frequency and rare genetic variants are defined herein as those with a MAF of ≥5%, ≥1% but <5% and <1%, respectively (Welter et al., 2014). The architecture of a trait does not only influence the choice of analysis method but also the whole design of a study. Since complex traits are assumed to be polygenic, the contribution of many, common and ancient variants with small effect sizes are implicated resulting in large population-based cohorts and genome-wide analysis strategies. Even though this work focuses on common variants the contribution of rare,

de novo or structural variants e.g. the recurrent 22q11.2 deletion in SCZ has been widely established (Bassett, Marshall, Lionel, Chow, & Scherer, 2008; Kirov, 2015). Whole-exome sequencing (WES) and whole-genome sequencing (WGS) studies which have only recently become feasible, contribute greatly to the understanding of disease mechanisms since they allow the accurate detection of rare and structural variants more accurately (Sanders et al., 2017; Zarrei, MacDonald, Merico, & Scherer, 2015). Interestingly, early micro-array-based studies already allowed to reliably detect rare variants (large microdeletions and –duplications covering at least several hundred kilobases) associated with psychiatric disorders and suggest that the rare variants converge on the same biological pathways as the common variants (Fromer et al., 2014; Purcell et al., 2014; Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014).

Pleiotropy describes the phenomenon that a specific variant has an effect on multiple traits. This phenomenon has often been described in psychiatric disorders (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013; Pickrell et al., 2016). Despite the high pleiotropy and phenomenological overlap, studies also revealed that the genetic architecture varies between psychiatric disorders with for example smaller rates of rare, de novo variants and CNVs in schizophrenia and bipolar disorder than in autism (Visscher et al., 2017).

In the case of BD, GWAS have identified the first susceptibility genes (Cichon et al., 2011; Mühleisen et al., 2014; Sklar et al., 2011; Stahl et al., 2018). So far, GWAS as well as gene-set analyses have suggested major roles for calcium signal transmission, neurodevelopmental genes, and microRNAs/non-coding RNAs (Forstner et al., 2015; Mühleisen et al., 2017; Sklar et al., 2011). However, the majority of underlying pathways and regulatory networks remain unknown (Nurnberger et al., 2014). For bipolar disorder, common alleles are estimated to explain 25-38% of the phenotypic variance resulting in a substantial part of the heritability unexplained (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013; Sang H. Lee, Wray, Goddard, & Visscher, 2011). It is hypothesized that rare variants with higher penetrance may contribute to BD susceptibility and account at least for a part of the hidden heritability (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013; Goes, 2016). This hypothesis may be particularly promising in severely affected individuals with a strong family history of the disease. Even though some studies suggest an influence of large CNVs in the etiology of BD, it seems that they do not play

a major role in BD (Green et al., 2016; Priebe et al., 2012). Preliminary results from sequencing studies suggest an enrichment of rare variants in specific gene-sets, such as axon guidance, calcium signaling, G protein-coupled receptors and potassium channels (Ament et al., 2015; Cruceanu et al., 2017; Fiorentino et al., 2014; Georgi et al., 2014; Goes et al., 2016; Strauss et al., 2014).

As briefly outlined above, impressive progress has been made over the last 10 years with regard to understanding the genetic architecture of psychiatric disorders. Part of this success is attributable to technological (such as SNP microarray and next-generation sequencing technology) and methodological (GWAS, WES, WGS; bioinformatics analyses) developments, part to better insights into the number and nature of the involved variants and the enormous sample sizes needed to successfully identify them. These findings will enable more accurate and biology-informed diagnosis, screenings, prognosis and therapies (Timpson, Greenwood, Soranzo, Lawson, & Richards, 2017).

2.3 Data analytics for common variants

GWAS were developed to systematically analyze common variants and CNVs. One of the first milestone papers in the field of psychiatric genomics was published for schizophrenia by O'Donovan et al. (2008). The success of GWAS was accelerated by the decreasing costs of genotyping resulting in an excess of GWAS-based publications. However, since common variants associated with complex disorders individually have small effect sizes it soon became clear that only through consortia and large-scale collaborations the sample sizes become large enough to reliably detect these effects (Psychiatric GWAS Consortium Steering Committee, 2009). Individual variants meeting stringent statistical criteria (genome-wide significance plus replication) that were found within these collaborative efforts still only account for a fraction of the estimated heritability of the disorders under study. Therefore, improved methods allowing a combined view at different genetic factors and better addressing the polygenic nature of psychiatric disorders were developed. A method suggested to aggregate these individual effects is the polygenic risk score (PRS) analysis that captures the effects of all variants below a certain threshold (Purcell et al., 2009). PRS have also been used to investigate the variants shared across disorders and successfully found genetic overlap not just between psychiatric disorders (Duncan et al., 2017; Tesli et al., 2014) but also with psychological phenotypes (Hatzimanolis et

al., 2015; Kauppi et al., 2015). LD score regression represents an efficient and powerful method not just to estimate the genetic correlation between traits but also to estimate the SNP-chip heritability without the need of individual-level data (Bulik-Sullivan, Finucane, et al., 2015; Bulik-Sullivan, Loh, et al., 2015). A different way to combine single variants is the gene-set analysis approach. Methods based on this approach rely on the assumption that the aggregation of SNPs within biologically meaningful sets of genes has greater statistical power to detect the polygenic architecture underlying psychiatric disorders than a single-SNP approach (Ramanan et al., 2012).

3 Methods

3.1 Genome-wide Association Analysis

GWAS were based on the hypothesis that common genetic variants (with low to moderate penetrance) explain some of the observed phenotypic variance for complex traits (Hirschhorn & Daly, 2005). GWAS do not require a biological hypothesis (such as candidate gene studies) and can be applied to discrete or quantitative phenotypes. A typical GWAS workflow comprises 5 steps: 1) SNP and sample quality control of raw data 2) Principal Component Analysis (PCA) 3) Imputation 4) GWAS and 5) Replication or meta-analysis. The following section briefly describes these steps.

First, since millions of SNPs undergo association testing in GWAS, rigorous quality control procedures need to be in place. It is crucial for further analysis to filter out SNPs and samples that do not meet standard quality control thresholds (Balding, 2006; Carvalho, Bengtsson, Speed, & Irizarry, 2007; Teo et al., 2007).

Second, the presence of systematic differences in allele frequencies in subgroups possibly due to different ancestry is called population stratification and represents one of the major confounding factors in GWAS (Lander & Schork, 1994). Owing to this, the genomic inflation factor λ is usually computed to assess whether the test statistics are inflated and need to be adjusted. Most often, population stratification is corrected for by excluding individuals based on their eigenvalues from PCA or by including principal components as covariates in the analysis model later on (L. Liu, Zhang, Liu, & Arendt, 2013; Price et al., 2006).

Third, statistical imputation of unobserved variants is an efficient way to improve comparability between different genotyping arrays and studies. It is facilitated by the fact that the genotypes of not directly genotyped variants can be estimated by the haplotypes inferred from directly genotyped SNPs and the haplotypes observed from a fully sequenced reference panel (Biernacka et al., 2009; Delaneau, Marchini, & Zagury, 2011; Howie, Donnelly, & Marchini, 2009; Howie, Fuchsberger, Stephens, Marchini, & Abecasis, 2012). Genetic Imputation also enables the fine-mapping of causal variants and has become a routine step in most GWAS pipelines.

Forth, linear regression is generally used to perform an analysis on quantitative traits and logistic regression on dichotomous traits. Even though other methods are

proposed for GWAS, these two are the most frequently used methods since they also allow to adjust for confounders. Pe'er, Yelensky, Altshuler, and Daly (2008) postulated that a p-value smaller than $5x10^{-8}$ should be considered genome-wide significant corresponding to the Bonferroni correction for 1'000'000 independent tests, representing the estimated number of common variants across the European genome.

Fifth, to control for false positive findings, the replication of GWAS findings in an independent sample or meta-analysis is strongly recommended. Replication studies need to consist of a sample size large enough to be able to detect the effect of the susceptibility allele. The replication sample needs to be independent but of the same population and the identical phenotype must be investigated (Chanock et al., 2007). It is vital to validate that the direction of effect of the associated allele is the same in both GWAS. However, Skol, Scott, Abecasis, and Boehnke (2006) reported that a joint analysis of the replication and the discovery sample together almost always has more power than the two-stage approach. Currently, also mega-analysis (a method to jointly analyze individual-level data from different studies) are performed, however, statistics have shown that meta-analysis are as efficient as a mega-analysis, with the benefits of having less privacy restrictions and logistical challenges since only summary statistics are shared between groups (Lin & Zeng, 2010).

3.2 Gene-Based Analysis

Testing only the associations of single SNPs has been demonstrated to be insufficient to dissect the complex genetic architecture of psychiatric disorders. The focus on genes rather than single SNPs as the unit of analysis has long been proposed and gene-based association scores are now reported more frequently (Hammerschlag et al., 2017; Kang, Jiang, & Cui, 2013; Neale & Sham, 2004). GWAS results often are difficult to replicate due to factors such as population differences, lack of power, allelic heterogeneity or diverse genotyping coverage (Hägg et al., 2015; Yang et al., 2012). In contrast, gene-based association analyses are suited to detect genes that may increase susceptibility to complex diseases since they are able to aggregate the cumulative effect of alleles within one gene and its regulatory region. Since gene association scores often are the basis for downstream analysis, it is prudent to carefully decide which method to choose. Several methods of how to compute a gene-based p-value have been proposed, mainly differing in their assumption of the underlying genetic architecture. A common and simple way is to choose the most significant SNP

to represent the association of the whole gene (Segrè et al., 2010). Albeit this is a fast method, it is most sensitive when only one SNP in a gene shows association and neglecting the additive effect of SNPs can lead to a loss of power (Ramanan et al., 2012). Alternatively, the calculation of a mean-based association p-value has been proposed, considering all SNPs within the gene boundaries. Still, this measurement can be biased by different aspects such as LD, coverage or gene size (de Leeuw et al., 2016). Other approaches allow the definition of a user-defined percentage of top associated SNPs to be used as a proxy for a gene association score consequently excluding SNPs diluting the summary statistics for a gene (Mishra & Macgregor, 2015). Novel methods also provide the opportunity to calculate multiple gene-based p-values and then aggregate them into a joint p-value which has the advantage of being more sensitive to different genetic architectures (de Leeuw et al., 2015).

The definition of gene boundaries and therefore the assignment of SNPs to genes represents an important decision since it may influence not only the power of the gene-based analysis but also follow-up analysis such as GSA. The regulatory effects of SNPs located outside a gene have been widely demonstrated however the inclusion of SNPs within regulatory regions also enables the inclusion of SNPs not relevant to the gene (Holmans, 2010; Maston, Evans, & Green, 2006). Definition of gene windows ranging from 0kb up to 500kb have been reported in various studies (The Network Pathway Analysis Subgroup of the Psychiatric Genomics Consortium, 2015; Veyrieras et al., 2008; K. Wang, Li, & Bucan, 2007). More recent approaches also allow the inclusion of distant SNPs in high LD with genic SNPs (Mishra & Macgregor, 2015). The inclusion of SNPs outside the gene boundaries increases the possibility of overlapping gene definitions and the number of SNPs that may be assigned to more than one gene, therefore potentially leading to an overinflated test statistic if not corrected for. This multiple-counting issue becomes even more momentous in the context of GSA, where one SNP can account for the strong association signals of several genes located within the same target gene-set (Dixson et al., 2014; Sedeño-Cortés & Pavlidis, 2014). A prominent region often excluded from analysis is the major histocompatibility complex (MHC) on chromosome 6 since it is known for its strong LD.

3.3 Gene-Set Analysis

The advantages of analyzing the cumulative effect of genes rather than single SNPs has long been discussed and GSA have become a requirement when publishing genome-wide association studies. Especially in polygenic traits, the assumption that SNPs underlying a disorder are enriched in genes constituting to a set of biologically meaningful genes has been widely accepted. Since the development of gene-set enrichment methods mainly for gene expression studies, the field has come a long way improving the methods not only in their purely computational burden but also the awareness and correction of confounding factors yielded in more reliable and statistically reproducible results (de Leeuw et al., 2016; Mooney & Wilmot, 2015). Nowadays, a broad range of GSA tools are freely available, however, despite some differences, the fundamental structures are highly comparable. The typical analysis of GSA consists of the following steps: 1) defining target gene-sets 2) formulating null hypothesis 3) mapping SNPs to genes 4) calculating gene association scores 5) calculating gene-set association scores 6) assessing gene-set significance.

3.3.1 Gene-Set Definition

In GSA, gene-sets are defined as a group of related genes that share a particular attribute, and the aim is to determine whether this attribute is associated with the phenotype of interest. Information on biological pathways and processes is available through a vast number of databases differing in e.g. curation-level, organisms included or functional areas covered. Reactome is an open access, peer-reviewed and well-curated database of biological pathways and processes which is extensively cross-referenced to other resources (Croft et al., 2014; Fabregat et al., 2018). Other frequently used open-source resources are the Gene Ontology (GO; (Ashburner et al., 2000; The Gene Ontology Consortium, 2017)) or the Kyoto Encyclopedia of Genes and Genomes databases (KEGG; (Kanehisa, Furumichi, Tanabe, Sato, & Morishima, 2017; Kanehisa & Goto, 2000; Kanehisa, Sato, Kawashima, Furumichi, & Tanabe, 2016)). Since GO is structured in a hierarchical way, it is necessary to account for the vastly overlapping gene-sets when systematically used for GSA. An easy way to download annotated gene-sets deriving from various sources is by accessing the Molecular Signatures Database (MSigDB, http://software.broadinstitute.org/gsea/msigdb/index.jsp). Ingenuity Knowledge Base

(Ingenuity Systems, Redwood City, CA, USA), a highly comprehensive and well-structured database, represents a well-curated but commercial alternative. However, researchers are not restricted to limit themselves to previously described gene-sets from public resources since the definition of customized target gene-sets such as genes known to contribute to a disorder or genes based on a cellular function are a promising tool to unravel biological mechanisms underlying diseases (Jansen et al., 2017; Nurnberger et al., 2014; Thapar et al., 2015).

3.3.2 Statistical structures of GSA

Even though a plethora of various gene-set analysis tools exist, only two different null hypothesis definitions are applied. The self-contained methods only consider genes in the target gene-set and test the association signal within this set against no signal. Whereas competitive methods consider all genes and test association of genes in the target gene-set against other genes not in this set. The competitive method may have less power to detect significant gene-sets when the genes associated with the trait are located in multiple gene-sets (Goeman & Bühlmann, 2007). Nevertheless, the competitive methods are better suited for the analysis of polygenic traits since they are able to robustly account for systematic inflation arising from various sources whereas self-contained methods are not (Devlin & Roeder, 1999; Moskvina, Craddock, Holmans, Owen, & O'Donovan, 2006; Price et al., 2006). Caution has to be exercised when interpreting the results gained from a self-contained GSA since this method is only able to show how strong the association for some of the genes in the gene-set are but not its relevance compared to other gene-sets (Mooney & Wilmot, 2015).

The second main difference between methods is the test statistic used for the computation of the gene-set association score where three main approaches can be distinguished. The simplest way is to apply a significance threshold to the gene-based p-values and count the number of genes designated as significant. A different but related approach is to rank the genes based on their p-values and then to check whether the target gene-set is enriched for highly ranked genes. Lastly a mean or sum-based approach can be used to summarize the gene-based p-values within the target gene-set (Mooney & Wilmot, 2015; Ramanan et al., 2012). Systematic comparisons between methods have shown that mean-based methods yield the greatest power

since the ranking and partitioning of genes results in a loss of information (de Leeuw et al., 2016).

Care should be taken when interpreting GSA results since several confounders have been reported to introduce significance even though no true relation exists. Statistically important factors to consider when applying GSA are the gene-set sizes, the size of genes itself and LD between them in the sense that large gene-sets consisting of large genes in high LD have the highest risk of becoming significant. The inflation of false-positives is especially strong if the most significant SNP was chosen to represent the association of the whole gene. Also, high LD between a truly disease-associated gene with genes not related to the disease can lead to a significant enrichment of a non-causative gene-set when these genes cluster within the same gene-set (de Leeuw et al., 2016; L. Wang et al., 2011).

3.3.3 GSA tools

The main characteristics and differences of the competitive tools used within the context of this thesis will be described briefly (see also Table 1). The improved gene set enrichment analysis for genome-wide association study version 2 (i-GSEA4GWASv2) is an easy-accessible, web-based resource based on the competitive GSEA algorithm developed by Subramanian et al. (2005) with the adaptation of permuting SNP labels instead of phenotype labels enabling the analysis of GWAS data. However, the method is not able to correct for the LD patterns between SNPs and therefore input data should only consist of a LD-independent set of SNPs (Zhang, Chang, Guo, & Wang, 2015). An additional representative of top-SNP methods is GSA-SNP with the advantage of using the kth best p-value within each gene instead of the traditionally used best SNP expected to result in fewer spurious association (Nam, Kim, Kim, & Kim, 2010). Meta-Analysis Gene-set Enrichment of variaNT Associations (MAGENTA), an additional top-SNP method based on Subramanian et al. (2005), has specifically been designed to analyze summary statistics derived from meta-analysis with the benefit of applying an elaborate correction algorithm to well-known confounders (Segrè et al., 2010). In contrast, INterval enRICHment analysis (INRICH) needs genotype data to create LD independent genomic intervals and is characterized by a sophisticated permutation scheme enabling the method to account for methodological biases, e.g. multiple counting of genes in high LD, effectively (P. H. Lee, O'Dushlaine, Thomas, & Purcell,

2012). Since this algorithm relies on independent intervals spread over the genome, the method can also be applied when only a subset of SNPs is tested provided these loci are independent and could have been distributed all over the genome.

Multi-marker Analysis of GenoMic Annotation (MAGMA) represents a powerful and highly flexible tool enabling the analysis of genes and gene-sets. It is based on a regression approach for both gene and gene-set analysis but also consists of more traditional approaches making the comparison with other tools fast and easy (de Leeuw et al., 2015). Even though MAGMA is a fairly new tool it has become the tool of choice for many researchers resulting in publications in various high-impact journals (Gandal et al., 2018; Howard et al., 2018; Pardiñas et al., 2018).

Table 1. Overview of GSA methods used

Method	**Input Data**	**Hypothesis tested**	**Test-statistic**	**Description**
i-GSEA4GWASv2	SNP p-values or gene p-values	Competitive	Rank-based	GSEA-based method with SNP label permutation
GSA-SNP	SNP p-values	Competitive	Rank-based	Uses kth best p-value as gene association score, offers multiple test statistics
MAGENTA	SNP p-values	Competitive	Count-based	GSEA-based method for meta-analytic data
INRICH	SNP p-values	Competitive (self-contained)	Count-based	Permutation-based method that uses LD independent genomic interval regions
MAGMA	Raw genotypes, SNP p-values or gene p-values	Competitive or self-contained	Mean-based	Regression-based, highly flexible and fast tool

The application of multiple methods to the same data has been recommended since the genetic architecture of complex disorders is not yet entirely clear and therefore the most appropriate method may not be determined a priori. Furthermore, the so-called technical replication has often been proposed for validation since most studies are not able to replicate their GSA findings in an independent sample (Gui, Li,

Sham, & Cherny, 2011; The Network Pathway Analysis Subgroup of the Psychiatric Genomics Consortium, 2015).

4 Original Research Papers

4.1 Identification of shared risk loci and pathways for bipolar disorder and schizophrenia

PLOS ONE

RESEARCH ARTICLE

Identification of shared risk loci and pathways for bipolar disorder and schizophrenia

Andreas J. Forstner[1,2,3,4,5☯], Julian Hecker[6☯], Andrea Hofmann[1,2,7], Anna Maaser[1,2], Céline S. Reinbold[4,5], Thomas W. Mühleisen[1,2,4,5,8], Markus Leber[9], Jana Strohmaier[10], Franziska Degenhardt[1,2], Jens Treutlein[10], Manuel Mattheisen[1,11,12], Johannes Schumacher[1,2], Fabian Streit[10], Sandra Meier[10,12,13], Stefan Herms[1,2,4,5], Per Hoffmann[1,2,4,5], André Lacour[14], Stephanie H. Witt[10], Andreas Reif[15], Bertram Müller-Myhsok[16,17,18], Susanne Lucae[16], Wolfgang Maier[19], Markus Schwarz[20], Helmut Vedder[20], Jutta Kammerer-Ciernioch[21], Andrea Pfennig[22], Michael Bauer[22], Martin Hautzinger[23], Susanne Moebus[24], Lorena M. Schenk[1,2], Sascha B. Fischer[4,5], Sugirthan Sivalingam[1,2], Piotr M. Czerski[25], Joanna Hauser[25], Jolanta Lissowska[26], Neonila Szeszenia-Dabrowska[27], Paul Brennan[28], James D. McKay[29], Adam Wright[30,31], Philip B. Mitchell[30,31], Janice M. Fullerton[32,33], Peter R. Schofield[32,33], Grant W. Montgomery[34], Sarah E. Medland[34], Scott D. Gordon[34], Nicholas G. Martin[34], Valery Krasnov[35], Alexander Chuchalin[36], Gulja Babadjanova[36], Galina Pantelejeva[37], Lilia I. Abramova[37], Alexander S. Tiganov[37], Alexey Polonikov[38,39], Elza Khusnutdinova[40,41], Martin Alda[42,43], Cristiana Cruceanu[44,45,46], Guy A. Rouleau[44], Gustavo Turecki[45,46,47], Catherine Laprise[48], Fabio Rivas[49], Fermin Mayoral[49], Manolis Kogevinas[50], Maria Grigoroiu-Serbanescu[51], Tim Becker[14,52], Thomas G. Schulze[53], Marcella Rietschel[10], Sven Cichon[1,2,4,5,8], Heide Fier[6‡], Markus M. Nöthen[1,2‡*]

1 Institute of Human Genetics, University of Bonn, Bonn, Germany, **2** Department of Genomics, Life & Brain Center, University of Bonn, Bonn, Germany, **3** Department of Psychiatry (UPK), University of Basel, Basel, Switzerland, **4** Human Genomics Research Group, Department of Biomedicine, University of Basel, Basel, Switzerland, **5** Institute of Medical Genetics and Pathology, University Hospital Basel, Basel, Switzerland, **6** Institute for Genomics Mathematics, University of Bonn, Bonn, Germany, **7** Institute of Medical Microbiology, Immunology and Parasitology, University of Bonn, Bonn, Germany, **8** Institute of Neuroscience and Medicine (INM-1), Research Center Juelich, Juelich, Germany, **9** Department of Psychiatry & Psychotherapy, University of Cologne, Cologne, Germany, **10** Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim/University of Heidelberg, Mannheim, Germany, **11** Department of Biomedicine and Centre for integrative Sequencing, iSEQ, Aarhus University, Aarhus, Denmark, **12** The Lundbeck Foundation Initiative for integrative Psychiatric Research, iPSYCH, Aarhus and Copenhagen, Denmark, **13** National Centre Register-Based Research, Aarhus University, Aarhus, Denmark, **14** German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany, **15** Department of Psychiatry, Psychosomatic Medicine and Psychotherapy, University Hospital Frankfurt am Main, Frankfurt am Main, Germany, **16** Max Planck Institute of Psychiatry, Munich, Germany, **17** Munich Cluster for Systems Neurology (SyNergy), Munich, Germany, **18** University of Liverpool, Institute of Translational Medicine, Liverpool, United Kingdom, **19** Department of Psychiatry, University of Bonn, Bonn, Germany, **20** Psychiatric Center Nordbaden, Wiesloch, Germany, **21** Center of Psychiatry Weinsberg, Weinsberg, Germany, **22** Department of Psychiatry and Psychotherapy, University Hospital Carl Gustav Carus, TU Dresden, Dresden, Germany, **23** Department of Psychology, Clinical Psychology and Psychotherapy, Eberhard Karls University Tübingen, Tübingen, Germany, **24** Institute of Medical Informatics, Biometry and Epidemiology, University Duisburg-Essen, Essen, Germany, **25** Laboratory of Psychiatric Genetics, Department of Psychiatry, Poznan University of Medical Sciences, Poznan, Poland, **26** Department of Cancer Epidemiology and Prevention, Maria Sklodowska-Curie Memorial Cancer Centre and Institute of Oncology Warsaw, Warsaw, Poland, **27** Department of Epidemiology, Nofer Institute of Occupational Medicine, Lodz, Poland, **28** Genetic Epidemiology Group, International Agency for Research on Cancer (IARC), Lyon, France, **29** Genetic Cancer Susceptibility Group, International Agency for Research on Cancer (IARC), Lyon, France, **30** School of Psychiatry, University of New South Wales, Randwick, Australia, **31** Black Dog Institute, Prince of Wales Hospital, Randwick, Australia, **32** Neuroscience Research Australia, Sydney, Australia, **33** School of Medical Sciences Faculty of Medicine, University of New South Wales, Sydney, Australia, **34** Queensland Institute of Medical Research (QIMR), Brisbane, Australia, **35** Moscow Research Institute of Psychiatry, Moscow, Russian Federation, **36** Institute of Pulmonology, Russian State Medical University, Moscow, Russian Federation, **37** Russian Academy of Medical Sciences, Mental Health Research Center, Moscow, Russian Federation, **38** Department of Biology, Medical Genetics

Citation: Forstner AJ, Hecker J, Hofmann A, Maaser A, Reinbold CS, Mühleisen TW, et al. (2017) Identification of shared risk loci and pathways for bipolar disorder and schizophrenia. PLoS ONE 12(2): e0171595. doi:10.1371/journal.pone.0171595

Editor: Consuelo Walss-Bass, University of Texas Health Science Center at Houston, UNITED STATES

Received: October 10, 2016

Accepted: January 23, 2017

Published: February 6, 2017

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: The study was supported by the German Federal Ministry of Education and Research (BMBF) through the Integrated Network IntegraMent (Integrated Understanding of Causes and Mechanisms in Mental Disorders), under the auspices of the e:Med Programme (grant 01ZX1314A to M.M.N. and S.C., grant 01ZX1314G to M.R.). M.M.N. is a member of the DFG-funded

Excellence-Cluster ImmunoSensation. M.M.N. also received support from the Alfried Krupp von Bohlen und Halbach-Stiftung. The study was supported by the German Research Foundation (DFG; grant FOR2107; RI908/11-1 to M.R.; NO246/10-1 to M.M.N.). The study was also supported by the Swiss National Science Foundation (SNSF, grant 156791 to S.C.). M.G.S. received grant no. 89/2012 from UEFISCDI, Romania. Canadian patients were genotyped within the ConLiGen project (www.ConLiGen.org), with the support of a grant from the DFG to M.R., M.B., and T.G.S. (RI 908/7-1). Controls for Germany II were drawn from the Heinz Nixdorf Recall Study (HNR) cohort, which was established with the support of the Heinz Nixdorf Foundation. Recruitment of the Australian sample was supported by an Australian NHMRC program grant (number 1037196). The recruitment of the Canadian patients was supported by a grant from the Canadian Institutes of Health Research #64410 to M.A. The study also used data generated by the GABRIEL consortium (controls for the sample Russia). Funding for the generation of these data was provided by the European Commission as part of GABRIEL contract number 018996 under the Integrated Program LSH-2004-1.2.5-1. Post genomic approaches to understand the molecular basis of asthma aiming at a preventive or therapeutic control and the Wellcome Trust under award 084703. Canadian controls were drawn from the French Canadian study (SLSJ), which was supported in part by the Canada Research Chair Environment and genetics of respiratory diseases and allergy, the Canadian Institutes of Health Research (Operating grant No. MOP-13506), and the Quebec Respiratory Network of the Fonds de rechercheen Santé du Québec (FRQS). Polish controls were recruited by the International Agency for Research on Cancer (IARC)/Centre National de Genotypage (CNG) GWAS Initiative. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

and Ecology, Kursk State Medical University, Kursk, Russian Federation, **39** Research Institute for Genetic and Molecular Epidemiology, Kursk State Medical University, Kursk, Russian Federation, **40** Institute of Biochemistry and Genetics, Ufa Scientific Center of Russian Academy of Sciences, Ufa, Russian Federation, **41** Department of Genetics and Fundamental Medicine of Bashkir State University, Ufa, Russian Federation, **42** Department of Psychiatry, Dalhousie University, Halifax, Canada, **43** National Institute of Mental Health, Klecany, Czech Republic, **44** Montreal Neurological Institute, McGill University, Montreal, Canada, **45** Department of Human Genetics, McGill University, Montreal, Canada, **46** McGill Group for Suicide Studies & Douglas Research Institute, Montreal, Canada, **47** Department of Psychiatry, McGill University, Montreal, Canada, **48** Département des sciences fondamentales, Université du Québec à Chicoutimi (UQAC), Saguenay, Canada, **49** Department of Psychiatry, Hospital Regional Universitario, Biomedical Institute of Malaga, Malaga, Spain, **50** Center for Research in Environmental Epidemiology (CREAL), Barcelona, Spain, **51** Biometric Psychiatric Genetics Research Unit, Alexandru Obregia Clinical Psychiatric Hospital, Bucharest, Romania, **52** Institute for Medical Biometry, Informatics and Epidemiology, University of Bonn, Bonn, Germany, **53** Institute of Psychiatric Phenomics and Genomics, Ludwig-Maximilians-University Munich, Munich, Germany

☯ These authors contributed equally to this work.
‡ These authors also contributed equally to this work.
* markus.noethen@uni-bonn.de

Abstract

Bipolar disorder (BD) is a highly heritable neuropsychiatric disease characterized by recurrent episodes of mania and depression. BD shows substantial clinical and genetic overlap with other psychiatric disorders, in particular schizophrenia (SCZ). The genes underlying this etiological overlap remain largely unknown. A recent SCZ genome wide association study (GWAS) by the Psychiatric Genomics Consortium identified 128 independent genome-wide significant single nucleotide polymorphisms (SNPs). The present study investigated whether these SCZ-associated SNPs also contribute to BD development through the performance of association testing in a large BD GWAS dataset (9747 patients, 14278 controls). After re-imputation and correction for sample overlap, 22 of 107 investigated SCZ SNPs showed nominal association with BD. The number of shared SCZ-BD SNPs was significantly higher than expected ($p = 1.46x10^{-8}$). This provides further evidence that SCZ-associated loci contribute to the development of BD. Two SNPs remained significant after Bonferroni correction. The most strongly associated SNP was located near *TRANK1*, which is a reported genome-wide significant risk gene for BD. Pathway analyses for all shared SCZ-BD SNPs revealed 25 nominally enriched gene-sets, which showed partial overlap in terms of the underlying genes. The enriched gene-sets included calcium- and glutamate signaling, neuropathic pain signaling in dorsal horn neurons, and calmodulin binding. The present data provide further insights into shared risk loci and disease-associated pathways for BD and SCZ. This may suggest new research directions for the treatment and prevention of these two major psychiatric disorders.

Introduction

Bipolar disorder (BD) is a severe neuropsychiatric disease characterized by recurrent episodes of mania and depression. BD has an estimated lifetime prevalence of around 1% [1], and a heritability of around 70% [2]. BD shows substantial clinical and genetic overlap with other

psychiatric disorders [3, 4]. An analysis of the genome-wide genotype data of the Psychiatric Genomics Consortium (PGC) revealed a 68% genetic correlation between BD and schizophrenia (SCZ), which was the highest correlation with BD of all psychiatric diseases investigated [3]. However, the genes involved in this etiological overlap remain largely unknown.

Although research into BD and SCZ has identified a number of susceptibility genes, the respective biological pathways still await identification. For BD, recent genome wide association studies (GWAS) have identified a number of risk loci [5–13].

For SCZ, a PGC meta-analysis of data from >36,000 patients and 113,000 controls identified 128 independent genome-wide significant single nucleotide polymorphisms (SNPs) in 108 genetic loci [14].

The aim of the present study was to investigate whether these 128 SCZ-associated SNPs also contribute to the development of BD. For this purpose, we performed association testing of these SNPs in our large BD GWAS dataset [12]. In addition, we analyzed whether the genome-wide significant BD-associated SNPs identified in our BD GWAS [12] show association with SCZ.

Materials and methods

Sample description

The analyses were performed using data from our previous GWAS of BD (9,747 patients and 14,278 controls) [12]. This GWAS dataset combined: (i) the MooDS data (collected from Canada, Australia, and four European countries); and (ii) the GWAS results for BD of the large multinational PGC [5]. The patients were assigned the following diagnoses (DSM-IV, DSM-IIR, Research Diagnostic Criteria): BD type 1 (n = 8,001; 82.1%); BD type 2 (n = 1,212; 12.4%); schizoaffective disorder (bipolar type; n = 269; 2.8%); and BD not otherwise specified (n = 265, 2.7%) [12]. The study was approved by the local ethics committees of the participating centers (University Hospital Würzburg, Germany; Central Institute of Mental Health, Mannheim, Germany; University of Essen, Germany; Ludwig Maximilians University, Munich, Germany; Prince of Wales Hospital, Sydney, Australia; Queensland Institute of Medical Research, Brisbane, Australia; Poznan University of Medical Sciences, Poland; University of Szczecin, Poland; speciality mood disorders clinics in Halifax and Ottawa, Canada; Russian State Medical University, Moscow, Russian Federation; Kursk State Medical University, Russian Federation; Regional University Hospital of Malaga, Spain; and Instituto Municipal de Asistencia Sanitaria, IMAS-IMIM, Barcelona, Spain) [12]. Written informed consent was obtained from all participants prior to inclusion [12].

Genome-wide significant loci for SCZ and BD

For the 128 linkage disequilibrium (LD)-independent genome-wide significant SNPs for SCZ, genetic information was obtained from the supplementary information of the SCZ GWAS of the PGC [14]. This is the largest GWAS of SCZ to date.

Genome-wide significant SNPs for BD were obtained from our BD GWAS [12].

Imputation and meta-analysis

Different reference panels were used for the imputation of the MooDS and PGC BD genotype data (1,000 Genomes Project, February 2012 release; and HapMap phase 2 CEU, respectively). Therefore, the summary statistics of the PGC BD GWAS [5] were imputed using the 1,000 Genomes Project reference panel and ImpG-Summary. The latter is a recently proposed method for the rapid and accurate imputation of summary statistics [15]. This resulted in

z-scores for >20 million SNPs. A total of 111 SCZ-associated SNPs could be mapped to the re-imputed PGC BD GWAS data. The remaining variants were either located on the X-chromosome (n = 3), or represented insertions or deletions (n = 14) which could not be imputed by the applied method. In total, 107 of the 111 SCZ-associated SNPs could be identified in the MooDS BD GWAS.

A meta-analysis for these 107 SNPs was then performed by combining the PGC BD GWAS and the MooDS BD GWAS, and using the sample size based strategy implemented in METAL [16].

Analysis of shared BD-SCZ SNPs

The risk alleles for all nominally significant SNPs in our BD GWAS [12] were compared to those reported in the PGC SCZ GWAS.

The SCZ discovery meta-analysis comprised data from 35,476 patients and 46,839 controls. Our BD GWAS comprised data from 9,747 patients and 14,278 controls [12]. To correct for an overlap between the two studies of around 500 patients and 9,200 controls [17, 18], we applied the framework of a bivariate normal distribution for the z-scores from both studies, corresponding to a specific SNP. Since the significant hits from a study were selected from different chromosomal regions, we assumed that the z-scores within a study are independent. According to the LD Score regression method [19], the mean inflation of the test statistics provides an approximation of the variance of the z-scores. By considering the set of SNPs in the HapMap3 reference panel [20], the calculated variance was approximately 1.82 for SCZ and 1.24 for BD. From equation (16) in Bulik-Sullivan et al. [19] (Supplementary Material), the covariance between z-scores was calculated to be 0.1644, under the assumption of no genetic correlation. This yielded a correlation of approximately 0.109. To confirm the validity of these theoretical calculations, we estimated the covariance of z-scores due to sample overlap by applying the LD Score regression software directly to the results of the PGC SCZ GWAS and our BD GWAS. After restriction to the well-imputed SNPs of HapMap3, the software estimated a covariance of 0.1707. This result provides further evidence that the degree of sample overlap was correctly estimated in the present study.

The z-scores for the 107 SCZ-associated SNPs were extracted from the PGC SCZ discovery study. The corresponding z-scores were extracted from our BD GWAS [12]. Using the values above, the mean and the variance of the normal distribution for the BD z-scores were determined, given the z-scores from the PGC SCZ discovery study. After the transformation of the initial z-scores from our BD GWAS, a total of 22 of 107 z-scores for BD had corresponding two-sided association p-values of <5% (Table 1).

Analogously, the z-scores for the genome-wide significant BD SNPs were extracted from our BD GWAS [12], and the corresponding z-scores were extracted from the PGC SCZ discovery study. Of the five BD-associated lead SNPs in our BD GWAS, one SNP (rs6550435) was in high LD (r^2 = 0.897, SNAP [21]) with a genome-wide SCZ-associated SNP (rs75968099), and was thus excluded from this additional analysis. For the remaining four SNPs, the transformation was computed in the other direction. After correction for sample overlap, no BD SNP showed association with SCZ.

Bonferroni correction for multiple testing was performed by multiplying the nominal p-values with the number of investigated SNPs (n = 107+4 = 111).

Pathway analysis

Pathway analysis for all 22 shared SCZ-BD SNPs was performed using Ingenuity Pathway Analysis (IPA; http://www.ingenuity.com/) [22, 23] and INRICH [24].

Table 1. Schizophrenia-associated SNPs with a p-value of <0.05 in our bipolar disorder GWAS data after correction for sample overlap.

SNP	Chr	Position	Alleles	P BD Meta	P_{corr} BD Meta	P PGC SCZ	Nearby Gene/s
rs75968099	3	36858583	T/C	2.03×10^{-5}	0.0022	1.05×10^{-13}	*TRANK1*
rs2535627	3	52845105	T/C	4.68×10^{-5}	0.0052	4.26×10^{-11}	*ITIH3-ITIH4*
rs6704641	2	200164252	A/G	0.0030	0.3331	8.33×10^{-9}	*SATB2*
rs140505938	1	150031490	T/C	0.0032	0.3597	4.49×10^{-10}	*VPS45*
rs7893279	10	18745105	T/G	0.0043	0.4770	1.97×10^{-12}	*CACNB2*
rs6704768	2	233592501	A/G	0.0063	0.6991	2.32×10^{-12}	*GIGYF2*
rs12704290	7	86427626	A/G	0.0075	0.8315	3.33×10^{-10}	*GRM3*
rs211829	7	110048893	T/C	0.0088	0.9778	3.71×10^{-8}	-
rs3735025	7	137074844	T/C	0.0098	>0.9999	3.28×10^{-9}	*DGKI*
rs324017	12	57487814	A/C	0.0098	>0.9999	2.13×10^{-8}	*NAB2*
rs2909457	2	162845855	A/G	0.0109	>0.9999	4.62×10^{-8}	*SLC4A10-DPP4*
rs9922678	16	9946319	A/G	0.0120	>0.9999	1.28×10^{-8}	*GRIN2A*
rs950169	15	84706461	T/C	0.0181	>0.9999	1.62×10^{-11}	*ADAMTSL3*
rs55661361	11	124613957	A/G	0.0301	>0.9999	2.8×10^{-12}	*NRGN*
rs10043984	5	137712121	T/C	0.0307	>0.9999	1.09×10^{-8}	*KDM3B*
rs1498232	1	30433951	T/C	0.0323	>0.9999	2.86×10^{-9}	*LOC101929406*
rs6434928	2	198304577	A/G	0.0351	>0.9999	2.06×10^{-11}	*SF3B1-COQ10B*
rs2007044	12	2344960	A/G	0.0367	>0.9999	3.22×10^{-18}	*CACNA1C*
rs8044995	16	68189340	A/G	0.0380	>0.9999	1.51×10^{-8}	*NFATC3*
rs56205728	15	40567237	A/G	0.0387	>0.9999	4.18×10^{-9}	*PAK6*
rs2693698	14	99719219	A/G	0.0429	>0.9999	4.8×10^{-9}	*BCL11B*
rs832187	3	63833050	T/C	0.0465	>0.9999	1.43×10^{-8}	*THOC7*

Single nucleotide polymorphisms (SNPs) are shown according to their p-values in our bipolar disorder (BD) GWAS [12] following correction for sample overlap. Chromosomal positions refer to genome build GRCh37 (hg19). Abbreviations: Chr, chromosome; P BD Meta, p-value in our BD GWAS [12] after correction for sample overlap; P_{corr} BD Meta, p-value in our BD GWAS [12] after correction for sample overlap and Bonferroni correction for multiple testing; P PGC SCZ, p-value in the PGC schizophrenia GWAS [14].

doi:10.1371/journal.pone.0171595.t001

In IPA, each gene is represented in a global molecular network, which is designed using information from the Ingenuity Pathway Knowledge Base. 'Networks' were generated algorithmically, and on the basis of their connectivity in terms of activation, expression, and transcription. Molecular relationships between genes are represented by connecting lines between nodes, as supported by published data stored in the Ingenuity Pathway Knowledge Base and/or PubMed. For the purposes of the present study, the canonical pathway analysis available in IPA was applied. Here, an SNP is mapped to a gene if it falls within the gene-coding region or within the 2 kilobase (kb) upstream/ 0.5 kb downstream range of the gene-coding region. This resulted in the inclusion of 13 genes in the pathway analysis. Significant pathways were filtered in order to achieve a minimum of two genes per set. The significance of the association between the SNP-associated genes mapped by IPA and the canonical pathway was measured using Fisher's exact test.

INRICH [24] was used as a secondary pathway analysis tool, as it enables examination of enriched association signals of LD-independent genomic intervals. Gene Ontology (GO) gene sets were extracted from the Molecular Signatures Database (MSigDB), version 5.0 (Broad Institute, http://software.broadinstitute.org/gsea/msigdb/index.jsp, downloaded in September 2015). The size of the extracted gene sets ranged from 10 to 200 genes, resulting in 1,268 target sets for testing. The intervals around the 22 SNPs of interest were based on empirical estimates of LD from PLINK (http://pngu.mgh.harvard.edu/purcell/plink/). SNPs were assigned to

genes using 50 kb up- and downstream windows. In total, 21 intervals were tested for the 1,268 target sets.

In IPA, correction for multiple testing was performed using the Benjamini Hochberg method. In INRICH, the empirical gene set p-value was corrected for multiple testing using bootstrapping-based re-sampling.

Results

A total of 107 of the 128 SCZ-associated SNPs could be mapped to both the re-imputed PGC BD GWAS and the MooDS BD GWAS data. A meta-analysis of these 107 SNPs was then performed using METAL [16].

After correction for sample overlap, 22 of the 107 SCZ-associated SNPs showed nominally significant p-values in our BD GWAS (Table 1, S1 Table). For all 22 SNPs, the direction of the effect was identical to that observed in the PGC SCZ GWAS [14]. Of the five genome-wide significant BD-associated SNPs identified in our BD GWAS, one SNP (rs6550435) was in high LD ($r^2 = 0.897$) with a genome-wide SCZ-associated SNP (rs75968099). None of the remaining four genome-wide significant BD-associated SNPs showed a nominally significant association with SCZ after correction for sample overlap (data not shown).

The number of SCZ SNPs with a p-value of <0.05 in our BD GWAS (n = 22) was significantly higher than expected ($p = 1.46x10^{-8}$, binomial test). This provides further evidence that SCZ-associated loci contribute to the development of BD.

The most strongly associated SNP was located near the gene *TRANK1* (Table 1, $p = 2.03x10^{-5}$), which is a reported genome-wide significant risk gene for BD [7, 12]. The other nominally associated SCZ-BD SNPs implicated loci which contain interesting candidate genes for BD and SCZ. These include the chromatin remodeling gene *SATB2*, the glutamate receptor genes *GRM3* and *GRIN2A*, and the calcium channel subunit gene *CACNB2*. The latter is a reported genome-wide significant risk gene for a number of psychiatric disorders, including BD and SCZ [17].

After Bonferroni correction for multiple testing, two SNPs (rs75968099, rs2535627) showed significant association with BD ($p_{corr} = 2.25x10^{-3}$ and $p_{corr} = 5.19x10^{-3}$, respectively).

Pathway analysis using IPA revealed nine pathways with nominally significant enrichment (Fig 1). Of these, eight remained significantly enriched after Benjamini Hochberg correction for multiple testing. The pathway with the strongest enrichment was synaptic long term potentiation ($p_{corr} = 0.003$, Fig 2, S2 Table). In addition, significant enrichment was found for glutamate receptor- and calcium signaling; neuropathic pain signaling in dorsal horn neurons; and CREB signaling in neurons.

These findings are consistent with previous pathway analyses of BD and SCZ [5, 25–27]. The present analysis also confirmed the glutamatergic signaling pathway, which was considered provisional in a recent review [28].

Pathway analysis using INRICH identified a total of 16 nominally significant gene-sets, which showed partial overlap in terms of the underlying genes. The enriched gene-sets include voltage-gated calcium channel complex/activity; calmodulin binding; glutamate receptor activity; and M phase of the mitotic cell cycle (Fig 3). None of these gene-sets remained significantly enriched for associations after correction for multiple testing (Fig 3, S3 Table).

Discussion

The present analyses revealed a significant enrichment of BD-associated SNPs within known SCZ-associated loci ($p = 1.46x10^{-8}$). This is consistent with previous reports of overlapping genetic susceptibility for BD and SCZ [4, 29, 30].

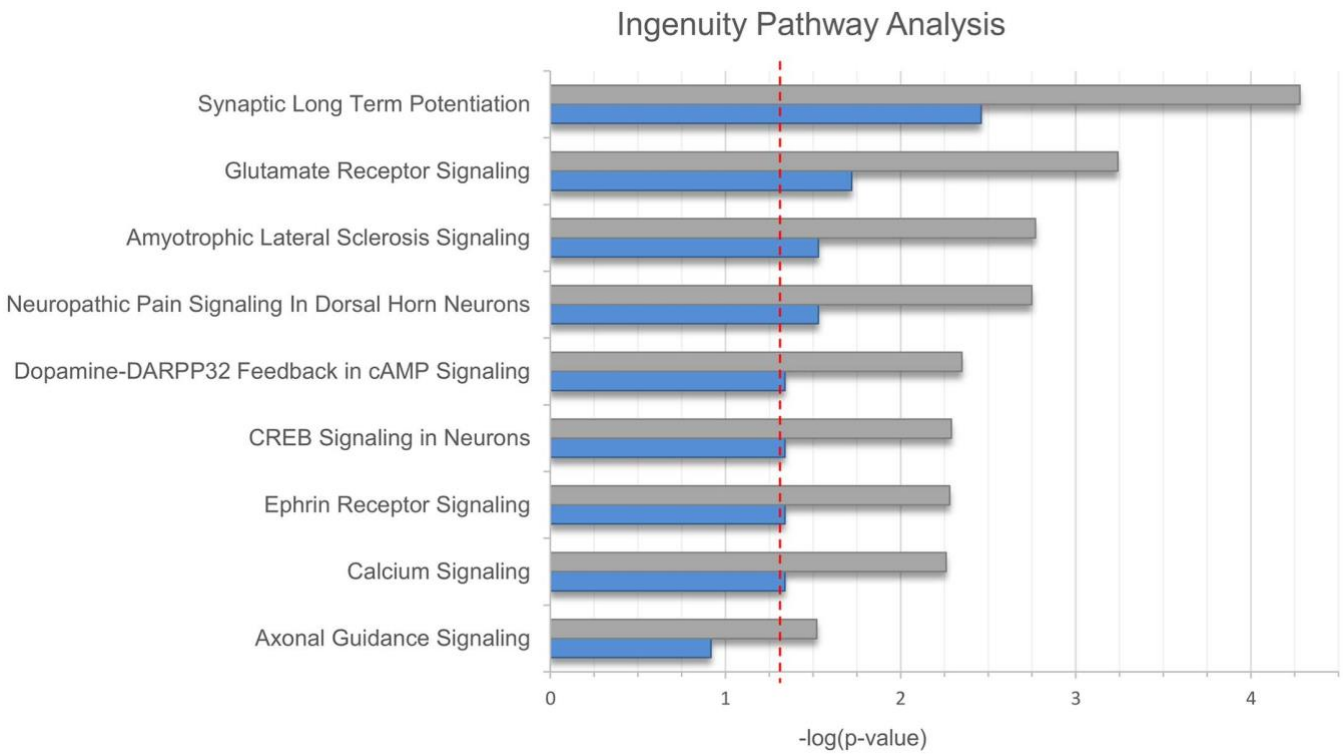


Fig 1. Results of the Ingenuity Pathway Analysis. Results of the Ingenuity Pathway Analysis (IPA) are shown in bar plot format. The x-axis shows negative logarithmic enrichment p-values for all associated pathways containing two and more genes prior to- (gray) and after- (blue) Benjamini Hochberg correction for multiple testing. The red horizontal line indicates a p-value of 0.05.

doi:10.1371/journal.pone.0171595.g001

The most strongly associated SNP was located near *TRANK1*, which is a reported genome-wide significant risk gene for BD [7]. The second SNP with significant BD association after correction for multiple testing (rs2535627, Table 1) was located in a genomic region on chromosome 3. This region contains multiple genes, including inter-alpha-trypsin inhibitor heavy chain 3 (*ITIH3*) and -4 (*ITIH4*). Common variation at the *ITIH3-ITIH4* region has been identified as a genome-wide significant risk factor for five different psychiatric disorders, including SCZ and BD [17].

Interestingly, the GWAS index SNP rs2535627 represents a Bonferroni-significant fetal brain methylation quantitative trait locus (mQTL), as it has been associated with DNA methylation at cg11645453. The latter is located in the 5' untranslated region of *ITIH4* [31]. This suggests that the SCZ-BD associated SNP rs2535627 might contribute to disease susceptibility by altering the expression of ITIH4 in the brain [32]. This hypothesis is supported by a recent study, which found that the G-allele of the SNP rs4687657—which is in moderate LD with rs2535627 (r^2 = 0.426, D' = 1.000, SNAP [21])—was significantly associated with reduced ITIH4 expression in the postmortem dorsolateral prefrontal cortex of controls [33].

SNPs with nominal association implicated several other plausible susceptibility genes for BD and SCZ (Table 1). These include *SATB2*, which is a highly conserved chromatin remodeling gene [34]. A previous animal study demonstrated that *SATB2* was an essential regulator of axonal connectivity in the developing neocortex [35]. In addition, mutations spanning *SATB2* have been reported in patients with neurodevelopmental disorders, including autism [36, 37].

The present SCZ-BD associated SNPs implicated three promising candidate genes for shared BD-SCZ etiology, i.e., *CACNB2*, *GRM3*, and *GRIN2A*. The gene *CACNB2* encodes an

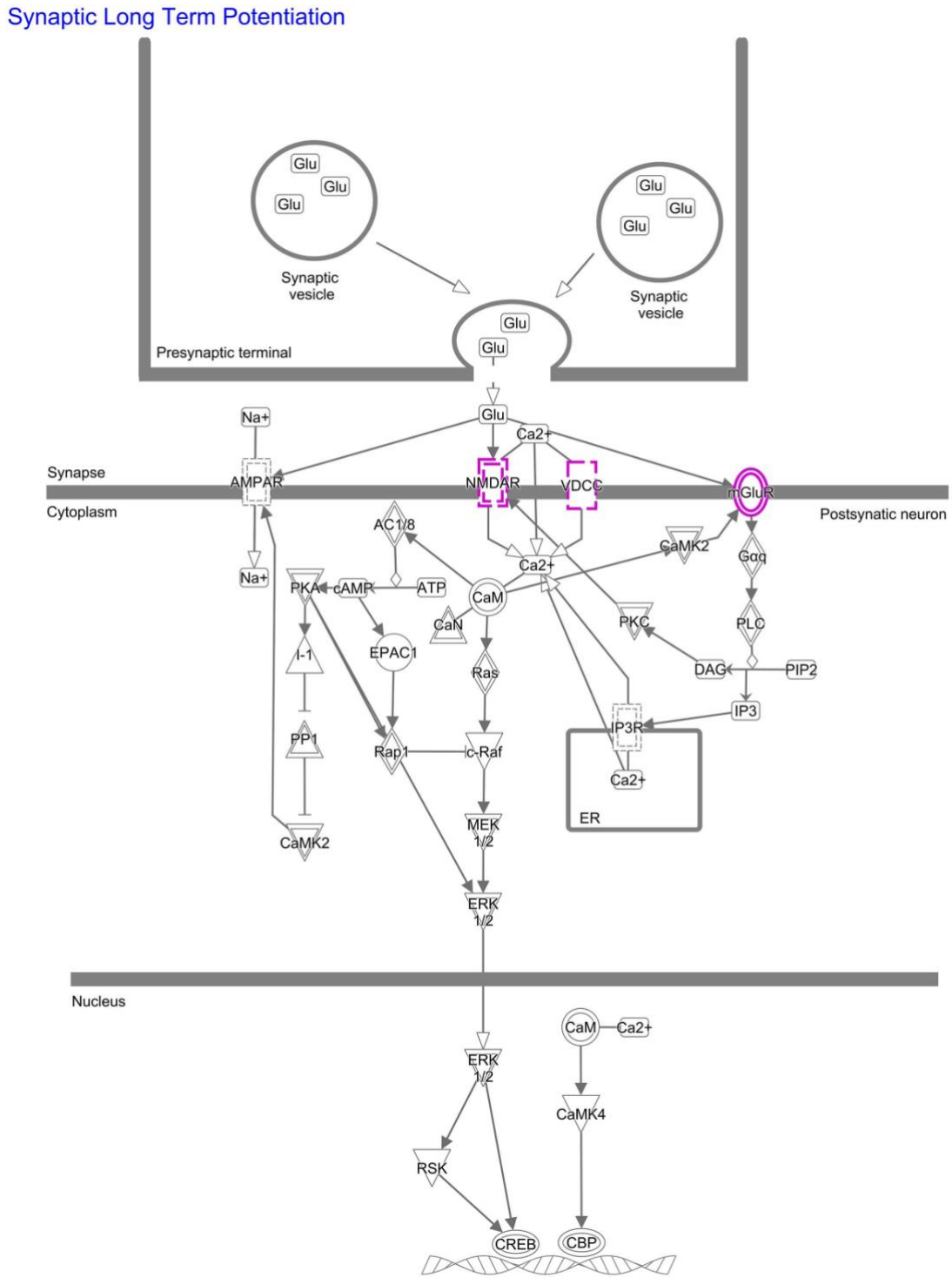


Fig 2. IPA pathway synaptic long term potentiation. Results of the Ingenuity pathway analysis (**I**PA) for the pathway "Synaptic Long Term Potentiation" are shown. Shared schizophrenia-bipolar disorder associated genes (*GRIN2A*, *GRM3*, *CACNA1C*) are highlighted in purple.

doi:10.1371/journal.pone.0171595.g002

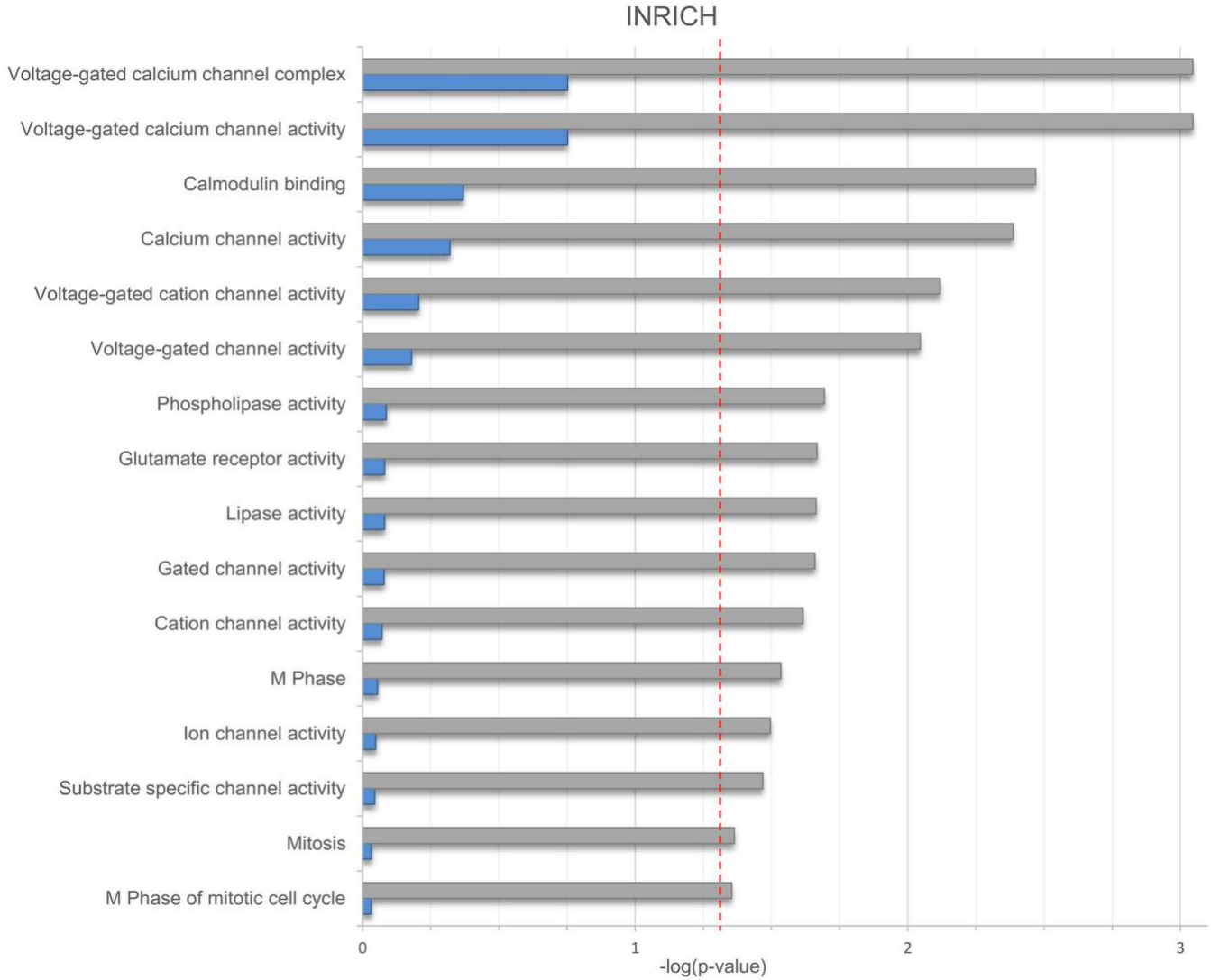


Fig 3. Results of the INRICH pathway analysis. Results of the INRICH pathway analysis are shown in bar plot format. The x-axis shows negative logarithmic enrichment p-values for all nominally associated pathways containing two and more genes prior to- (gray) and after- (blue) correction for multiple testing. The red horizontal line indicates a p-value of 0.05.

doi:10.1371/journal.pone.0171595.g003

L-type voltage-gated calcium channel subunit, and is a reported genome-wide significant risk gene for several psychiatric disorders, including SCZ and BD [17].

The gene *GRM3* encodes a metabotropic glutamate receptor. *GRM3* is expressed predominantly in astrocytes, and has been investigated by previous authors as a potential therapeutic target in SCZ [14]. A further SCZ-BD SNP was located near *GRIN2A*, which encodes an NMDA receptor subunit involved in glutamatergic neurotransmission and synaptic plasticity [14]. Interestingly, rare mutations in *GRIN2A* have been reported in patients with SCZ [38].

The present pathway analysis implicated calcium- and glutamate signaling, and neuropathic pain signaling in dorsal horn neurons. These findings are consistent with previous pathway analyses of BD and SCZ [5, 25–27]. These results thus provide further evidence that neurotransmitter signaling and synaptic processes are involved in the development of BD and SCZ.

Our enrichment analysis identified a total of 25 enriched gene-sets, which showed partial overlap in terms of the underlying genes. One of the major characteristics of the GO database is its hierarchical structure. This structure involves the use of broad 'parent' terms, which can be divided into more distinctive 'child' terms [39]. After taking these relations into account, we categorized our findings from the GO database into five different parent gene-set families: channel activity, lipase activity, mitotic cell cycle, calmodulin binding, and glutamate receptor signaling (S3 Table).

The results generated by IPA and INRICH were broadly consistent, despite the fact that the underlying databases were different. In some cases, pathways were implicated by the same genes, e.g., glutamate signaling was implicated by *GRIN2A* and *GRM3* in both IPA and INRICH. In other cases, pathways were implicated by differing genes, e.g., calcium channel activity/calcium signaling was implicated by *NFATC3* and *GRIN2A* in IPA, and by *CACNB2* and *CACNA1C* in INRICH (S2 and S3 Tables). This provides further support for the involvement of these pathways in the development of BD and SCZ.

The most strongly enriched pathway according to IPA was synaptic long term potentiation (Fig 2). This pathway has been implicated in learning and memory mechanisms [40]. Interestingly, several previous studies have provided evidence for the involvement of impaired long term potentiation in the pathophysiology of SCZ [41, 42]. In the present study, this pathway result was driven by the genes *GRIN2A*, *GRM3*, and *CACNA1C*. The products of all three genes are located in the postsynaptic membrane (Fig 2), which may suggest that dysfunction at the postsynaptic level is an early step in the development of BD and SCZ [43].

The identified pathways support specific hypotheses regarding the shared neurobiology of BD and SCZ. Notably, our results provide further evidence that glutamate signaling might be involved in the development of both SCZ and BD [44]. This would be consistent with the observation from routine clinical practice that SCZ drugs which target glutamate signaling are also effective in BD patients with psychosis or mania [44].

A limitation of the present study was the substantial sample overlap between our BD GWAS [12] and the SCZ GWAS of the PGC [14], since this creates an inflation of effect. To address this, the correlation of z-scores between the two studies was calculated. Based on this information, the initial z-scores were then transformed to correct for sample overlap. To estimate the correlation of test statistics, the publically available summary statistics of the PGC SCZ GWAS were used, which comprise the results of the discovery phase (35,476 patients, 46,839 controls). As the effect of shared samples might be stronger in the discovery sample than in the complete meta-analysis, we may have overestimated the correlation of test statistics between the two GWAS. Therefore our correction for sample overlap may have been too conservative. However, since the inflation effect introduced by shared samples might be different for independent SNPs compared to the average correlation of test statistics, we assume that our conservative approach was appropriate in terms of reducing false positive results. In future cross-disorder studies, shared samples should be identified and removed from one study on the basis of individual genotype data. This was not possible in the present study, as the analyses were based on summary statistics.

The present data provide further insights into shared risk loci and disease-associated pathways for BD and SCZ.

However, further research is required to determine precisely how the genetic risk variants correlate with particular diagnoses or clinical symptoms. For example, in a previous study, we showed that common variation at the *NCAN* locus was associated with both BD [8] and SCZ [45]. Genetic variation at the *NCAN* locus thus represents a cross-diagnosis contributory factor, which may relate to a specific mania symptom-complex [46]. Therefore, future studies are warranted to determine the specific BD and SCZ phenotypic dimensions to which the present

variants contribute. Such findings may suggest new research directions for the treatment and prevention of BD and SCZ.

Supporting information

S1 Table. Overview of the 107 investigated schizophrenia-associated SNPs and respective test statistics. Single nucleotide polymorphisms (SNPs) are shown according to their p-values in our bipolar disorder (BD) GWAS [12] following correction for sample overlap. Chromosomal positions refer to genome build GRCh37 (hg19). An imputation accuracy metric of 1 indicates that the respective SNP was not imputed using ImpG-Summary. Abbreviations: Chr, chromosome; A1, the allele to which the z-score is predicted; A2, other allele; Z/P BD Meta, z-score/p-value in our BD GWAS [12] after correction for sample overlap; Pcorr BD Meta, p-value in our BD GWAS [12] after correction for sample overlap and Bonferroni correction for multiple testing; Z/P PGC SCZ (discovery), derived z-score/p-value in the PGC schizophrenia GWAS (discovery phase) [14].
(XLSX)

S2 Table. Results of the Ingenuity Pathway Analysis. Enrichment p-values for all nine nominally associated pathways containing two and more genes are shown both prior to and after Benjamini Hochberg (B-H) correction for multiple testing. Abbreviation: No. Genes in Pathway, total number of genes in each pathway.
(DOCX)

S3 Table. Results of the INRICH pathway analysis. Empirical gene set p-values for all 16 nominally associated pathways containing two and more genes are shown. The p-values were corrected for multiple testing using bootstrapping-based re-sampling (corrected p-value). Abbreviations: GO, Gene Ontology; No. Genes in Pathway, total number of genes in each pathway.
(DOCX)

Acknowledgments

The authors are grateful to all patients and control subjects for their participation. We thank the Schizophrenia and Bipolar Disorder Working Groups of the Psychiatric Genomics Consortium (PGC) for providing access to the relevant data.

Author contributions

Conceptualization: AJF J. Hecker HF MMN.

Data curation: AJF J. Hecker HF MMN.

Formal analysis: AJF J. Hecker AH AM CSR LMS SBF SS MR SC HF MMN.

Funding acquisition: MB PBM NGM MA CL MG-S TGS MR SC MMN.

Investigation: AJF J. Hecker AH AM CSR TWM ML FD JT MM J. Schumacher FS S. Meier SH PH AL SHW TB HF.

Methodology: AJF J. Hecker HF MMN.

Project administration: AJF HF MMN.

Resources: J. Strohmaier FS AR BM-M SL WM MS HV JK-C A. Pfennig MB MH S. Moebus PMC J. Hauser JL NS-D PB JDM AW PBM JMF PRS GWM SEM SDG NGM VK AC GB GP LIA AST A. Polonikov EK MA CC GAR GT CL FR FM MK MG-S TGS MR MMN.

Software: J. Hecker AH AM CSR HF.

Supervision: AJF HF MMN.

Validation: AJF J. Hecker AM CSR HF MMN.

Visualization: AJF J. Hecker AH AM CSR HF MMN.

Writing – original draft: AJF J. Hecker AH CSR HF MMN.

Writing – review & editing: AC AH AJF AL AM A. Pfennig A. Polonikov AR AST AW BM-M CC CL CSR EK FD FR FM FS GAR GB GP GT GWM HF HV JDM J. Hecker J. Hauser JK-C JL JMF J. Strohmaier J. Schumacher JT LIA LMS MA MB MG-S MH MK ML MM MMN MR MS NGM NS-D PB PBM PH PMC PRS SBF SC SDG SEM SH SHW SL S. Meier S. Moebus SS TB TGS TWM VK WM.

References

1. Craddock N, Sklar P. Genetics of bipolar disorder. Lancet. 2013; 381(9878):1654–62. doi: 10.1016/S0140-6736(13)60855-7 PMID: 23663951
2. Nothen MM, Nieratschker V, Cichon S, Rietschel M. New findings in the genetics of major psychoses. Dialogues Clin Neurosci. 2010; 12(1):85–93. PMID: 20373670
3. Lee SH, Ripke S, Neale BM, Faraone SV, Purcell SM, Perlis RH, et al. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. Nat Genet. 2013; 45(9):984–94. doi: 10.1038/ng.2711 PMID: 23933821
4. Ruderfer DM, Fanous AH, Ripke S, McQuillin A, Amdur RL, Gejman PV, et al. Polygenic dissection of diagnosis and clinical dimensions of bipolar disorder and schizophrenia. Mol Psychiatry. 2014; 19 (9):1017–24. doi: 10.1038/mp.2013.138 PMID: 24280982
5. 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. doi: 10.1038/ng.943 PMID: 21926972
6. 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. doi: 10.1038/sj.mp.4002012 PMID: 17486107
7. Chen DT, Jiang X, Akula N, Shugart YY, Wendland JR, Steele CJ, 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. doi: 10.1038/mp.2011.157 PMID: 22182935
8. Cichon S, Muhleisen TW, Degenhardt FA, Mattheisen M, Miro 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. doi: 10.1016/j.ajhg.2011.01.017 PMID: 21353194
9. Ferreira MA, 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. doi: 10.1038/ng.209 PMID: 18711365
10. 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. doi: 10.1038/mp.2012.48 PMID: 22565781
11. 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. doi: 10.1038/mp.2012.142 PMID: 23070075
12. Muhleisen 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. doi: 10.1038/ncomms4339 PMID: 24618891
13. Sullivan PF, Daly MJ, O'Donovan M. Genetic architectures of psychiatric disorders: the emerging picture and its implications. Nat Rev Genet. 2012; 13(8):537–51. doi: 10.1038/nrg3240 PMID: 22777127

14. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. Nature. 2014; 511(7510):421–7. doi: 10.1038/nature13595 PMID: 25056061

15. Pasaniuc B, Zaitlen N, Shi H, Bhatia G, Gusev A, Pickrell J, et al. Fast and accurate imputation of summary statistics enhances evidence of functional enrichment. Bioinformatics. 2014; 30(20):2906–14. doi: 10.1093/bioinformatics/btu416 PMID: 24990607

16. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 2010; 26(17):2190–1. doi: 10.1093/bioinformatics/btq340 PMID: 20616382

17. Cross-Disorder Group of the Psychiatric Genomics Consortium. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. Lancet. 2013; 381(9875):1371–9. doi: 10.1016/S0140-6736(12)62129-1 PMID: 23453885

18. Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium. Genome-wide association study identifies five new schizophrenia loci. Nat Genet. 2011; 43(10):969–76. doi: 10.1038/ng.940 PMID: 21926974

19. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR, et al. An atlas of genetic correlations across human diseases and traits. Nat Genet. 2015; 47(11):1236–41. doi: 10.1038/ng.3406 PMID: 26414676

20. Altshuler DM, Gibbs RA, Peltonen L, Dermitzakis E, Schaffner SF, Yu F, et al. Integrating common and rare genetic variation in diverse human populations. Nature. 2010; 467(7311):52–8. doi: 10.1038/nature09298 PMID: 20811451

21. Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PI. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics. 2008; 24(24):2938–9. doi: 10.1093/bioinformatics/btn564 PMID: 18974171

22. Fillman SG, Cloonan N, Catts VS, Miller LC, Wong J, McCrossin T, et al. Increased inflammatory markers identified in the dorsolateral prefrontal cortex of individuals with schizophrenia. Mol Psychiatry. 2013; 18(2):206–14. doi: 10.1038/mp.2012.110 PMID: 22869038

23. Winham SJ, Cuellar-Barboza AB, McElroy SL, Oliveros A, Crow S, Colby CL, et al. Bipolar disorder with comorbid binge eating history: a genome-wide association study implicates APOB. J Affect Disord. 2014; 165:151–8. doi: 10.1016/j.jad.2014.04.026 PMID: 24882193

24. Lee PH, O'Dushlaine C, Thomas B, Purcell SM. INRICH: interval-based enrichment analysis for genome-wide association studies. Bioinformatics. 2012; 28(13):1797–9. doi: 10.1093/bioinformatics/bts191 PMID: 22513993

25. Nurnberger JI Jr., Koller DL, Jung J, Edenberg HJ, Foroud T, Guella I, et al. Identification of pathways for bipolar disorder: a meta-analysis. JAMA Psychiatry. 2014; 71(6):657–64. doi: 10.1001/jamapsychiatry.2014.176 PMID: 24718920

26. Ripke S, O'Dushlaine C, Chambert K, Moran JL, Kahler AK, Akterin S, et al. Genome-wide association analysis identifies 13 new risk loci for schizophrenia. Nat Genet. 2013; 45(10):1150–9. doi: 10.1038/ng.2742 PMID: 23974872

27. Xu W, Cohen-Woods S, Chen Q, Noor A, Knight J, Hosang G, et al. Genome-wide association study of bipolar disorder in Canadian and UK populations corroborates disease loci including SYNE1 and CSMD1. BMC Med Genet. 2014; 15:2. doi: 10.1186/1471-2350-15-2 PMID: 24387768

28. Harrison PJ. Molecular neurobiological clues to the pathogenesis of bipolar disorder. Curr Opin Neurobiol. 2016; 36:1–6. doi: 10.1016/j.conb.2015.07.002 PMID: 26210959

29. Lichtenstein P, Yip BH, Bjork C, Pawitan Y, Cannon TD, Sullivan PF, et al. Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. Lancet. 2009; 373(9659):234–9. doi: 10.1016/S0140-6736(09)60072-6 PMID: 19150704

30. 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. doi: 10.1038/nature08185 PMID: 19571811

31. Hannon E, Spiers H, Viana J, Pidsley R, Burrage J, Murphy TM, et al. Methylation QTLs in the developing brain and their enrichment in schizophrenia risk loci. Nat Neurosci. 2016; 19(1):48–54. doi: 10.1038/nn.4182 PMID: 26619357

32. Stranger BE, Nica AC, Forrest MS, Dimas A, Bird CP, Beazley C, et al. Population genomics of human gene expression. Nat Genet. 2007; 39(10):1217–24. doi: 10.1038/ng2142 PMID: 17873874

33. Ohi K, Shimada T, Nitta Y, Kihara H, Okubo H, Uehara T, et al. Schizophrenia risk variants in ITIH4 and CALN1 regulate gene expression in the dorsolateral prefrontal cortex. Psychiatr Genet. 2016; 26(3):142–3. doi: 10.1097/YPG.0000000000000127 PMID: 26991396

34. Gyorgy AB, Szemes M, de Juan Romero C, Tarabykin V, Agoston DV. SATB2 interacts with chromatin-remodeling molecules in differentiating cortical neurons. Eur J Neurosci. 2008; 27(4):865–73. doi: 10.1111/j.1460-9568.2008.06061.x PMID: 18333962

35. Alcamo EA, Chirivella L, Dautzenberg M, Dobreva G, Farinas I, Grosschedl R, et al. Satb2 regulates callosal projection neuron identity in the developing cerebral cortex. Neuron. 2008; 57(3):364–77. doi: 10.1016/j.neuron.2007.12.012 PMID: 18255030

36. Rosenfeld JA, Ballif BC, Lucas A, Spence EJ, Powell C, Aylsworth AS, et al. Small deletions of SATB2 cause some of the clinical features of the 2q33.1 microdeletion syndrome. PLoS One. 2009; 4(8): e6568. doi: 10.1371/journal.pone.0006568 PMID: 19668335

37. Talkowski ME, Rosenfeld JA, Blumenthal I, Pillalamarri V, Chiang C, Heilbut A, et al. Sequencing chromosomal abnormalities reveals neurodevelopmental loci that confer risk across diagnostic boundaries. Cell. 2012; 149(3):525–37. doi: 10.1016/j.cell.2012.03.028 PMID: 22521361

38. Tarabeux J, Kebir O, Gauthier J, Hamdan FF, Xiong L, Piton A, et al. Rare mutations in N-methyl-D-aspartate glutamate receptors in autism spectrum disorders and schizophrenia. Transl Psychiatry. 2011; 1:e55. doi: 10.1038/tp.2011.52 PMID: 22833210

39. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 2000; 25(1):25–9. doi: 10.1038/75556 PMID: 10802651

40. Shim SS, Hammonds MD, Tatsuoka C, Feng IJ. Effects of 4-weeks of treatment with lithium and olanzapine on long-term potentiation in hippocampal area CA1. Neurosci Lett. 2012; 524(1):5–9. doi: 10.1016/j.neulet.2012.06.047 PMID: 22750162

41. Goto Y, Yang CR, Otani S. Functional and dysfunctional synaptic plasticity in prefrontal cortex: roles in psychiatric disorders. Biol Psychiatry. 2010; 67(3):199–207. doi: 10.1016/j.biopsych.2009.08.026 PMID: 19833323

42. Salavati B, Rajji TK, Price R, Sun Y, Graff-Guerrero A, Daskalakis ZJ. Imaging-based neurochemistry in schizophrenia: a systematic review and implications for dysfunctional long-term potentiation. Schizophr Bull. 2015; 41(1):44–56. doi: 10.1093/schbul/sbu132 PMID: 25249654

43. Network and Pathway Analysis Subgroup of Psychiatric Genomics Consortium. Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. Nat Neurosci. 2015; 18(2):199–209. doi: 10.1038/nn.3922 PMID: 25599223

44. Cherlyn SY, Woon PS, Liu JJ, Ong WY, Tsai GC, Sim K. Genetic association studies of glutamate, GABA and related genes in schizophrenia and bipolar disorder: a decade of advance. Neurosci Biobehav Rev. 2010; 34(6):958–77. doi: 10.1016/j.neubiorev.2010.01.002 PMID: 20060416

45. Muhleisen TW, Mattheisen M, Strohmaier J, Degenhardt F, Priebe L, Schultz CC, et al. Association between schizophrenia and common variation in neurocan (NCAN), a genetic risk factor for bipolar disorder. Schizophr Res. 2012; 138(1):69–73. doi: 10.1016/j.schres.2012.03.007 PMID: 22497794

46. Miro X, Meier S, Dreisow ML, Frank J, Strohmaier J, Breuer R, et al. Studies in humans and mice implicate neurocan in the etiology of mania. Am J Psychiatry. 2012; 169(9):982–90. doi: 10.1176/appi.ajp.2012.11101585 PMID: 22952076

4.2 Genome-wide association study of borderline personality disorder reveals genetic overlap with bipolar disorder, major depression and schizophrenia

OPEN

Citation: Transl Psychiatry (2017) 7, e1155; doi:10.1038/tp.2017.115

www.nature.com/tp

ORIGINAL ARTICLE

Genome-wide association study of borderline personality disorder reveals genetic overlap with bipolar disorder, major depression and schizophrenia

SH Witt[1,123], F Streit[1,123], M Jungkunz[2,3], J Frank[1], S Awasthi[4], CS Reinbold[5], J Treutlein[1], F Degenhardt[6,7], AJ Forstner[5,6,7,8], S Heilmann-Heimbach[6], L Dietl[9], CE Schwarze[10], D Schendel[1], J Strohmaier[1], A Abdellaoui[11], R Adolfsson[12], TM Air[13], H Akil[14], M Alda[15], N Alliey-Rodriguez[16], OA Andreassen[17,18], G Babadjanova[19], NJ Bass[20], M Bauer[21], BT Baune[13], F Bellivier[22], S Bergen[23], A Bethell[24], JM Biernacka[25], DHR Blackwood[26], MP Boks[27], DI Boomsma[11], AD Børglum[28,29,30], M Borrmann-Hassenbach[31], P Brennan[32], M Budde[33,34], HN Buttenschøn[35], EM Byrne[36], P Cervantes[37], T-K Clarke[26], N Craddock[38], C Cruceanu[39], D Curtis[40,41], PM Czerski[42], U Dannlowski[43,44], T Davis[13], EJC de Geus[11], A Di Florio[45], S Djurovic[46,47], E Domenici[48], HJ Edenberg[49,50], B Etain[51], SB Fischer[5], L Forty[45], C Fraser[45], MA Frye[52], JM Fullerton[53,54], K Gade[33,34], ES Gershon[16], I Giegling[55], SD Gordon[56], K Gordon-Smith[57], HJ Grabe[58], EK Green[59], TA Greenwood[60], M Grigoroiu-Serbanescu[61], J Guzman-Parra[62], LS Hall[26,63], M Hamshere[38], J Hauser[42], M Hautzinger[64], U Heilbronner[34], S Herms[5,6,7], S Hitturlingappa[13], P Hoffmann[5,6,7], P Holmans[38], J-J Hottenga[11], S Jamain[51,65], I Jones[24], LA Jones[57], A Juréus[23], RS Kahn[66], J Kammerer-Ciernioch[67], G Kirov[38], S Kittel-Schneider[68], S Kloiber[69,70,71], SV Knott[57], M Kogevinas[72], M Landén[23,73], M Leber[74], M Leboyer[75], QS Li[76], J Lissowska[77], S Lucae[71], NG Martin[56,78], F Mayoral-Cleries[62], SL McElroy[79], AM McIntosh[26,80], JD McKay[81], A McQuillin[82], SE Medland[56], CM Middeldorp[11], Y Milaneschi[83], PB Mitchell[84,85], GW Montgomery[86], G Morken[87,88], O Mors[89,90], TW Mühleisen[91,92], B Müller-Myhsok[39,93,94], RM Myers[95], CM Nievergelt[60], JI Nurnberger[96], MC O'Donovan[97], LMO Loohuis[98], R Ophoff[99], L Oruc[100], MJ Owen[97], SA Paciga[101], BWJH Penninx[83], A Perry[57], A Pfennig[21], JB Potash[102], M Preisig[103], A Reif[68], F Rivas[62], GA Rouleau[104,105], PR Schofield[53,54], TG Schulze[1,33,34,106,107], M Schwarz[108], L Scott[109], GCB Sinnamon[110], EA Stahl[111,112], J Strauss[69], G Turecki[37], S Van der Auwera[58], H Vedder[108], JB Vincent[113], G Willemsen[11], CC Witt[114], NR Wray[36,115], HS Xi[116], Bipolar Disorders Working Group of the Psychiatric Genomics Consortium, Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium, Schizophrenia Working Group of the Psychiatric Genomics Consortium, A Tadic[117,118], N Dahmen[118], BH Schott[4,119], S Cichon[6,91,92,120], MM Nöthen[6,7], S Ripke[4,121,122], A Mobascher[118], D Rujescu[55], K Lieb[118], S Roepke[9], C Schmahl[2], M Bohus[3] and M Rietschel[1]

Borderline personality disorder (BOR) is determined by environmental and genetic factors, and characterized by affective instability and impulsivity, diagnostic symptoms also observed in manic phases of bipolar disorder (BIP). Up to 20% of BIP patients show comorbidity with BOR. This report describes the first case–control genome-wide association study (GWAS) of BOR, performed in one of the largest BOR patient samples worldwide. The focus of our analysis was (i) to detect genes and gene sets involved in BOR and (ii) to investigate the genetic overlap with BIP. As there is considerable genetic overlap between BIP, major depression (MDD) and schizophrenia (SCZ) and a high comorbidity of BOR and MDD, we also analyzed the genetic overlap of BOR with SCZ and MDD. GWAS, gene-based tests and gene-set analyses were performed in 998 BOR patients and 1545 controls. Linkage disequilibrium score regression was used to detect the genetic overlap between BOR and these disorders. Single marker analysis revealed no significant association after correction for multiple testing. Gene-based analysis yielded two significant genes: *DPYD* ($P = 4.42 \times 10^{-7}$) and *PKP4* ($P = 8.67 \times 10^{-7}$); and gene-set analysis yielded a significant finding for exocytosis (GO:0006887, $P_{FDR} = 0.019$; FDR, false discovery rate). Prior studies have implicated *DPYD*, *PKP4* and exocytosis in BIP and SCZ. The most notable finding of the present study was the genetic overlap of BOR with BIP ($r_g = 0.28$ [$P = 2.99 \times 10^{-3}$]), SCZ ($r_g = 0.34$ [$P = 4.37 \times 10^{-5}$]) and MDD ($r_g = 0.57$ [$P = 1.04 \times 10^{-3}$]). We believe our study is the first to demonstrate that BOR overlaps with BIP, MDD and SCZ on the genetic level. Whether this is confined to transdiagnostic clinical symptoms should be examined in future studies.

Translational Psychiatry (2017) **7,** e1155; doi:10.1038/tp.2017.115; published online 20 June 2017

INTRODUCTION

Borderline personality disorder (BOR; for the sake of readability, we have decided to use the rather unconventional abbreviation 'BOR' for Borderline Personality Disorder and the abbreviation 'BIP' for Bipolar Disorder) is a complex neuropsychiatric disorder with a lifetime prevalence of around 3%.[1] Untreated cases often have a chronic and severely debilitating clinical course.[1] BOR affects up to 20% of all psychiatric inpatients, and is associated with high health-care utilization. BOR therefore represents a substantial socio-economic burden.[2,3]

Correspondence: Dr SH Witt, Central Institute of Mental Health, Department of Genetic Epidemiology in Psychiatry, University of Heidelberg, 68159 Mannheim, Germany.
E-mail: Stephanie.Witt@zi-mannheim.de
[123]These authors contributed equally to this work.
A full list of affiliations appears at the end of the paper.
Received 16 February 2017; accepted 10 April 2017

BOR is characterized by affective instability, emotional dysregulation and poor interpersonal functioning.[3] Suicide rates in BOR range between 6 and 8%, and up to 90% of patients engage in non-suicidal self-injurious behavior.[4] Other prototypical features include high-risk behaviors and impulsive aggression. Current theories view dysfunctions in emotion processing, social interaction and impulsivity as core psychological mechanisms of BOR.[5]

To date, genetic research into BOR has been limited. Available genetic studies have involved small samples and focused on candidate genes, while no genome-wide association study (GWAS) of BOR patients has yet been performed.[6] However, Lubke *et al.*[7] conducted a GWAS of borderline personality features using data from three cohorts comprising $n = 5802$, $n = 1332$ and $n = 1301$ participants, respectively. Using the borderline subscale of the Personality Assessment Inventory (PAI-BOR), four borderline personality features (affect instability, identity problems, negative relations and self-harm) were assessed. The most promising signal in the combined analysis of two samples was for seven SNPs in the gene *SERINC5*, which encodes a protein involved in myelination. Two of the SNPs could be replicated in the third sample. Interestingly, here, the effect was highest for the affect instability items, that is, features that are key characteristics of manic phases of bipolar disorder (BIP).

Understanding of the pathogenesis of BOR remains limited. Both environmental and genetic factors are known to have a role in BOR etiology. Familial aggregation has been demonstrated,[8,9] and heritability estimates from twin studies range from 35 to 65%, with higher heritability estimates being obtained with self-ratings.[10–12]

The potential comorbidity between BOR and BIP is part of an ongoing debate. For example, Fornaro *et al.*[13] report substantial comorbidity of ~20% with BIP, whereas Tsanas *et al.*[14] find clear symptomatic differences between these two diagnostic groups. BOR displays an overlap of some symptoms with BIP, such as affective instability. In contrast, features such as dissociative symptoms, a feeling of chronic emptiness and identity disturbances are specific to BOR.[15] To date, no twin or family study has generated conclusive results concerning a genetic overlap between the two disorders.[16,17] However, a twin study[18] and a large-population-based study using polygenic risk score analyses[19] indicate a genetic overlap between borderline personality features and neuroticism, an established risk factor for BIP and other psychiatric disorders.[20]

To the best of our knowledge, the present study represents the first case–control GWAS in BOR, and was performed in one of the largest BOR patient samples worldwide. Given the limited heritability and the expected complex genetic architecture of BOR, the sample is too small to generate significant results for single markers. Instead, the main aim of the investigation was to detect (i) genes and gene sets with a potential involvement in BOR; and (ii) potential genetic overlap with BIP. As a substantial overlap of common risk variants exists between BIP and schizophrenia (SCZ), and to a lesser extent between BIP and major depressive disorder (MDD), and as there is also a high comorbidity of BOR and MDD, a further aim of the study was to determine whether any observed genetic overlap between BOR and BIP, MDD and SCZ was driven by disorder-specific genetic factors using linkage disequilibrium (LD)-score regression and polygenic risk scores (PRS).

MATERIALS AND METHODS

Participants

The present sample comprised 1075 BOR patients and 1675 controls.[21] All the participants provided written informed consent before inclusion. The study was approved by the respective local ethics committees.

The patients were recruited at the following German academic institutions: Department of Psychosomatic Medicine, Central Institute of Mental Health, Mannheim ($n = 350$); Department of Psychiatry and Psychotherapy, University Medical Center Mainz ($n = 231$); and the Department of Psychiatry, Charité, Campus Benjamin Franklin, Berlin ($n = 494$). Inclusion criteria for patients were: age 16 to 65 years; Central European ancestry; and a lifetime DSM-IV diagnosis of BOR. The control sample comprised 1583 unscreened blood donors from Mannheim, and 92 subjects recruited by the University Medical Center Mainz.

Clinical assessment

The diagnoses of BOR were assigned according to DSM-IV criteria and on the basis of structured clinical interviews. The diagnostic criteria for BOR were assessed using the German version of the IPDE[22] or the SKID-II.[23] All the diagnostic interviews were conducted by trained and experienced raters. BOR patients with a comorbid diagnosis of BIP or SCZ assessed with SKID-I[23] were excluded.

Genotyping

Automated genomic DNA extraction was performed using the chemagic Magnetic Separation Module I (Chemagen Biopolymer-Technologie, Baesweiler, Germany). Genotyping was performed using the Infinium PsychArray-24 Bead Chip (Illumina, San Diego, CA, USA).

Quality control and imputation

A detailed description of the quality control and imputation procedures is provided elsewhere.[24]

Briefly, quality control parameters for the exclusion of subjects and single-nucleotide polymorphisms (SNPs) were: subject missingness > 0.02; autosomal heterozygosity deviation ($|\text{Fhet}| > 0.2$); SNP missingness > 0.02; difference in SNP missingness between cases and controls > 0.02; and SNP Hardy–Weinberg equilibrium ($P < 10^{-6}$ in controls; $P < 10^{-10}$ in cases).

Genotype imputation was performed using the pre-phasing/imputation stepwise approach in IMPUTE2/SHAPEIT (default parameters and a chunk size of 3 Mb),[25,26] using the 1000 Genomes Project reference panel (release 'v3.macGT1').[27]

Relatedness testing and population structure analysis were performed using a SNP subset that fulfilled strict quality criteria (INFO > 0.8, missingness $< 1\%$, minor allele frequency > 0.05), and which had been subjected to LD pruning ($r^2 > 0.02$). This subset comprised 63 854 SNPs. In cryptically related subjects, one member of each pair ($\hat{\pi} > 0.2$) was removed at random following the preferential retention of cases over controls. Principal components (PCs) were estimated from genotype data (see Supplementary Figures 1–6), and phenotype association was tested using logistic regression. The impact of the PCs on genome-wide test statistics was assessed using λ.

Association analysis

Including the first four PCs as covariates, an additive logistic regression model was used to test single marker associations, as implemented in PLINK.[28] The *P*-value threshold for genome-wide significance was set at 5×10^{-8}.

Gene-based analysis

To determine whether genes harbored an excess of variants with small *P*-values, a gene-based test was performed with MAGMA Version 1.04 (http://ctg.cncr.nl/software/magma)[29] using genotyped markers only, filtered with a minor allele frequency $> 1\%$ ($n = 284\,220$). This test uses summary data and takes LD between variants into account. SNPs within ± 10 kb of the gene boundary were assigned to each gene. Obtained *P*-values were Bonferroni-corrected for the number of tested genes ($n = 17\,755$, $P = 2.8 \times 10^{-6}$).

Gene-set analysis

Gene-set-based analysis was implemented using genotyped markers only, filtered as above. As in the gene-based analysis, SNPs within ± 10 kb of the gene boundary were assigned to each gene. Gene-set analyses were carried out using Gene Ontology (GO, http://software.broadinstitute.org/gsea/msigdb/) terms.

The discovery gene-set-based analysis was carried out using i-GSEA4GWASv2 (http://gsea4gwas-v2.psych.ac.cn/).[30] The size of the gene sets was restricted to 20–200 genes, and the major histocompatibility complex region was excluded. In total, 674 gene sets were tested. The results were

adjusted for multiple testing using false discovery rate (FDR). To validate the significant finding, the respective gene set was investigated with (i) GSA-SNP, using the *P*-value of the second-best SNP in each gene (https://gsa.muldas.org)[31] and (ii) MAGMA using summary data and a nominal *P*-value threshold of $P < 0.05$.

LD-score regression

To investigate a possible genetic overlap between BOR and SCZ, BIP and MDD, LD-score regression was performed.[32] Genetic correlations between BOR and (i) BIP, (ii) SCZ and (iii) MDD were calculated[33] using the result files of the Psychiatric Genomics Consortium (PGC) meta-analyses for SCZ (33 640 cases and 43 456 controls),[34] BIP (20 352 cases and 31 358 controls)[35] and MDD (16 823 cases and 25 632 controls).[35] There was no overlap in cases or controls of the present BOR GWAS sample with the PGC samples.

Polygenic risk score

To determine the impact of polygenic risk on BOR and subgroups (that is, BOR with and without MDD), PRS were calculated for each subject based on the above-mentioned PGC data sets.

To obtain a highly informative SNP set with minimal statistical noise, the following were excluded: low frequency SNPs (minor allele frequency < 0.1); low-quality variants (imputation INFO < 0.9) and indels. Subsequently, these SNPs were clumped discarding markers within 500 kb of, and in high LD ($r^2 \geqslant 0.1$) with, another more significant marker. From the major histocompatibility complex region, only one variant with the strongest significance was retained. PRS were calculated as described elsewhere.[36] This involved *P*-value thresholds 5×10^{-8}, 1×10^{-6}, 1×10^{-4}, 0.001, 0.01, 0.05, 0.1, 0.2, 0.5 and 1.0, and multiplication of the natural logarithm of the odds ratio of each variant by the imputation probability for the risk allele. The resulting values were then totaled. For each subject, this resulted in one PRS for SCZ, MDD and BIP for each *P*-value threshold.

In a first step, the association of the PRS for BIP, SCZ and MDD with BOR case–control status was analyzed using standard logistic regression and by including the four PCs as covariates. For each *P*-value threshold, the proportion of variance explained (Nagelkerke's R^2) in BOR case–control status was computed by comparison of a full model (covariates+PRS) score to a reduced model (covariates only).

For further exploratory analysis, the $P < 0.05$ PRS for each disorder was selected (that is, including all markers that reached nominal significance in the training samples). To determine whether the different scores contribute independently to the case–control status, a regression including the PRS for MDD, SCZ and BIP and the four PCs was computed. In a secondary analysis, two further models were computed. These included the PRS for BIP and the PRS of either MDD or SCZ, while controlling for the four PCs.

Furthermore, PRS were analyzed by differentiating between controls, and patients with or without comorbid MDD. For each PRS, a linear model was computed using the PRS as a dependent variable, disease state as an independent variable and the four PCs as covariates. Differences between groups were assessed using *post hoc* tests (Bonferroni-corrected).

RESULTS

Sample characteristics

Genetic quality control led to the exclusion of 207 subjects. Reasons for exclusion were: (i) insufficient data quality (low call rate), $n = 6$; (ii) relatedness, $n = 63$; and (iii) population outlier status, $n = 138$. After quality control, the sample comprised 998 BOR cases (914 female/84 male) and 1545 controls (868 female/677 male). Mean age for cases was 29.58 years (range: 18–65 years, standard deviation (s.d. = 8.64)). Mean age for controls was 44.19 years (range: 18–72 years, s.d. = 13.24; details see Supplementary Table 1). Of the 998 cases, 666 had comorbid lifetime MDD, and 262 did not (data missing for 40 cases).

Single marker analysis

A total of 10 736 316 single markers were included in the analysis. As expected for GWAS on a complex psychiatric disorder with the current sample size, the single marker analysis revealed no significant hit after correction for multiple testing (see Figures 1 and 2). The most significant marker was rs113507694 in *DPPA3* on chromosome 12 ($P = 2.01 \times 10^{-07}$; odds ratio = 0.35, minor allele frequency = 0.03, INFO = 0.59). Single markers with $P < 1 \times 10^{-5}$ are listed in Supplementary Table 2.

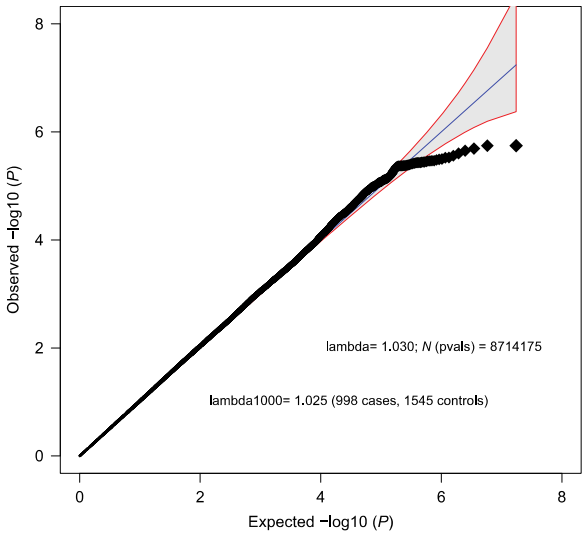


Figure 1. Quantile–Quantile plot. Quantile–Quantile plot of the case–control analysis (998 cases; 1545 controls) showing expected and observed −log10 *P*-values. The shaded region indicates the 95% confidence interval of expected *P*-values under the null hypothesis.

Gene-based analysis

In the gene-based analysis, a total of 17 755 genes were tested. Two genes showed significant association with BOR after correction for multiple testing: the gene coding for Plakophilin-4 on chromosome 2 (*PKP4*; $P = 8.24 \times 10^{-7}$); and the gene coding for dihydropyrimidine dehydrogenase on chromosome 1 (*DPYD*, $P = 1.20 \times 10^{-6}$). The most significant genes ($P < 5 \times 10^{-4}$) are listed in Table 1. The top hit of the previous GWAS of borderline personality features, *SERINC5,* achieved nominal significance in the present study ($P_{uncorrected} = 0.016$).

Gene-set analysis

Gene-set analysis with i-GSEA4GWASv2 revealed one significant gene set: exocytosis (GO: 0006887; $P_{FDR} = 0.019$). Of 25 genes in this gene set, 22 were mapped with variants and 15 showed nominally significant associations. Details on significant and nonsignificant genes in this gene set are provided in Supplementary Table 3. All gene sets with $P_{uncorrected} < 0.01$ are shown in Table 2. A technical replication analysis with GSA-SNP and MAGMA confirmed the gene-set exocytosis (GSA-SNP: $P_{uncorrected} = 2.32 \times 10^{-4}$; MAGMA: $P_{uncorrected} = 0.056$).

LD-score regression

Significant genetic correlations with BOR were found for BIP ($r_g = 0.28$; s.e. = 0.094; $P = 2.99 \times 10^{-3}$), MDD ($r_g = 0.57$; s.e. = 0.18; $P = 1.04 \times 10^{-3}$) and SCZ ($r_g = 0.34$; s.e. = 0.082; $P = 4.37 \times 10^{-5}$). A meta-analytic comparison revealed no significant differences between the correlations (all $P > 0.13$).

Polygenic risk score

PRS analysis revealed significant associations with BOR for the PRS of BIP, MDD and SCZ. SCZ PRS were significant for all investigated thresholds. BIP and MDD scores were significant for all PRS that

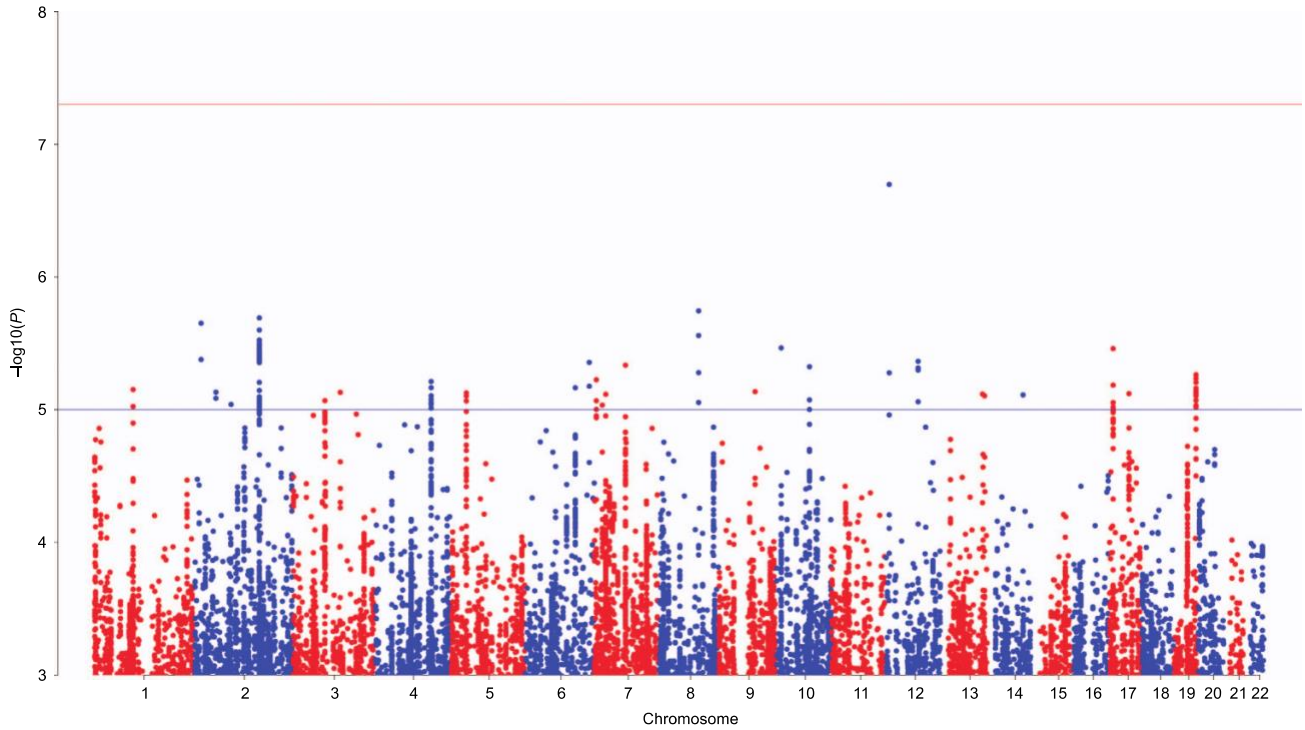


Figure 2. Manhattan plot showing association results. Manhattan plot of the case–control analysis (998 cases; 1545 controls). For each single-nucleotide polymorphism (SNP), the chromosomal position is shown on the *x* axis, and the –log10 *P*-value on the *y* axis. The red line indicates genome-wide significance ($P < 5 \times 10^{-8}$) and the blue line indicates suggestive evidence for association ($P < 1 \times 10^{-5}$).

Table 1. Results of the gene-based analysis using MAGMA

GENE	*CHR*	*START*	*STOP*	N_{SNPS}	N_{PARAM}	Z_{STAT}	P
PKP4	2	159303476	159547941	21	13	4.7924	8.24×10^{-7}
DPYD	1	97533299	98396615	105	68	4.7162	1.20×10^{-6}
GRAMD1B	11	123315191	123508478	34	28	3.8856	5.10×10^{-5}
STX8	17	9143788	9489275	38	33	3.7984	7.28×10^{-5}
BMP2	20	6738745	6770910	7	6	3.588	1.67×10^{-4}
TRAF3IP1	2	239219185	239319541	11	8	3.5389	2.01×10^{-4}
ZP3	7	76016841	76081388	9	7	3.5037	2.29×10^{-4}
PINX1	8	10612473	10707394	19	11	3.5034	2.30×10^{-4}
GTF3C4	9	135535728	135575471	4	4	3.4851	2.46×10^{-4}
DNAH1	3	52340335	52444513	11	8	3.4543	2.76×10^{-4}
YKT6	7	44230577	44263893	6	3	3.3841	3.57×10^{-4}
CCSER1	4	91038684	92533370	111	78	3.3804	3.62×10^{-4}
LRRC59	17	48448594	48484914	8	6	3.3716	3.74×10^{-4}
TMEM71	8	133712191	133782914	9	8	3.3668	3.80×10^{-4}
BAP1	3	52425020	52454121	3	3	3.345	4.11×10^{-4}
AQR	15	35138552	35271995	8	6	3.3299	4.34×10^{-4}
FGFR1	8	38258656	38336352	12	10	3.3162	4.56×10^{-4}

Abbreviations: CHR, chromosome; N_{PARAM}, number of parameters used in the model; N_{SNPS}, number of single-nucleotide polymorphisms; *P*, *P*-value of gene; Z_{STAT}, *z*-value of the gene. Most significant genes ($P < 5 \times 10^{-4}$) in the gene-based analysis and their chromosomal position. Genes in bold font were significant after correction for multiple testing.

included SNPs with *P*-values higher than 0.0001 and 0.001, respectively (see Supplementary Table 4). The share of variance explained in BOR case–control status (Nagelkerke's R^2) by the respective PRS was up to 0.86% for BIP; up to 3.1% for SCZ; and up to 2.1% for MDD (see Figure 3 and Supplementary Table 4).

Simultaneous addition of the PRS for SCZ, BIP and MDD (threshold $P < 0.05$) to the regression model explained 4.4% of the variance (Nagelkerke's R^2) in BOR case–control status. The PRS for SCZ and the PRS for MDD were significant predictors ($P = 9.78 \times 10^{-9}$ and $P = 1.9 \times 10^{-7}$, respectively). The PRS for BIP was not a significant predictor in this model ($P = 0.28$).

A secondary analysis was then performed including (i) BIP PRS with MDD PRS and (ii) BIP PRS with SCZ PRS. Here, BIP PRS explained variance independently of MDD PRS ($P = 0.0067$), but not of SCZ PRS ($P = 0.11$).

Differentiation between cases with and without comorbid MDD and controls revealed significant effects of BOR diagnosis on PRS for BIP, SCZ and MDD (all $P < 0.001$, see Figure 4). *Post hoc* analyses revealed no differences in PRS for the BIP, SCZ or MDD PRS of the BOR subgroup with comorbid MDD compared with the BOR subgroup without MDD (all $P > 0.5$).

Compared with controls, PRS for SCZ and MDD were significantly increased in the BOR subgroups with and without comorbid MDD (all $P < 0.001$). The PRS for BIP only showed a significant difference to controls in the BOR subgroup with comorbid MDD ($P < 0.001$, see Figure 4).

DISCUSSION

The present study is the first case–control GWAS of BOR. As expected, no genome-wide significant association was found for any single marker. In the gene-based test, however, two genes achieved genome-wide significance: dihydropyrimidine dehydrogenase (*DPYD*) and Plakophilin-4 (*PKP4*). *DPYD* encodes a pyrimidine catabolic enzyme, which is the initial and rate-limiting factor in the pathway of uracil and thymidine catabolism. Genetic deficiency of this enzyme results in an error in pyrimidine metabolism.[37] This is associated with thymine–uraciluria and an increased risk of toxicity in cancer patients receiving 5-fluorouracil chemotherapy (http://www.ncbi.nlm.nih.gov/gene/1806). Recent PGC meta-analyses revealed an association between *DPYD* and SCZ and BIP.[34,38,39] *DPYD* contains a binding site for the micro-RNA miR-137, which has previously been associated with schizophrenia,[40] and a previous exome-sequencing study reported two putative functional *de novo* variants in *DPYD* in cases with SCZ.[41] *PKP4* is involved in the regulation of cell adhesion and cytoskeletal organization.[42] In pathway analyses of PGC GWAS data, cell adhesion was associated with BIP,[43] and SCZ,[44] whereas cell junction was implicated in MDD, as well as in an integrative pathway analysis of all three disorders.[45]

SERINC5, which was the top hit of the previous GWAS of Borderline personality features,[7] achieved nominal significance in the present study. The protein SERINC5 incorporates serine into newly forming membrane lipids, and is enriched in myelin in the brain.[46] Previous research suggests that decreased myelination is associated with a reduced capacity for social interaction.[7,47]

The gene-set analyses yielded significant results for exocytosis. In neuronal synapses, exocytosis is triggered by an influx of calcium and critically underlies synaptic signaling. Dysregulated neuronal signaling and exocytosis are core features of neurodevelopmental psychiatric disorders such as the autism spectrum disorders and intellectual disability.[48,49] Moreover, recent findings from large meta-analyses have implicated dysregulated neuronal signaling and exocytosis in the molecular mechanisms of BIP, SCZ and MDD.[48,50,51] These processes may now represent promising starting points for further research into BOR.

The most interesting finding of this study is that BOR showed a genetic overlap with BIP, SCZ and MDD. Notably, BIP did not show a higher correlation with BOR ($r_g = 0.28$) than SCZ ($r_g = 0.34$) or MDD ($r_g = 0.57$). In view of the present sample size, these values must be viewed with caution. A more accurate estimation of these correlations will require calculations in larger cohorts.

Although comorbid BIP was excluded in the present BOR patients, the possibility that the observed genetic overlap between BOR and BIP was at least partly attributable to misdiagnosis cannot be excluded. However, an alternative explanation appears more likely, that is, that disorders currently categorized as BOR and BIP share a common genetic background,

Table 2. Results of the gene-set analysis

Gene-set name	*Number of genes*	P-*value*	*FDR* P-*value*
GO: EXOCYTOSIS	**25**	**0.001**	**0.019**
GO: RESPONSE TO ORGANIC SUBSTANCE	30	0.002	0.173
GO: BRAIN DEVELOPMENT	51	0.003	0.888
GO: HORMONE METABOLIC PROCESS	30	0.003	0.511
GO: PROTEIN C TERMINUS BINDING	73	0.003	0.536
GO: LYSOSOME	53	0.007	0.785
GO: LYTIC VACUOLE	53	0.007	0.785
GO: MULTI-ORGANISM PROCESS	143	0.007	0.920

Abbreviations: FDR, false discovery rate; GO, Gene Ontology; *P*-value, gene-set *P*-value. Most significant gene sets (uncorrected $P < 0.01$) in the gene-set analysis with i-GSEA4GWASv2 are listed. Gene sets in bold font were significant after correction for multiple testing.

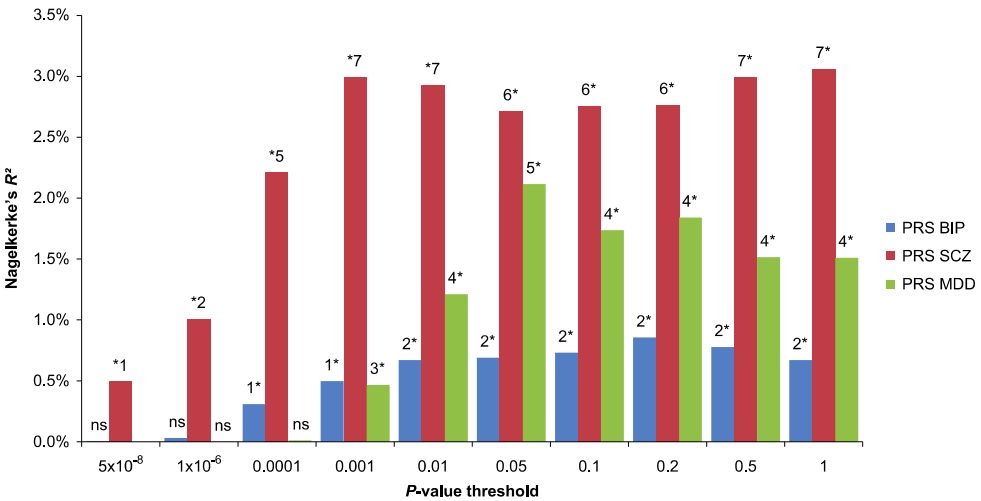


Figure 3. Polygenic risk score analysis. The proportion of variance explained in case–control status (*y* axis; Nagelkerke's R^2) by the PRS for BIP, SCZ and MDD is depicted for the different *P*-value cutoffs used in the calculation of the PRS. Principal components were included in the models to control for population stratification. 1*, $P < 0.05$; 2*, $P < 0.001$; 3*, $P < 1 \times 10^{-4}$; 4*, $P < 1 \times 10^{-6}$; 5*, $P < 1 \times 10^{-8}$; 6*, $P < 1 \times 10^{-10}$; 7*, $P < 1 \times 10^{-12}$. BIP, bipolar disorder; MDD, major depressive disorder; NS, nonsignificant; PRS, polygenic risk score; SCZ, schizophrenia.

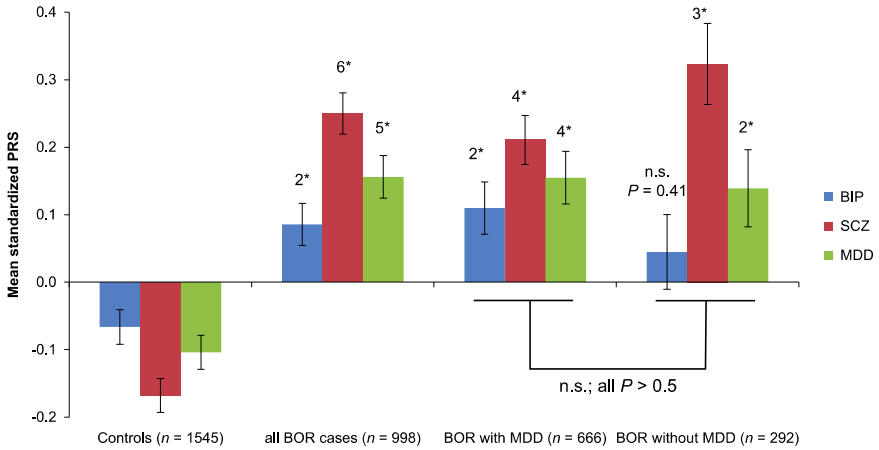


Figure 4. Polygenic risk score analysis in subgroups. Mean z-standardized PRS and standard error (s.e.) for BIP, SCZ and MDD are shown in the control group, all cases, and in cases with and without comorbid MDD. PRS with a *P*-value threshold of $P = 0.05$ were selected for this comparison and principal components were included in the models to control for population stratification. The numbers at the top of each bar indicate the significance of the difference in the respective PRS in comparison with the control group. 1*, $P < 0.05$; 2*, $P < 0.001$; 3*, $P < 1 \times 10^{-4}$; 4*, $P < 1 \times 10^{-6}$; 5*, $P < 1 \times 10^{-8}$; 6*, $P < 1 \times 10^{-10}$; 7*, $P < 1 \times 10^{-12}$. BIP, bipolar disorder; BOR, borderline personality disorder; MDD, major depressive disorder; NS, nonsignificant; PRS, polygenic risk score; SCZ, schizophrenia.

and they also do so with SCZ and MDD. This hypothesis is supported by the present observation of a genetic overlap between BOR and SCZ, two disorders that are rarely misdiagnosed by psychiatrists, despite the presence of common psychotic symptoms.

An explanation could also be that the genetic commonality between BOR and BIP, SCZ, and MDD might be due to a common effect of MDD. Prior to the introduction of DSM-IV, a history of MDD was required for a diagnosis of BIP, and MDD has a high prevalence in patients with SCZ (25-85%).[52,53] Therefore, the MDD genetic risk variants that are common to BOR, BIP, and SCZ may be responsible for the observed overlap. For this reason, we conducted two further analyses. First, we compared PRS of BIP, SCZ and MDD in subsamples of BOR patients with (~60%) and without comorbid MDD. Here, no differences in any of the PRS were found. Second, we performed a joint analysis of PRS of BIP, SCZ and MDD in a logistic regression analysis in BOR patients vs controls. Here, no differences were found in any of the PRS. Second, we performed a joint analysis of the PRS of BIP, SCZ and MDD in a logistic regression analysis in BOR patients vs controls. Here, both the SCZ and the MDD risk score explained variance in BOR case–control status independently. Secondary analysis revealed that the BIP risk score explained variance independently of the MDD risk score but not of the SCZ risk score. These results indicate that comorbidity with MDD does not explain the genetic overlap between BOR and BIP, SCZ and MDD. However, the training sets differ in terms of their power to detect underlying risk variants, and therefore the derived PRS differ in terms of the variance they can explain.

It must be noted, that in the PGC-BIP, -SCZ and -MDD samples, controls are partly overlapping. However, it is unlikely that this drives the genetic correlation of BOR with those disorders as the overlap of controls in these samples is rather small (under 10%).[54] Also, the joint logistic regression analysis demonstrated that polygenic risk for SCZ and MDD contributed independently to the BOR risk (see above).

The present study had several limitations. First, despite being one of the largest BOR samples available worldwide, the sample size was small in terms of the estimation of heritability. Replication of the present results is warranted in larger, independent cohorts. This should include the investigation of non-European samples. Second, no information was available on the presence of common clinical features such as psychotic symptoms and affect instability. This precluded detailed analysis of the identified genetic overlap. Future studies in larger cohorts should also investigate more detailed phenotypes, including comorbid axis I and axis II disorders, such as addiction and personality disorders, respectively. Third, the observation that psychiatric patients often establish non-random relationships with persons affected by the same or another psychiatric disorder,[55] and therefore have offspring with a higher genetic risk for psychiatric disorders, might contribute to the observed genetic correlation of BOR with BIP, SCZ and MDD. However, the LD-score method does not investigate the impact of assortative mating.[32] Therefore, assessment of the degree to which this phenomenon may have influenced the genetic correlation estimates was beyond the scope of the present study.

Despite these limitations, the results indicate that neither comorbidity with MDD nor risk variants that are exclusive to MDD explain the genetic overlap between BOR and BIP, SCZ and MDD. Future investigations of larger data sets for BOR and other psychiatric disorders are warranted to refine the analysis of shared and specific genetic risk.

Future studies are warranted to delineate the communalities and specificities of the respective disorders.

CONCLUSION

In summary, the present study is the first GWAS of patients diagnosed with BOR. The results suggest promising novel genes and a novel pathway for BOR, and demonstrate that, rather than being a discrete entity, BOR has an etiological overlap with the major psychoses. The genetic overlap with BIP is consistent with the observation that some diagnostic criteria for BOR overlap with those for BIP. The overlap between BOR and SCZ and MDD is consistent with previous observations of genetic overlap of other psychiatric disorders.[56] Given that BOR patients display specific clinical symptoms not observed in patients with other psychiatric disorders, knowledge of shared and non-shared genetic and clinical features will be important for the development of personalized treatment approaches.

CONFLICT OF INTEREST

JIN Jr is an investigator for Assurex and a consultant for Janssen. AT has received consultancy fees from Janssen and Novartis. The remaining authors declare no conflict of interest.

ACKNOWLEDGMENTS

We thank all patients and control subjects for their participation. We thank the KFO 256 workgroup of the CIMH, and Bipolar Disorder Working Group-, the Major Depression Working Group- and the Schizophrenia Working Group of the Psychiatric Genomics Consortium (PGC-BIP, PGC-MDD, PGC-SCZ) for providing access to the relevant data. The Romanian sample was funded by UEFISCDI, Romania, grant no. 89/2012 to MGS. The CoLaus|PsyCoLaus study was and is supported by research grants from GlaxoSmithKline, the Faculty of Biology and Medicine of Lausanne and the Swiss National Science Foundation (grants 3200B0–105993, 3200B0-118308, 33CSCO-122661, 33CS30-139468 and 33CS30-148401). The study was supported by the German Federal Ministry of Education and Research (BMBF) through the Integrated Network IntegraMent (Integrated Understanding of Causes and Mechanisms in Mental Disorders), under the auspices of the e:Med Programme (grant 01ZX1314A to MMN and SC; grant 01ZX1314G to MR). The study was supported by the German Research Foundation (DFG; grant FOR2107; RI908/11-1 to MR; WI3429/3-1 to SHW; NO246/10-1 to MMN; DA1151/5-1 to UD; KFO 256 BO 1487/12-1 to MB;.SFB 779 TP A08 to BHS).

REFERENCES

1 Tomko RL, Trull TJ, Wood PK, Sher KJ. Characteristics of borderline personality disorder in a community sample: comorbidity, treatment utilization, and general functioning. *J Pers Disord* 2014; **28**: 734–750.
2 Bohus M, Schmahl C. [Psychopathology and treatment of borderline personality disorder]. *Nervenarzt* 2007; **78**: 1069–1080.
3 Lieb K, Zanarini MC, Schmahl C, Linehan MM, Bohus M. Borderline personality disorder. *Lancet* 2004; **364**: 453–461.
4 Zanarini MC, Frankenburg FR, Reich DB, Fitzmaurice G, Weinberg I, Gunderson JG. The 10-year course of physically self-destructive acts reported by borderline patients and axis II comparison subjects. *Acta Psychiatr Scand* 2008; **117**: 177–184.
5 Tsanas A, Saunders KE, Bilderbeck AC, Palmius N, Osipov M, Clifford GD *et al.* Daily longitudinal self-monitoring of mood variability in bipolar disorder and borderline personality disorder. *J Affect Disord* 2016; **205**: 225–233.
6 Calati R, Gressier F, Balestri M, Serretti A. Genetic modulation of borderline personality disorder: systematic review and meta-analysis. *J Psychiatr Res* 2013; **47**: 1275–1287.
7 Lubke GH, Laurin C, Amin N, Hottenga JJ, Willemsen G, van Grootheest G *et al.* Genome-wide analyses of borderline personality features. *Mol Psychiatry* 2014; **19**: 923–929.
8 Gunderson JG, Zanarini MC, Choi-Kain LW, Mitchell KS, Jang KL, Hudson JI. Family study of borderline personality disorder and its sectors of psychopathology. *Arch Gen Psychiatry* 2011; **68**: 753–762.
9 Torgersen S, Lygren S, Oien PA, Skre I, Onstad S, Edvardsen J *et al.* A twin study of personality disorders. *Compr Psychiatry* 2000; **41**: 416–425.
10 Distel MA, Willemsen G, Ligthart L, Derom CA, Martin NG, Neale MC *et al.* Genetic covariance structure of the four main features of borderline personality disorder. *J Pers Disord* 2010; **24**: 427–444.
11 Kendler KS, Myers J, Reichborn-Kjennerud T. Borderline personality disorder traits and their relationship with dimensions of normative personality: a web-based cohort and twin study. *Acta Psychiatr Scand* 2011; **123**: 349–359.
12 Reichborn-Kjennerud T, Ystrom E, Neale MC, Aggen SH, Mazzeo SE, Knudsen GP *et al.* Structure of genetic and environmental risk factors for symptoms of DSM-IV borderline personality disorder. *JAMA Psychiatry* 2013; **70**: 1206–1214.
13 Fornaro M, Orsolini L, Marini S, De Berardis D, Perna G, Valchera A *et al.* The prevalence and predictors of bipolar and borderline personality disorders comorbidity: systematic review and meta-analysis. *J Affect Disord* 2016; **195**: 105–118.
14 Ghaemi SN, Dalley S, Catania C, Barroilhet S. Bipolar or borderline: a clinical overview. *Acta Psychiatr Scand* 2014; **130**: 99–108.
15 Ghaemi SN, Barroilhet S. Confusing borderline personality with severe bipolar illness. *Acta Psychiatr Scand* 2015; **132**: 281–282.
16 Loranger AW, Oldham JM, Tulis EH. Familial transmission of DSM-III borderline personality disorder. *Arch Gen Psychiatry* 1982; **39**: 795–799.
17 Pope HG Jr., Jonas JM, Hudson JI, Cohen BM, Gunderson JG. The validity of DSM-III borderline personality disorder. A phenomenologic, family history, treatment response, and long-term follow-up study. *Arch Gen Psychiatry* 1983; **40**: 23–30.
18 Distel MA, Trull TJ, Willemsen G, Vink JM, Derom CA, Lynskey M *et al.* The five-factor model of personality and borderline personality disorder: a genetic analysis of comorbidity. *Biol Psychiatry* 2009; **66**: 1131–1138.
19 Gale CR, Hagenaars SP, Davies G, Hill WD, Liewald DC, Cullen B *et al.* Pleiotropy between neuroticism and physical and mental health: findings from 108 038 men and women in UK Biobank. *Transl Psychiatry* 2016; **6**: e791.
20 Malouff JM, Thorsteinsson EB, Rooke SE, Schutte NS. Alcohol involvement and the Five-Factor model of personality: a meta-analysis. *J Drug Educ* 2007; **37**: 277–294.
21 Witt S, Dukal H, Hohmeyer C, Radosavljevic-Bjelic S, Schendel D, Frank J *et al.* Biobank of Psychiatric Diseases Mannheim - BioPsy. *Open J Bioresour* 2016; **3**: e2.
22 Loranger A, Sartorius N, Andreoli A, Berger P, Buchheim P, Chanabasavanna S. *German Version of the International Personality Disorder Examination: IPDE*. WHO: Geneva, Switzerland, 1998.
23 First MB, Spitzer RL, Robert L, Gibbon M, Williams JB. *Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version*, Patient Edition. Biometrics Research, New York State Psychiatric Institute: New York, NY, USA, 2002.
24 Consortium SWGotPG. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 2014; **511**: 421–427.
25 Howie B, Marchini J, Stephens M. Genotype imputation with thousands of genomes. *G3 Genes Genomes Genetics* 2011; **1**: 457–470.
26 Delaneau O, Marchini J, Zagury J-F. A linear complexity phasing method for thousands of genomes. *Nat Methods* 2012; **9**: 179–181.
27 Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA *et al.* A map of human genome variation from population-scale sequencing. *Nature* 2010; **467**: 1061–1073.
28 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559–575.
29 de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol* 2015; **11**: e1004219.
30 Zhang K, Cui S, Chang S, Zhang L, Wang J. i-GSEA4GWAS: a web server for identification of pathways/gene sets associated with traits by applying an improved gene set enrichment analysis to genome-wide association study. *Nucleic Acids Res* 2010; **38**(Web Server issue): W90–W95.
31 Nam D, Kim J, Kim SY, Kim S. GSA-SNP: a general approach for gene set analysis of polymorphisms. *Nucleic Acids Res* 2010; **38**(Web Server issue): W749–W754.
32 Bulik-Sullivan BK, Loh P-R, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics C*et al.* LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* 2015; **47**: 291–295.
33 Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR *et al.* An atlas of genetic correlations across human diseases and traits. *Nat Genet* 2015; **47**: 1236–1241.
34 Schizophrenia Working Group of the Psychiatric Genomics C. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 2014; **511**: 421–427.
35 Abstracts of the XXIIIrd World Congress of Psychiatric Genetics (WCPG): final symposia and penary abstracts. *Eur Neuropsychopharmacol* 2015; 13 & 23.
36 International Schizophrenia C, Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC *et al.* Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009; **460**: 748–752.
37 Edwards L, Gupta R, Filipp FV. Hypermutation of DPYD deregulates pyrimidine metabolism and promotes malignant progression. *Mol Cancer Res* 2016; **14**: 196–206.
38 Duan J, Shi J, Fiorentino A, Leites C, Chen X, Moy W *et al.* A rare functional noncoding variant at the GWAS-implicated MIR137/MIR2682 locus might confer risk to schizophrenia and bipolar disorder. *Am J Hum Genet* 2014; **95**: 744–753.
39 Ripke S, O'Dushlaine C, Chambert K, Moran JL, Kahler AK, Akterin S *et al.* Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nat Genet* 2013; **45**: 1150–1159.
40 Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium. Genome-wide association study identifies five new schizophrenia loci. *Nat Genet* 2011; **43**: 969–976.
41 Xu B, Ionita-Laza I, Roos JL, Boone B, Woodrick S, Sun Y *et al. De novo* gene mutations highlight patterns of genetic and neural complexity in schizophrenia. *Nat Genet* 2012; **44**: 1365–1369.
42 Keil R, Schulz J, Hatzfeld M. p0071/PKP4, a multifunctional protein coordinating cell adhesion with cytoskeletal organization. *Biol Chem* 2013; **394**: 1005–1017.
43 Kao CF, Chen HW, Chen HC, Yang JH, Huang MC, Chiu YH *et al.* Identification of susceptible loci and enriched pathways for bipolar II disorder using genome-wide association studies. *Int J Neuropsychopharmacol* 2016; **19**: pyw064.
44 Zhang Z, Yu H, Jiang S, Liao J, Lu T, Wang L *et al.* Evidence for association of cell adhesion molecules pathway and NLGN1 polymorphisms with schizophrenia in Chinese Han population. *PLoS ONE* 2015; **10**: e0144719.
45 Network, Pathway Analysis Subgroup of Psychiatric Genomics C. Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. *Nat Neurosci* 2015; **18**: 199–209.
46 Krueger WH, Gonye GE, Madison DL, Murray KE, Kumar M, Spoerel N *et al.* TPO1, a member of a novel protein family, is developmentally regulated in cultured oligodendrocytes. *J Neurochem* 1997; **69**: 1343–1355.
47 Liu J, Dietz K, DeLoyht JM, Pedre X, Kelkar D, Kaur J *et al.* Impaired adult myelination in the prefrontal cortex of socially isolated mice. *Nat Neurosci* 2012; **15**: 1621–1623.

48 Cupertino RB, Kappel DB, Bandeira CE, Schuch JB, da Silva BS, Muller D *et al.* SNARE complex in developmental psychiatry: neurotransmitter exocytosis and beyond. *J Neural Transm (Vienna)* 2016; **123**: 867–883.
49 Pescosolido MF, Gamsiz ED, Nagpal S, Morrow EM. Distribution of disease-associated copy number variants across distinct disorders of cognitive development. *J Am Acad Child Adolesc Psychiatry* 2013; **52**: 414–430.e14.
50 Zhao Z, Xu J, Chen J, Kim S, Reimers M, Bacanu SA *et al.* Transcriptome sequencing and genome-wide association analyses reveal lysosomal function and actin cytoskeleton remodeling in schizophrenia and bipolar disorder. *Mol Psychiatry* 2015; **20**: 563–572.
51 Sullivan PF, de Geus EJ, Willemsen G, James MR, Smit JH, Zandbelt T *et al.* Genome-wide association for major depressive disorder: a possible role for the presynaptic protein piccolo. *Mol Psychiatry* 2009; **14**: 359–375.
52 Bosanac P, Castle DJ. Schizophrenia and depression. *Med J Australia* 2013; **199**(6 Suppl): S36–S39.
53 Buckley PF, Miller BJ, Lehrer DS, Castle DJ. Psychiatric comorbidities and schizophrenia. *Schizophr Bull* 2009; **35**: 383–402.
54 Anttila V, Bulik-Sullivan B, Finucane HK, Bras J, Duncan L, Escott-Price V *et al.* Analysis of shared heritability in common disorders of the brain. *bioRciv* 2016.
55 Nordsletten AE, Larsson H, Crowley JJ, Almqvist C, Lichtenstein P, Mataix-Cols D. Patterns of nonrandom mating within and across 11 major psychiatric disorders. *JAMA Psychiatry* 2016; **73**: 354–361.
56 Cross-Disorder Group of the Psychiatric Genomics C, Lee SH, Ripke S, Neale BM, Faraone SV, Purcell SM *et al.* Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet* 2013; **45**: 984–994.

Supplementary Information accompanies the paper on the Translational Psychiatry website (http://www.nature.com/tp)

[1]Central Institute of Mental Health, Department of Genetic Epidemiology in Psychiatry, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany; [2]Central Institute of Mental Health, Clinic of Psychosomatic and Psychotherapeutic Medicine, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany; [3]Central Institute of Mental Health, Institute for Psychiatric and Psychosomatic Psychotherapy (IPPP)/Psychosomatic Medicine and Psychotherapy, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany; [4]Department of Psychiatry and Psychotherapy, Charité Universitätsmedizin Berlin, Campus Mitte, Berlin, Germany; [5]Human Genomics Research Group, Department of Biomedicine, University of Basel, Basel, Switzerland; [6]Institute of Human Genetics, University of Bonn, Bonn, Germany; [7]Life and Brain Center, Department of Genomics, University of Bonn, Bonn, Germany; [8]Department of Psychiatry (UPK), University of Basel, Basel, Switzerland; [9]Department of Psychiatry, Charité-Universitätsmedizin Berlin, Campus Benjamin Franklin, Berlin, Germany; [10]Department of Clinical Psychology and Psychotherapy, University of Heidelberg, Heidelberg, Germany; [11]Department of Biological Psychology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands; [12]Department of Clinical Sciences, Psychiatry, Umeå University, Umeå, Sweden; [13]Discipline of Psychiatry, University of Adelaide, Adelaide, SA, Australia; [14]Molecular and Behavioral Neuroscience Institute, University of Michigan, Ann Arbor, MI, USA; [15]Department of Psychiatry, Dalhousie University, Halifax, NS, Canada; [16]Department of Psychiatry and Behavioral Neuroscience, University of Chicago, Chicago, IL, USA; [17]Division Mental Health and Addiction, Oslo University Hospital, Oslo, Norway; [18]NORMENT, University of Oslo, Oslo, Norway; [19]Institute of Pulmonology, Russian State Medical University, Moscow, Russian Federation; [20]Division of Psychiatry, University College London, London, UK; [21]Department of Psychiatry and Psychotherapy, University Hospital Carl Gustav Carus, Dresden, Germany; [22]Inserm, U1144, AP-HP, GH Saint-Louis, Département de Psychiatrie et de Médecine Addictologique, Paris, France; [23]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; [24]National Center for Mental Health, Cardiff University, Cardiff, UK; [25]Health Sciences Research, Mayo Clinic, Rochester, MN, USA; [26]Division of Psychiatry, University of Edinburgh, Edinburgh, UK; [27]Urain Center Rudolf Magnus, Department of Psychiatry, University Medical Center Utrecht, Utrecht, The Netherlands; [28]Department of Biomedicine, Aarhus University, Aarhus, Denmark; [29]iSEQ, Centre for Integrative Sequencing, Aarhus University, Aarhus, Denmark; [30]iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research, Aarhus, Denmark; [31]Medical and Quality Assurance, Clinics of Upper Bavaria, Munich, Germany; [32]Genetic Epidemiology Group, International Agency for Research on Cancer, Lyon, France; [33]Department of Psychiatry and Psychotherapy, University Medical Center Göttingen, Goettingen, Germany; [34]Medical Center of the University of Munich, Campus Innenstadt, Institute of Psychiatric Phenomics and Genomics (IPPG), Munich, Germany; [35]Translational Neuropsychiatry Unit, Department of Clinical Medicine, Aarhus University, Aarhus, Denmark; [36]Queensland Brain Institute, The University of Queensland, Brisbane, QLD, Australia; [37]Department of Psychiatry, McGill University, Montreal, QC, Canada; [38]Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff, UK; [39]Department of Translational Research in Psychiatry, Max Planck Institute of Psychiatry, Munich, Germany; [40]Centre for Psychiatry, Queen Mary University of London, London, UK; [41]UCL Genetics Institute, University College London, London, UK; [42]Laboratory of Psychiatric Genetics, Department of Psychiatry, Poznan University of Medical Sciences, Poznan, Poland; [43]Department of Psychiatry, University of Marburg, Marburg, Germany; [44]Department of Psychiatry, University of Münste, Münster, Germany; [45]Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff, UK; [46]Department of Medical Genetics, Oslo University Hospital Ullevål, Oslo, Norway; [47]NORMENT, KG Jebsen Centre for Psychosis Research, Department of Clinical Science, University of Bergen, Bergen, Norway; [48]Centre for Integrative Biology, Università degli Studi di Trento, Trento, Italy; [49]Indiana University School of Medicine, Department of Biochemistry and Molecular Biology, Indianapolis, IN, USA; [50]Indiana University School of Medicine, Department of Medical and Molecular Genetics, Indianapolis, IN, USA; [51]Faculté de Médecine, Université Paris Est, Créteil, France; [52]Department of Psychiatry and Psychology, Mayo Clinic, Rochester, MN, USA; [53]School of Medical Sciences, University of New South Wales, Sydney, NSW, Australia; [54]Neuroscience Research Australia, Sydney, NSW, Australia; [55]Department of Psychiatry, University of Halle, Halle, Germany; [56]Genetics and Computational Biology, QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia; [57]Department of Psychological Medicine, University of Worcester, Worcester, UK; [58]Department of Psychiatry and Psychotherapy, University Medicine Greifswald, Greifswald, Germany; [59]School of Biomedical and Healthcare Sciences, Plymouth University Peninsula Schools of Medicine and Dentistry, Plymouth, UK; [60]Department of Psychiatry, University of California San Diego, La Jolla, CA, USA;

[61]Biometric Psychiatric Genetics Research Unit, Alexandru Obregia Clinical Psychiatric Hospital, Bucharest, Romania; [62]Mental Health Department, Biomedicine Institute, University Regional Hospital, Málaga, Spain; [63]Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, UK; [64]Department of Psychology, Eberhard Karls Universität Tübingen, Tubingen, Germany; [65]Inserm U955, Psychiatrie Translationnelle, Créteil, France; [66]University Medical Center Utrecht, Division of Neuroscience, Department of Psychiatry, Utrecht, The Netherlands; [67]Center of Psychiatry Weinsberg, Weinsberg, Germany; [68]Department of Psychiatry, Psychosomatic Medicine and Psychotherapy, University Hospital Frankfurt am Main, Frankfurt am Main, Germany; [69]Centre for Addiction and Mental Health, Toronto, ON, Canada; [70]Department of Psychiatry, University of Toronto, Toronto, ON, Canada; [71]Max Planck Institute of Psychiatry, Munich, Germany; [72]Center for Research in Environmental Epidemiology (CREAL), Barcelona, Spain; [73]Institute of Neuroscience and Physiology, University of Gothenburg, Gothenburg, Sweden; [74]Clinic for Psychiatry and Psychotherapy, University Hospital Cologne, Cologne, Germany; [75]Inserm U955, Translational Psychiatry Laboratory, AP-HP, DHU PePSY, Department of Psychiatry, Université Paris Est, Créteil, France; [76]Janssen Research and Development, LLC, Neuroscience Therapeutic Area, Titusville, NJ, USA; [77]M. Sklodowska-Curie Cancer Center and Institute of Oncology, Cancer Epidemiology and Prevention, Warsaw, Poland; [78]School of Psychology, The University of Queensland, Brisbane, QLD, Australia; [79]Lindner Center of HOPE, Research Institute, Mason, OH, USA; [80]Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK; [81]Genetic Cancer Susceptibility Group, International Agency for Research on Cancer, Lyon, France; [82]Division of Psychiatry, University College London, London, UK; [83]VU University Medical Center and GGZ inGeest, Department of Psychiatry, Amsterdam, The Netherlands; [84]School of Psychiatry, University of New South Wales, Sydney, NSW, Australia; [85]Black Dog Institute, Sydney, NSW, Australia; [86]Institute for Molecular Biology, University of Queensland, Brisbane, QLD, Australia; [87]Department of Neuroscience, Norwegian University of Science and Technology, Trondheim, Norway; [88]Department of Psychiatry, St Olavs University Hospital, Trondheim, Norway; [89]Risskov, Psychosis Research Unit, Aarhus University Hospital, Aarhus, Denmark; [90]iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research, Department of Clinical Medicine, Aarhus University, Aarhus, Denmark; [91]Research Center Juelich, Institute of Neuroscience and Medicine (INM-1), Juelich, Germany; [92]Division of Medical Genetics, University of Basel, Basel, Switzerland; [93]Munich Cluster for Systems Neurology (SyNergy), Munich, Germany; [94]University of Liverpool, Liverpool, UK; [95]HudsonAlpha Institute for Biotechnology, Huntsville, AL, USA; [96]Department of Psychiatry, Indiana University School of Medicine, Indianapolis, IN, USA; [97]MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University, Cardiff, UK; [98]Center for Neurobehavioral Genetics, University of California Los Angeles, Los Angeles, CA, USA; [99]University Medical Center Utrecht, Division of Brain Research, Utrecht, The Netherlands; [100]Psychiatry Clinic, Clinical Center University of Sarajevo, Sarajevo, Bosnia-Herzegovina; [101]Pfizer Global Research and Development, Human Genetics and Computational Biomedicine, Groton, CT, USA; [102]Department of Psychiatry, University of Iowa, Iowa City, IA, USA; [103]Department of Psychiatry, Psychiatric University Hospital of Lausanne, Lausanne, Switzerland; [104]Department of Neurology and Neurosurgery, Faculty of Medicine, McGill University, Montreal, QC, Canada; [105]Montreal Neurological Institute and Hospital, Montreal, QC, Canada; [106]Department of Psychiatry and Behavioral Sciences, Johns Hopkins University, Baltimore, MD, USA; [107]NIMH Division of Intramural Research Programs, Human Genetics Branch, Bethesda, MD, USA; [108]Psychiatric Center Nordbaden, Wiesloch, Germany; [109]Center for Statistical Genetics, Department of Biostatistics, University of Michigan, Ann Arbor, MI, USA; [110]School of Medicine and Dentistry, James Cook University, Townsville, QLD, Australia; [111]Broad Institute of MIT and Harvard, Medical and Population Genetics, Cambridge, MA, USA; [112]Division of Psychiatric Genomics, Icahn School of Medicine at Mount Sinai, New York, NY, USA; [113]Centre for Addiction and Mental Health, Molecular Neuropsychiatry and Development Laboratory, Toronto, ON, Canada; [114]Department of Anaesthesiology and Operative Intensive Care, University Hospital Mannheim, Medical Faculty Mannheim/Heidelberg University, Mannheim, Germany; [115]Institute for Molecular Bioscience, The University of Queensland, Brisbane, QLD, Australia; [116]Pfizer Global Research and Development, Computational Sciences Center of Emphasis, Cambridge, MA, USA; [117]AGAPLESION Elisabethenstift gGmbh, Department of Psychiatry, Psychosomatics and Psychotherapy, Darmstadt, Germany; [118]University Medical Center, Department of Psychiatry and Psychotherapy, Mainz, Germany; [119]Leibniz Institute for Neurobiology, Magdeburg, Germany; [120]Department of Biomedicine, University of Basel, Basel, Switzerland; [121]Stanley Center for Psychiatric Research and Medical and Population Genetics Program, Broad Institute of MIT and Harvard, Cambridge, MA, USA and [122]Analytic and Translational Genetics Unit, Department of Medicine, Harvard Medical School, Massachusetts General Hospital, Boston, MA, USA

4.3 Gene set enrichment analysis and expression pattern exploration implicate an involvement of neurodevelopmental processes in bipolar disorder

Journal of Affective Disorders 228 (2018) 20–25

Contents lists available at ScienceDirect

Journal of Affective Disorders

journal homepage: www.elsevier.com/locate/jad

Gene set enrichment analysis and expression pattern exploration implicate an involvement of neurodevelopmental processes in bipolar disorder

Thomas W. Mühleisen[a,b,1], Céline S. Reinbold[b,1], Andreas J. Forstner[c,d], Lilia I. Abramova[e], Martin Alda[f,g], Gulja Babadjanova[h], Michael Bauer[i], Paul Brennan[j], Alexander Chuchalin[h], Cristiana Cruceanu[k,l,m], Piotr M. Czerski[n], Franziska Degenhardt[c,d], Sascha B. Fischer[b], Janice M. Fullerton[o,p], Scott D. Gordon[q], Maria Grigoroiu-Serbanescu[r], Paul Grof[s,t,u], Joanna Hauser[n], Martin Hautzinger[v], Stefan Herms[b,c,d], Per Hoffmann[b,c,d], Jutta Kammerer-Ciernioch[w], Elza Khusnutdinova[x,y], Manolis Kogevinas[z], Valery Krasnov[aa], André Lacour[ab], Catherine Laprise[ac,ad], Markus Leber[ae], Jolanta Lissowska[af], Susanne Lucae[ag], Anna Maaser[c,d], Wolfgang Maier[ah], Nicholas G. Martin[q], Manuel Mattheisen[ai,aj], Fermin Mayoral[ak], James D. McKay[al], Sarah E. Medland[q], Philip B. Mitchell[am,an], Susanne Moebus[ao], Grant W. Montgomery[q], Bertram Müller-Myhsok[ag,ap,aq], Lilijana Oruc[ar], Galina Pantelejeva[e], Andrea Pfennig[i], Lejla Pojskic[as], Alexey Polonikov[at,au], Andreas Reif[av], Fabio Rivas[ak], Guy A. Rouleau[k], Lorena M. Schenk[c,d], Peter R. Schofield[o,p], Markus Schwarz[aw], Fabian Streit[ax], Jana Strohmaier[ax], Neonila Szeszenia-Dabrowska[ay], Alexander S. Tiganov[e], Jens Treutlein[ax], Gustavo Turecki[l,m,az], Helmut Vedder[aw], Stephanie H. Witt[ax], Thomas G. Schulze[ba], Marcella Rietschel[ax], Markus M. Nöthen[c,d], Sven Cichon[a,b,c,d,*]

[a] Institute of Neuroscience and Medicine (INM-1), Research Centre Jülich, Jülich, Germany
[b] Department of Biomedicine & Institute of Medical Genetics and Pathology, Human Genomics Research Group and Division of Medical Genetics, Department of Biomedicine, University and University Hospital Basel, Basel, Switzerland
[c] Institute of Human Genetics, Institute of Human Genetics, University of Bonn School of Medicine & University Hospital Bonn, Bonn, Germany
[d] Department of Genomics, Life & Brain Research Center, University of Bonn, Bonn, Germany
[e] Russian Academy of Medical Sciences, Mental Health Research Center, Moscow, Russian Federation
[f] Department of Psychiatry, Dalhousie University, Halifax, Canada
[g] National Institute of Mental Health, Klecany, Czech Republic
[h] Institute of Pulmonology, Russian State Medical University, Moscow, Russian Federation
[i] Department of Psychiatry and Psychotherapy, University Hospital Carl Gustav Carus, TU Dresden, Dresden, Germany
[j] Genetic Epidemiology Group, International Agency for Research on Cancer (IARC), Lyon, France
[k] Montreal Neurological Institute, McGill University, Montreal, Canada
[l] Department of Human Genetics, McGill University, Montreal, Canada
[m] McGill Group for Suicide Studies & Douglas Research Institute, Montreal, Canada
[n] Laboratory of Psychiatric Genetics, Department of Psychiatry, Poznan University of Medical Sciences, Poznan, Poland
[o] Neuroscience Research Australia, Sydney, Australia
[p] School of Medical Sciences Faculty of Medicine, University of New South Wales, Sydney, Australia
[q] Queensland Institute of Medical Research (QIMR), Brisbane, Australia
[r] Biometric Psychiatric Genetics Research Unit, Alexandru Obregia Clinical Psychiatric Hospital, Bucharest, Romania
[s] The International Group for the Study of Lithium-Treated Patients (IGSLI), Berlin, Germany
[t] Mood Disorders Center of Ottawa, Ottawa, Ontario, Canada K1G 4G3
[u] Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada M5T 1R8
[v] Department of Psychology, Clinical Psychology and Psychotherapy, Eberhard Karls University Tübingen, Tübingen, Germany
[w] Center of Psychiatry Weinsberg, Weinsberg, Germany
[x] Institute of Biochemistry and Genetics, Ufa Scientific Center of Russian Academy of Sciences, Ufa, Russian Federation
[y] Department of Genetics and Fundamental Medicine of Bashkir State University, Ufa, Russian Federation
[z] Center for Research in Environmental Epidemiology (CREAL), Barcelona, Spain
[aa] Moscow Research Institute of Psychiatry, Moscow, Russian Federation
[ab] German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany

* Correspondence to: Department of Biomedicine, University of Basel, Hebelstrasse 20, CH-4031 Basel, Switzerland.
E-mail address: sven.cichon@unibas.ch (S. Cichon).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.jad.2017.11.068

Available online 14 November 2017

[ac] Département des sciences fondamentales, Université du Québec à Chicoutimi, Saguenay, QC, Canada
[ad] Centre intégré universitaire de santé et services sociaux du Saguenay–Lac-Saint-Jean, Saguenay, Québec, Canada
[ae] Department of Psychiatry & Psychotherapy, University of Cologne, Cologne, Germany
[af] Department of Cancer Epidemiology and Prevention, Maria Sklodowska-Curie Memorial Cancer Centre and Institute of Oncology Warsaw, Warsaw, Poland
[ag] Max Planck Institute of Psychiatry, Munich, Germany
[ah] Department of Psychiatry, University of Bonn, Bonn, Germany
[ai] Department of Biomedicine and Centre for integrative Sequencing, iSEQ, Aarhus University, Aarhus, Denmark
[aj] The Lundbeck Foundation Initiative for integrative Psychiatric Research, iPSYCH, Aarhus and Copenhagen, Denmark
[ak] Department of Psychiatry, Hospital Regional Universitario, Biomedical Institute of Malaga, Malaga, Spain
[al] Genetic Cancer Susceptibility Group, International Agency for Research on Cancer (IARC), Lyon, France
[am] School of Psychiatry, University of New South Wales, Randwick, Australia
[an] Black Dog Institute, Prince of Wales Hospital, Randwick, Australia
[ao] Institute of Medical Informatics, Biometry and Epidemiology, University Duisburg-Essen, Essen, Germany
[ap] Munich Cluster for Systems Neurology (SyNergy), Munich, Germany
[aq] University of Liverpool, Institute of Translational Medicine, Liverpool, United Kingdom
[ar] Psychiatric Clinic, Clinical Center University of Sarajevo, Bolnička 25, 71000 Sarajevo, Bosnia and Herzegovina
[as] Institute for Genetic Engineering and Biotechnology, University of Sarajevo, Zmaja od Bosne 8 - Campus, 71000 Sarajevo, Bosnia and Herzegovina
[at] Department of Biology, Medical Genetics and Ecology, Kursk State Medical University, Kursk, Russian Federation
[au] Research Institute for Genetic and Molecular Epidemiology, Kursk State Medical University, Kursk, Russian Federation
[av] Department of Psychiatry, Psychosomatic Medicine and Psychotherapy, University Hospital Frankfurt am Main, Frankfurt am Main, Germany
[aw] Psychiatric Center Nordbaden, Wiesloch, Germany
[ax] Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, Germany
[ay] Department of Epidemiology, Nofer Institute of Occupational Medicine, Lodz, Poland
[az] Department of Psychiatry, McGill University, Montreal, Canada
[ba] Institute of Psychiatric Phenomics and Genomics, Ludwig-Maximilians-University Munich, Munich, Germany

ARTICLE INFO

Keywords:
Bipolar disorder
Pathway analysis
GRB2 events in ERBB2 signaling
NCAM signaling for neurite out-growth
Neurodevelopmental disorder

ABSTRACT

Background: Bipolar disorder (BD) is a common and highly heritable disorder of mood. Genome-wide association studies (GWAS) have identified several independent susceptibility loci. In order to extract more biological information from GWAS data, multi-locus approaches represent powerful tools since they utilize knowledge about biological processes to integrate functional sets of genes at strongly to moderately associated loci.
Methods: We conducted gene set enrichment analyses (GSEA) using 2.3 million single-nucleotide polymorphisms, 397 Reactome pathways and 24,025 patients with BD and controls. RNA expression of implicated individual genes and gene sets were examined in post-mortem brains across lifespan.
Results: Two pathways showed a significant enrichment after correction for multiple comparisons in the GSEA: *GRB2 events in ERBB2 signaling*, for which 6 of 21 genes were BD associated (P_{FDR} = 0.0377), and *NCAM signaling for neurite out-growth*, for which 11 out of 62 genes were BD associated (P_{FDR} = 0.0451). Most pathway genes showed peaks of RNA co-expression during fetal development and infancy and mapped to neocortical areas and parts of the limbic system.
Limitations: Pathway associations were technically reproduced by two methods, although they were not formally replicated in independent samples. Gene expression was explored in controls but not in patients.
Conclusions: Pathway analysis in large GWAS data of BD and follow-up of gene expression patterns in healthy brains provide support for an involvement of neurodevelopmental processes in the etiology of this neuropsychiatric disease. Future studies are required to further evaluate the relevance of the implicated genes on pathway functioning and clinical aspects of BD.

1. Introduction

Bipolar disorder (BD) is a genetically complex mental illness. During the past ten years, several genome-wide association studies (GWAS) of BD were conducted and have identified 19 loci harboring common genetic susceptibility variants (Sullivan et al., 2017). It is assumed that with growing sample sizes the number of loci will increase, as has been successfully demonstrated for schizophrenia, where GWAS in 61,000 patients found 155 independent loci (Sullivan et al., 2017).

Gene set enrichment analysis (GSEA) is a powerful tool to retrieve more biological information from existing GWAS. Such multi-locus approaches utilize functional frameworks of ontologies or pathways to integrate genes at strongly to moderately associated loci. Using the same sample size, GSEA therefore has greater statistical power to detect a polygenic contribution of individually small effects to overall risk than single-locus analyses (Lee et al., 2012).

Here, we applied GSEA algorithms to a large published GWAS on BD, including approximately 9700 patients and 14,200 controls (Mühleisen et al., 2014). We found associations between BD and two signaling pathways involved in brain development.

2. Methods and materials

2.1. Phenotype and SNP data

For GSEA, we used combined data from the Systematic Investigation of the Molecular Causes of Major Mood Disorders and Schizophrenia (MooDS) and Psychiatric Genomics Consortium (PGC) consortia comprising 2,267,487 autosomal single-nucleotide polymorphisms (SNPs) from 9747 patients with life-time diagnoses of BD and 14,278 controls, as described by Mühleisen et al. (2014). Written informed consent was obtained from all patients and controls before participation in the study.

2.2. Gene set enrichment analyses

For discovery, we used Meta-Analysis Gene-set Enrichment of variaNT Associations (MAGENTA; (Segrè et al., 2010)) with its default settings. At genome-wide level, each gene was mapped to the GWAS SNP showing the lowest p-value within gene boundaries (RefSeq definitions), to minimize the effect of a potential confounding factor introduced by overlapping gene boundaries (Sedeño-Cortés and Pavlidis,

2014). P-values of these index SNPs were corrected for confounders such as gene size, SNP density and linkage disequilibrium-related properties in the stepwise multiple linear regression model of MAGENTA. Resulting gene scores were assigned to target gene sets. For each target gene set, the observed number of gene scores above the user-defined threshold (here 95%) is evaluated against the expected number of gene scores above this threshold for gene sets of identical size, randomly sampled from the genome multiple times. A non-parametric test produces the nominal p-value for each tested target gene set. False-discovery rate (FDR) was used to correct for multiple testing (P_{FDR}).

For secondary analysis of the significantly enriched pathways, we applied Gene Set Analysis SNP (GSA-SNP; (Nam et al., 2010)) on the same input data. GSA-SNP uses p-values of SNPs to calculate enrichment scores by using the Z-statistic method. But instead of using the maximum effect per gene as a proxy for the respective gene, we chose the second-best p-value to represent the effect of each gene to avoid spurious associations (Kwon et al., 2012).

For pathways, we used curated target gene sets (pathways) from Reactome as available through the Molecular Signature Database (v6.0; (Subramanian et al., 2005)). Their sizes were restricted from 20 to 200 to avoid overly narrow or broad gene sets. This resulted in 397 sets for GSEA.

2.3. Gene expression data

BrainScope enables interactive visual exploration of spatial and temporal human brain transcriptomes from the Allen Institute for Brain Science (Huisman et al., 2017). Here we focused on the dataset *Developmental Transcriptome* from the BrainSpan atlas that had been pre-processed and re-analyzed by BrainScope's developers resulting in the dataset *Developing human (comparative explorer)* with RNA expression levels of 18,233 genes (Entrez Gene definitions) that were z-score normalized, to have a zero mean and a standard deviation of 1.

To explore changes of co-expressed genes in brain regions and time windows, we used heat maps of the comparative explorer from BrainScope under default settings. Each square of a heat map displayed the average regional expression of the selected gene(s) across pooled tissue samples (replicates, developmental stages) from donor brains (controls). For BD-associated pathways from GSEA results, heat maps were assembled and annotated using standard graphical software. The brain regions covered neocortical areas including primary cortices (auditory, motor, somatosensory, visual), pre- and orbitofrontal cortices, the temporal cortex (inferolateral, posterior superior), the parietal cortex (posteroventral); principal structures of the diencephalon including parts of the basal ganglia (amygdala, striatum) and limbic system (anterior cingulate, amygdala, hippocampus) coiled around and connected to thalamus and hypothalamus; the hindbrain (cerebellar cortex). The time windows comprised fetal development (from early 2nd trimester to birth), infancy (from birth to one year), childhood (from two to eleven years), adolescence (from 13 to 19 years), and adulthood (from 21 to 40 years). BrainScope, BrainSpan, and Entrez Gene are publicly accessible at www.brainscope.nl, www.brainspan.org, and www.ncbi.nlm.nih.gov/gene.

3. Results

3.1. Discovery and validation of BD-associated pathways

GSEA by MAGENTA on MooDS-PGC data revealed two study-wide significant Reactome pathways when applying the significance criterion of FDR < 0.05 (Table 1). The best finding was *GRB2 events in ERBB2 signaling* (P_{FDR} = 0.0377), for which 6 genes were associated (*NRAS*, *KRAS*, *EGFR*, *ERBB2*, *MAPK1*, *HBEGF*) out of 21 in the pathway. The second finding was *NCAM signaling for neurite out-growth* (P_{FDR} = 0.0451) for which 11 of 62 genes were associated (*NCAN*, *SPTBN2*, *FYN*, *NRAS*, *CREB1*, *KRAS*, *CACNB3*, *COL2A1*, *CACNB2*, *MAPK1*, *SPTBN1*). Three significant genes were common to both pathways (*NRAS*, *KRAS*, *MAPK1*). The associated genes showed a balanced contribution to the total significance of the two target gene sets (Supplementary Table 1). The subsequent GSEA by GSA-SNP on the same input data validated the enrichments in the two target gene sets (P = 4.80E-06 and P = 3.28E-08, respectively; Table 1).

3.2. Exploration of gene expression in BD-associated pathways

To assess patterns of co-expressed genes from both pathways in the developing and adult brain, we used data from BrainSpan accessed through BrainScope and screened (i) expression of each single associated gene, (ii) expression of the combined set of associated genes (Combined), and (iii) expression of associated genes in context of target gene sets (Whole pathway). We found that five of the six genes enriched in *GRB2 events in ERBB2 signaling* demonstrated expression peaks during fetal development and infancy, while *MAPK1* expression was lower during prenatal stages and higher during postnatal stages. The combined pattern of the six genes emphasized neural development and was similar to the whole pathway pattern. In *NCAM signaling for neurite out-growth*, four of the ten enriched genes (*NCAN*, *FYN*, *NRAS*, *CREB1*) revealed high expression during fetal and early postnatal development. *CACNB2*, *MAPK1*, *SPTBN1*, and *SPTBN2* showed low expression during fetal stages but increased later on, especially in infancy. Overall, most genes showed peaks of co-expression during fetal development (early second to third trimester) and infancy (birth to 18 months) in many neocortical areas and parts of the limbic system. Spatio-temporal expression patterns of genes stratified by pathway are displayed in Fig. 1.

Table 1
Association results of the GSEA. MAGENTA and GSA-SNP were used for discovery and validation steps.

			MAGENTA, 95th percentile enrichment cutoff		**GSA-SNP**, 2nd best SNP
Gene set name	**Gene set identifier**	***N* genes**	**P_{FDR}**	**Sign. genes (gene p-value)**	**Empirical p-value**
GRB2 events in ERBB2 signaling	R-HSA-1963640	21	0.0377	*NRAS* (1.94E-03), *KRAS* (2.20E-03), *EGFR* (6.18E-03), *ERBB2* (0.0196), *MAPK1* (0.0222), *HBEGF* (0.0306)	4.80E-06
NCAM signaling for neurite out-growth	R-HSA-375165	62	0.0451	*NCAN* (1.40E-05), *SPTBN2* (6.64E-05), *FYN* (2.75E-04), *NRAS* (1.94E-03), *CREB1* (2.11E-03), *KRAS* (2.20E-03), *CACNB3* (4.86E-03), *COL2A1* (0.0127), *CACNB2* (0.0138), MAPK1 (0.0222), *SPTBN1* (0.0251)	3.28E-08

Abbreviations: *CACNB2*, calcium voltage-gated channel auxiliary subunit beta 2; *CACNB3*, calcium voltage-gated channel auxiliary subunit beta 3; *COL2A1*, collagen type II alpha 1 chain; *CREB1*, cAMP responsive element binding protein 1; *EGF*, epidermal growth factor; *EGFR*, EGF receptor; *ERBB2*, Erb-B2 receptor tyrosine kinase 2; *FYN*, FYN proto-oncogene; *GRB2*, Growth factor receptor-bound protein 2; *HBEGF*, heparin-binding EGF-like growth factor; *KRAS*, KRAS proto-oncogene, GTPase; *MAPK1*, mitogen-activated protein kinase 1; N, number; *NCAM1*, neural cell adhesion molecule 1; *NCAN*, neurocan; *NRAS*, neuroblastoma RAS Viral oncogene homolog; P_{FDR}, FDR-adjusted p-value; *SPTBN1*, spectrin beta, non-erythrocytic 1; *SPTBN2*, spectrin beta, non-erythrocytic 2.

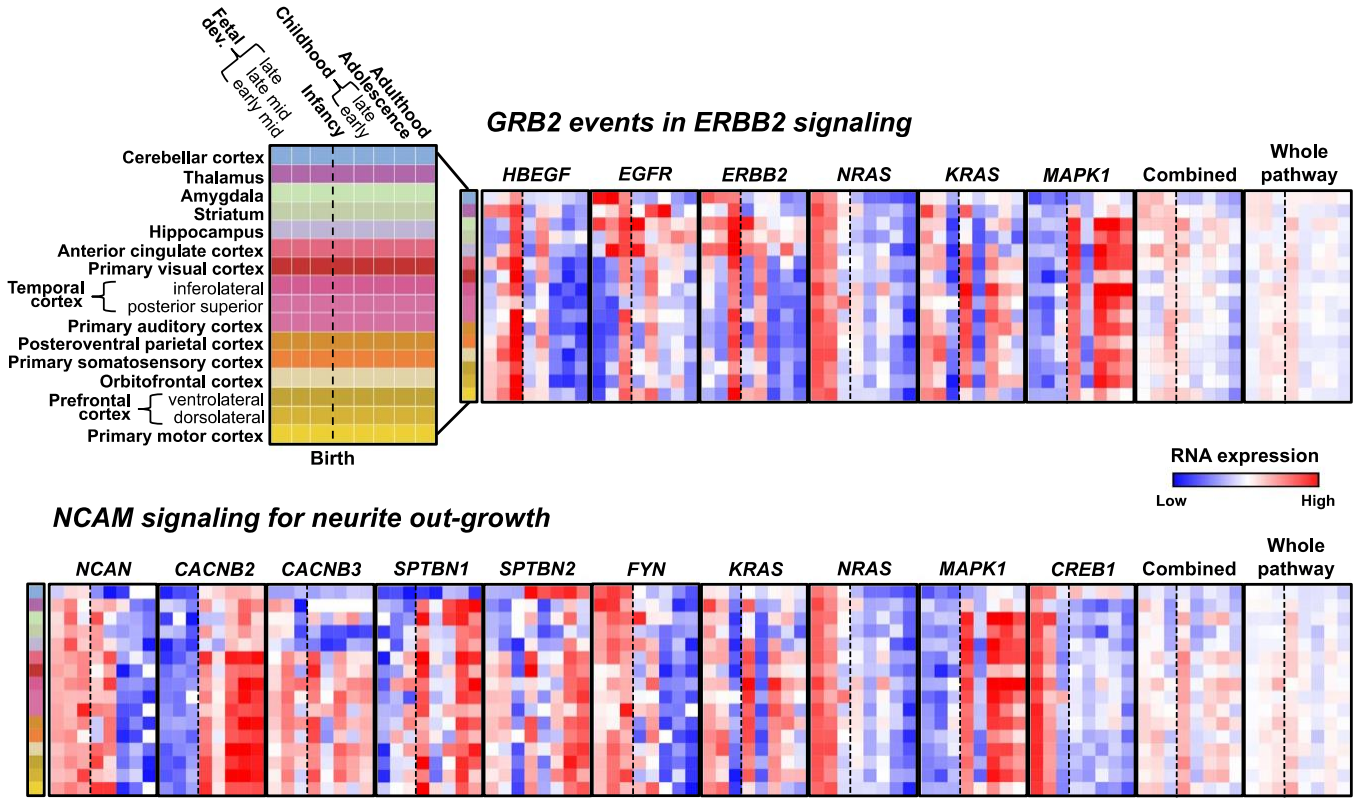


Fig. 1. Expression patterns of genes in BD-associated pathways during normal brain development. Each square of a heat map displays the spatio-temporal expression of the selected gene(s) in the indicated regions and stages in control brains. Levels of RNA expression are z-score normalized ranging from blue (low) over white (zero mean) to red (high). Patterns are shown for single enriched genes (gene symbols), the combined set of enriched genes (combined), and the target gene set (whole pathway). ERBB2 is a member of the EGF receptor family. Since ERBB2 has no ligand-binding domain, it needs a co-receptor to become activated. Upon binding of an EGF ligand, the ERBB2-EGFR heterodimer recruits adaptor protein GRB2 leading to SOS1-mediated guanine-nucleotide exchange on RAS (KRAS, NRAS) and activation of RAF and the MAP kinase cascade (MAPK1). NCAM1 works on modulation of intracellular signaling, either by activation of FGF receptors or cytoplasmic tyrosine kinases (FYN) that initiate MAP kinase cascades (MAPK1) and a transcription factor (CREB1) which regulates expression of genes for growth and survival of neurites. Spectrins (SPTBN1, SPTBN2) are cytoskeletal molecules and manage to link RPTP-alpha to the cytoplasmic domain of NCAM1. L-type channels (CACNB2, CACNB3) associate with NCAM1 in growth cones at the sites of NCAM1 clusters leading to processes that promote neurite out-growth. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

4. Discussion

Current disease models of BD suggest a multifactorial etiology resulting from the additive effects of many gene variants at different loci together with the effect of environmental factors. GWAS have demonstrated that the genotype relative risks of the involved common susceptibility variants are small and that large sample sizes are necessary to achieve sufficient statistical power to identify them (Sullivan et al., 2017). In the present analysis, we chose to apply GSEA to our GWAS data because this approach should have greater statistical power to detect a polygenic contribution of individually small effects to overall risk than single-locus analyses (Lee et al., 2012). To further strengthen our findings, we investigated genes within the implicated pathways for expression at milestones of normal brain development to obtain information on their relevance during ontogenetic stages. Biological pathway studies of BD so far have found evidence for genes involved in calcium channels, hormonal regulation, glutamate signaling, neural development, and histone methylation (Nurnberger et al., 2014; O'Dushlaine et al., 2015).

Our strongest finding was *GRB2 events in ERBB2 signaling* which functions to promote cell proliferation, survival, and differentiation, not only in the brain. Biologically, an association with *ERBB2*, *EGFR*, and *HBEGF* is plausible because they form a ligand-activated receptor complex for signaling and thus seem to be key players of that pathway. The importance of *ERBB2* in BD is further supported by a genome-wide significant association finding (Hou et al., 2016) and by the observation of dysregulated *ERBB2* expression in the dorsolateral prefrontal cortex in both BD and schizophrenia (Shao and Vawter, 2008). This expression alteration is significantly related to lifetime antipsychotic exposure, supporting *ERBB2* as target for clinical research. *ErbB2/B4*-deficient mice exhibit elevated aggression and reduced prepulse inhibition that both can be rescued by clozapine treatment, a frequently used antipsychotic medication (Barros et al., 2009). *EGFR* (alias *ERBB1*) is reported to play an essential role in axon myelination during the first postnatal weeks and can therefore be considered as an important regulator of neurodevelopment (Aguirre et al., 2007). The gene was also supported by single SNP and haplotype analysis in a GWAS of BD (Sklar et al., 2008). HBEGF is a EGF-like binding partner of EGFR and mice lacking *Hb-egf* in the ventral forebrain showed abnormalities in psychomotor behavior and neurotransmission which can be ameliorated by typical or atypical antipsychotics (Oyagi et al., 2009).

Our second finding was *NCAM signaling for neurite out-growth* which modulates neural differentiation and synaptic plasticity. Homophilic binding of NCAM1 molecules at the cell-surface induces signaling that leads to cell-cell adhesion and axon elongation. Association with *NCAN* in this pathway is of major importance since experiments in rats have demonstrated that interference of Ncam1-Ncam1 bindings by concurrent Ncan inhibits these cellular processes (Retzler et al., 1996). *NCAN* encodes an extracellular matrix proteoglycan and has been described as important susceptibility gene for BD (Cichon et al., 2011). Furthermore, *NCAN* was reported to be associated with brain development in health and disease, specifically to gray matter loss in central limbic regions and higher folding in the lateral occipital and prefrontal cortex suggesting impairments of emotion perception and regulation and top-down cognitive functioning (Dannlowski et al., 2015). Behavioral abnormalities in *Ncan*-deficient mice show striking similarities

with mania symptoms in humans that can be rescued by lithium treatment, an established mood stabilizer (Miró et al., 2012). Association with *CACNB2* and *CACNB3* represents another highlight of this pathway, since abnormal calcium channel activity is considered to be important for BD (Nurnberger et al., 2014). Unexpectedly, *CACNA1C* was not found among enriched pathway genes, despite strong support of this gene from SNP data. Further evaluation revealed that *CACNA1C* was absent from the pathway definition. A possible link to our other finding exists through a gene overlap of *KRAS*, *NRAS* and *MAPK1* as well as binding between *NCAM1* and *EGFR*.

In both pathways, most genes showed high co-expression during fetal development and infancy in many neocortical and subcortical areas indicating co-expression and possibly co-working of encoded protein functions. These observations provide links to brain regions where known pathophysiological changes in BD patients occur, for instance, in the limbic system which is concerned with many aspects of emotion and behavior.

4.1. Limitations

Although both pathway findings were technically reproduced by two different approaches and are based on one of the largest GWAS data of BD so far, association replication in independent samples was not attempted. Gene expression was explored in control brains only, which may show co-expression differences compared with BD brains. Follow-up studies are required to further evaluate the relevance of our findings for etiological and clinical aspects of BD.

5. Conclusion

The present study found evidence for associations between BD and two signaling pathways. Integration of evidence from genetic studies, brain developmental expression patterns and molecular functions of these pathways support the hypothesis that neurodevelopmental processes play an important role in the etiology of BD.

Acknowledgments

We are grateful to all patients and control subjects who contributed to this study. We thank the Bipolar Disorder Working Groups of the MooDS and PGC consortia for providing access to the relevant data. C.L. is the Canada Research Chair in Environment and Genetics of Respiratory Disorders and Allergy, Director of the Asthma Strategic Group of the Respiratory Health Network (RHN) of Fonds de la recherche en santé du Québec (FRSQ) and researcher of the CHILD Study.

Role of the funding source

The study was supported by the German Federal Ministry of Education and Research (BMBF) through the Integrated Network IntegraMent (Integrated Understanding of Causes and Mechanisms in Mental Disorders), under the auspices of the e:Med Programme (grant 01ZX1314A to M.M.N., F.D. and S.C., grant 01ZX1314G to M.R.). M.M.N. is a member of the DFG-funded Excellence-Cluster ImmunoSensation. M.M.N. also received support from the Alfried Krupp von Bohlen und Halbach-Stiftung. The study was supported by the German Research Foundation (DFG; grant FOR2107; RI908/11-1 to M.R.; NO246/10-1 to M.M.N.). The study was also supported by the Swiss National Science Foundation (SNSF, grant 156791 to S.C.). M.G.S. received grant no. 89/2012 from UEFISCDI, Romania. Canadian patients were genotyped within the ConLiGen project (www.ConLiGen.org), with the support of a grant from the DFG to M.R., M.B., and T.G.S. (RI 908/7-1). Controls for Germany II were drawn from the Heinz Nixdorf Recall Study (HNR) cohort, which was established with the support of the Heinz Nixdorf Foundation. Recruitment of the Australian sample was supported by an Australian National Health and Medical Research Council (NHMRC) program grant (number 1037196). The recruitment of the Canadian patients was supported by a grant from the Canadian Institutes of Health Research #64410 to M.A. The study also used data generated by the GABRIEL consortium (controls for the sample Russia). Funding for the generation of these data was provided by the European Commission as part of GABRIEL contract number 018996 under the Integrated Program LSH-2004-1.2.5-1. Post genomic approaches to understand the molecular basis of asthma aiming at a preventive or therapeutic control and the Wellcome Trust under award 084703. Canadian controls were drawn from the French Canadian study (SLSJ), which was supported in part by the Canada Research Chair Environment and genetics of respiratory diseases and allergy, The Canadian Institutes of Health Research (Operating grant No. MOP-13506), and the Quebec Respiratory Network of the Fonds de rechercheen santé du Québec (FRQS). Polish controls were recruited by the International Agency for Research on Cancer (IARC)/Centre National de Genotypage (CNG) GWAS Initiative. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.jad.2017.11.068.

References

Aguirre, A., Dupree, J.L., Mangin, J.M., Gallo, V., 2007. A functional role for EGFR signaling in myelination and remyelination. Nat. Neurosci. 10, 990–1002.
Barros, C.S., Calabrese, B., Chamero, P., Roberts, A.J., Korzus, E., Lloyd, K., et al., 2009. Impaired maturation of dendritic spines without disorganization of cortical cell layers in mice lacking NRG1/ErbB signaling in the central nervous system. Proc. Natl. Acad. Sci. USA 106, 4507–4512.
Cichon, S., Mühleisen, T.W., Degenhardt, F.A., Mattheisen, M., Miró, X., Strohmaier, J., et al., 2011. Genome-wide association study identifies genetic variation in neurocan as a susceptibility factor for bipolar disorder. Am. J. Hum. Genet. 88, 372–381.
Dannlowski, U., Kugel, H., Grotegerd, D., Redlich, R., Suchy, J., Opel, N., et al., 2015. NCAN cross-disorder risk variant is associated with limbic gray matter deficits in healthy subjects and major depression. Neuropsychopharmacology 2510–2516.
Hou, L., Bergen, S.E., Akula, N., Song, J., Hultman, C.M., Landén, M., et al., 2016. Genome-wide association study of 40,000 individuals identifies two novel loci associated with bipolar disorder. Hum. Mol. Genet. 25, 3383–3394.
Huisman, S.M.H.H., van Lew, B., Mahfouz, A., Pezzotti, N., Höllt, T., Michielsen, L., et al., 2017. BrainScope: interactive visual exploration of the spatial and temporal human brain transcriptome. Nucleic Acids Res. gkx046.
Kwon, J.-S., Kim, J., Nam, D., Kim, S., 2012. Performance comparison of two gene set analysis methods for genome-wide association study results: GSA-SNP vs i-GSEA4GWAS. Genom. Inform. 10, 123–127.
Lee, S.H., Yang, J., Goddard, M.E., Visscher, P.M., Wray, N.R., 2012. Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. Bioinformatics 28, 2540–2542.
Miró, X., Meier, S., Dreisow, M.L., Frank, J., Strohmaier, J., Breuer, R., et al., 2012. Studies in humans and mice implicate neurocan in the etiology of mania. Am. J. Psychiatry 169, 982–990.
Mühleisen, T.W., Leber, M., Schulze, T.G., Strohmaier, J., Degenhardt, F., Treutlein, J., et al., 2014. Genome-wide association study reveals two new risk loci for bipolar disorder. Nat. Commun. 5, 3339.
Nam, D., Kim, J., Kim, S.-Y., Kim, S., 2010. GSA-SNP: a general approach for gene set analysis of polymorphisms. Nucleic Acids Res. 38, W749–W754.
Nurnberger Jr., J.I., Koller, D.L., Jung, J., Edenberg, H.J., Foroud, T., Guella, I., et al., 2014. Identification of pathways for bipolar disorder: a meta-analysis. JAMA Psychiatry 71, 657–664.
O'Dushlaine, C., Rossin, L., Lee, P.H., Duncan, L., Parikshak, N.N., Newhouse, S., et al., 2015. Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. Nat. Neurosci. 18, 199–209.
Oyagi, A., Oida, Y., Kakefuda, K., Shimazawa, M., Shioda, N., Moriguchi, S., et al., 2009. Generation and characterization of conditional heparin-binding EGF-like growth factor knockout mice. PLoS One 4.
Retzler, C., Göhring, W., Rauch, U., 1996. Analysis of neurocan structures interacting with the neural cell adhesion molecule N-CAM. J. Biol. Chem. 271, 27304–27310.
Sedeño-Cortés, A.E., Pavlidis, P., 2014. Pitfalls in the application of gene-set analysis to genetics studies. Trends Genet. 30, 513–514.

Segrè, A.V., Consortium, D., investigators, M., Groop, L., Mootha, V.K., Daly, M.J., et al., 2010. Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. PLoS Genet. 6, e1001058.
Shao, L., Vawter, M.P., 2008. Shared gene expression alterations in schizophrenia and bipolar disorder. Biol. Psychiatry 64, 89–97.
Sklar, P., Smoller, J.W., Fan, J., Ferreira, M.A.R., Perlis, R.H., Chambert, K., et al., 2008. Whole-genome association study of bipolar disorder. Mol. Psychiatry 13, 558–569.
Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., et al., 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA 102, 15545–15550.
Sullivan, P.F., Agrawal, A., Bulik, C., Andreassen, O.A., Borglum, A., Breen, G., et al., 2017. Psychiatric genomics: an update and an agenda. Am J Psychiatry. http://dx.doi.org/10.1176/appi.ajp.2017.17030283. [Epub ahead of print]. (bioRxiv).

4.4 Analysis of the Influence of microRNAs in Lithium Response in Bipolar Disorder

frontiers in Psychiatry

ORIGINAL RESEARCH
published: xx May 2018
doi: 10.3389/fpsyt.2018.00207

Analysis of the Influence of microRNAs in Lithium Response in Bipolar Disorder

*Céline S. Reinbold[1,2†], Andreas J. Forstner[1,2,3,4,5†], Julian Hecker[6], Janice M. Fullerton[7,8], Per Hoffmann[1,2,3,4], Liping Hou[9], Urs Heilbronner[10], Franziska Degenhardt[3,4], Mazda Adli[11], Kazufumi Akiyama[12], Nirmala Akula[9], Raffaella Ardau[13], Bárbara Arias[14], Lena Backlund[15], Antonio Benabarre[16], Susanne Bengesser[17], Abesh K. Bhattacharjee[18], Joanna M. Biernacka[19,20], Armin Birner[17], Cynthia Marie-Claire[21,22], Pablo Cervantes[23], Guo-Bo Chen[24], Hsi-Chung Chen[25], Caterina Chillotti[13], Scott R. Clark[26], Francesc Colom[27], David A. Cousins[28], Cristiana Cruceanu[29], Piotr M. Czerski[30], Alexandre Dayer[31], Bruno Étain[21,22,32], Peter Falkai[33], Louise Frisén[15], Sébastien Gard[34], Julie S. Garnham[35], Fernando S. Goes[36], Paul Grof[37], Oliver Gruber[38], Ryota Hashimoto[39], Joanna Hauser[30], Stefan Herms[1,2,3,4], Stéphane Jamain[40], Esther Jiménez[16], Jean-Pierre Kahn[41], Layla Kassem[9], Sarah Kittel-Schneider[42], Sebastian Kliwicki[43], Barbara König[44], Ichiro Kusumi[45], Nina Lackner[17], Gonzalo Laje[9], Mikael Landén[46,47], Catharina Lavebratt[15], Marion Leboyer[48], Susan G. Leckband[49], Carlos A. López Jaramillo[50], Glenda MacQueen[51], Mirko Manchia[52,53], Lina Martinsson[54], Manuel Mattheisen[55], Michael J. McCarthy[56], Susan L. McElroy[57], Marina Mitjans[14], Francis M. Mondimore[36], Palmiero Monteleone[58,59], Caroline M. Nievergelt[18], Urban Ösby[60], Norio Ozaki[61], Roy H. Perlis[62], Andrea Pfennig[63], Daniela Reich-Erkelenz[10], Guy A. Rouleau[64], Peter R. Schofield[8,65], K. Oliver Schubert[26], Barbara W. Schweizer[36], Florian Seemüller[33], Giovanni Severino[66], Tatyana Shekhtman[18], Paul D. Shilling[18], Kazutaka Shimoda[67], Christian Simhandl[68,69], Claire M. Slaney[35], Jordan W. Smoller[62], Alessio Squassina[66], Thomas J. Stamm[11,69], Pavla Stopkova[71], Sarah K. Tighe[72], Alfonso Tortorella[59], Gustavo Turecki[29], Julia Volkert[42], Stephanie H. Witt[73], Adam J. Wright[74], L. Trevor Young[75], Peter P. Zandi[76], James B. Potash[71], J. Raymond DePaulo[36], Michael Bauer[63], Eva Reininghaus[17], Tomáš Novák[70], Jean-Michel Aubry[31], Mario Maj[77], Bernhard T. Baune[26], Philip B. Mitchell[74], Eduard Vieta[16], Mark A. Frye[20], Janusz K. Rybakowski[43], Po-Hsiu Kuo[78], Tadafumi Kato[79], Maria Grigoroiu-Serbanescu[80], Andreas Reif[42], Maria Del Zompo[66], Frank Bellivier[21,22,32], Martin Schalling[15], Naomi R. Wray[24], John R. Kelsoe[18], Martin Alda[35,70], Francis J. McMahon[9], Thomas G. Schulze[9,10,36,38,73], Marcella Rietschel[73], Markus M. Nöthen[3,4] and Sven Cichon[1,2,3,4,81]**

[1] Human Genomics Research Group, Department of Biomedicine, University of Basel, Basel, Switzerland, [2] Institute of Medical Genetics and Pathology, University Hospital Basel, Basel, Switzerland, [3] Institute of Human Genetics, University of Bonn, Bonn, Germany, [4] Department of Genomics, Life & Brain Center, University of Bonn, Bonn, Germany, [5] Department of Psychiatry (UPK), University of Basel, Basel, Switzerland, [6] Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, United States, [7] Neuroscience Research Australia, Sydney, NSW, Australia, [8] School of Medical Sciences, University of New South Wales, Sydney, NSW, Australia, [9] Intramural Research Program, US Department of Health & Human Services, National Institute of Mental Health, National Institutes of Health, Bethesda, MD, United States, [10] Department Psychiatry and Psychotherapy, Institute of Psychiatric Phenomics and Genomics, University Hospital, LMU Munich, Munich, Germany, [11] Department of Psychiatry and Psychotherapy, Charité Universitätsmedizin, Berlin, Germany, [12] Department of Biological Psychiatry and Neuroscience, Dokkyo Medical University School of Medicine, Mibu, Japan, [13] Unit of Clinical Pharmacology, University Hospital of Cagliari, Cagliari, Italy, [14] Zoology and Biological Anthropology Section (Department of Evolutive Biology, Ecology and Environmental Sciences), Facultat de Biologia and Institut de Biomedicina, CIBERSAM, Universitat de Barcelona, Barcelona, Spain, [15] Department of Molecular Medicine and Surgery, Karolinska Institutet and Center for Molecular Medicine, Karolinska University Hospital, Stockholm, Sweden, [16] Bipolar Disorder Program, Institute of Neuroscience, Hospital Clinic, CIBERSAM, IDIBAPS, University of Barcelona, Barcelona, Spain, [17] Special Outpatient Center for Bipolar Affective Disorder, Medical University of Graz, Graz, Austria, [18] Health Sciences Research, Mayo Clinic, Rochester, MN, United States, [19] Department of Psychiatry and Psychology, Mayo Clinic, Rochester, MN, United States, [20] Institut National de la Santé et de la Recherche Médicale, U1144, Paris, France, [21] Université Paris Diderot, Sorbonne Paris Cité, UMR-S 1144, Paris, France, [22] Department of Psychiatry, University of California San Diego, San Diego, San Diego, CA, United States, [23] The Neuromodulation Unit, McGill University Health Centre, Montreal, QC, Canada, [24] Queensland Brain Institute, The University of Queensland, Brisbane, QLD, Australia

OPEN ACCESS

Edited by:
Elizabeth A. Thomas, The Scripps Research Institute, United States

Reviewed by:
Janine Adele Lamb, University of Manchester, United Kingdom
Emma Sprooten, Donders Institute for Brain, Cognition and Behaviour, Radboud University, Netherlands

***Correspondence:**
Sven Cichon
sven.cichon@unibas.ch

[†] *These authors have contributed equally to this work.*

[25] Department of Psychiatry & Center of Sleep Disorders, National Taiwan University Hospital, Taipei, Taiwan, [26] Discipline of Psychiatry, University of Adelaide, Adelaide, SA, Australia, [27] Mental Health Research Group, IMIM-Hospital del Mar, Barcelona, Spain, [28] Campus for Ageing and Vitality, Newcastle University, Newcastle upon Tyne, United Kingdom, [29] Douglas Mental Health University Institute, McGill University, Montreal, QC, Canada, [30] Psychiatric Genetic Unit, Poznan University of Medical Sciences, Poznan, Poland, [31] Mood Disorders Unit, HUG - Geneva University Hospitals, Geneva, Switzerland, [32] AP-HP, GH Saint-Louis – Lariboisière – F. Widal, Département de Psychiatrie et de Médecine Addictologique, Paris, France, [33] Department of Psychiatry and Psychotherapy, Ludwig-Maximilians-University Munich, Munich, Germany, [34] Service de Psychiatrie, Hôpital Charles Perrens, Bordeaux, France, [35] Department of Psychiatry, Dalhousie University, Halifax, NS, Canada, [36] Department of Psychiatry and Behavioral Sciences, Johns Hopkins University, Baltimore, MD, United States, [37] Mood Disorders Center of Ottawa, Ottawa, ON, Canada, [38] Department of Psychiatry and Psychotherapy, University Medical Center, Georg-August University Göttingen, Göttingen, Germany, [39] Molecular Research Center for Children's Mental Development, United Graduate School of Child Development, Osaka University, Osaka, Japan, [40] Institut National de la Santé et de la Recherche Médicale U955, Psychiatrie Translationnelle, Créteil, France, [41] Service de Psychiatrie et Psychologie Clinique, Centre Psychothérapique de Nancy - Université de Lorraine, Nancy, France, [42] Department of Psychiatry, Psychosomatic Medicine and Psychotherapy, University Hospital Frankfurt, Frankfurt, Germany, [43] Department of Adult Psychiatry, Poznan University of Medical Sciences, Poznan, Poland, [44] Department of Psychiatry and Psychotherapeuthic Medicine, Landesklinikum Neunkirchen, Neunkirchen, Austria, [45] Department of Psychiatry, Hokkaido University Graduate School of Medicine, Sapporo, Japan, [46] Institute of Neuroscience and Physiology, The Sahlgrenska Academy at the Gothenburg University, Gothenburg, Sweden, [47] Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden, [48] Assistance Publique-Hôpitaux de Paris, Hôpital Albert Chenevier - Henri Mondor, Pôle de Psychiatrie, Créteil, France, [49] Department of Pharmacy, VA San Diego Healthcare System, San Diego, CA, United States, [50] Department of Psychiatry, University of Antioquia, Medellín, Medellín, Colombia, [51] Department of Psychiatry, University of Calgary, Calgary, AB, Canada, [52] Section of Psychiatry, Department of Medical Sciences and Public Health, University of Cagliari, Cagliari, Italy, [53] Department of Pharmacology, Dalhousie University, Halifax, NS, Canada, [54] Department of Clinical Neurosciences, Karolinska Institutet, Stockholm, Sweden, [55] Department of Biomedicine, Aarhus University, Aarhus, Denmark, [56] Department of Psychiatry, VA San Diego Healthcare System, San Diego, CA, United States, [57] Department of Psychiatry, Lindner Center of Hope/University of Cincinnati, Mason, OH, United States, [58] Neurosciences Section, Department of Medicine and Surgery, University of Salerno, Salerno, Italy, [59] Department of Psychiatry, University of Campania "Luigi Vanvitelli", Naples, Italy, [60] Department of Neurobiology, Care Sciences, and Society, Karolinska Institutet and Center for Molecular Medicine, Karolinska University Hospital, Stockholm, Sweden, [61] Department of Psychiatry, Nagoya University Graduate School of Medicine, Nagoya, Japan, [62] Department of Psychiatry, Massachusetts General Hospital and Harvard Medical School, Boston, MA, United States, [63] Department of Psychiatry and Psychotherapy, Medical Faculty, University Hospital Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany, [64] Department of Neurology and Neurosurgery, Montreal Neurological Institute and Hospital, McGill University, Montreal, QC, Canada, [65] Mental Illness, Neuroscience Research Australia, Sydney, NSW, Australia, [66] Department of Biomedical Sciences, University of Cagliari, Cagliari, Italy, [67] Department of Psychiatry, Dokkyo Medical University School of Medicine, Mibu, Japan, [68] Medical school, Sigmund Freud University, Vienna, Austria, [69] Bipolar Center Wiener Neustadt,Vienna, Austria, [70] Department of Psychiatry, Psychotherapy and Psychosomatics, Medical School Brandenburg, Neuruppin, Germany, [71] Department of Psychiatry, National Institute of Mental Health, Klecany, Czechia, [72] Department of Psychiatry, University of Iowa, Iowa, IA, United States, [73] Department of Psychiatry, University of Naples, SUN, Naples, Italy, [74] Department of Genetic Epidemiology in Psychiatry, Medical Faculty Mannheim, Central Institute of Mental Health, University of Heidelberg, Mannheim, Germany, [75] School of Psychiatry, University of New South Wales, and Black Dog Institute, Sydney, NSW, Australia, [76] Department of Psychiatry, University of British Columbia, Vancouver, BC, Canada, [77] Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States, [78] Department of Public Health and Institute of Epidemiology and Preventive Medicine, National Taiwan University, Taipei, Taiwan, [79] Laboratory for Molecular Dynamics of Mental Disorders, RIKEN Brain Science Institute, Saitama, Japan, [80] Biometric Psychiatric Genetics Research Unit, Alexandru Obregia Clinical Psychiatric Hospital, Bucharest, Romania, [81] Research Centre Jülich, Institute of Neuroscience and Medicine (INM-1), Jülich, Germany

Bipolar disorder (BD) is a common, highly heritable neuropsychiatric disease characterized by recurrent episodes of mania and depression. Lithium is the best-established long-term treatment for BD, even though individual response is highly variable. Evidence suggests that some of this variability has a genetic basis. This is supported by the largest genome-wide association study (GWAS) of lithium response to date conducted by the International Consortium on Lithium Genetics (ConLiGen). Recently, we performed the first genome-wide analysis of the involvement of miRNAs in BD and identified nine BD-associated miRNAs. However, it is unknown whether these miRNAs are also associated with lithium response in BD. In the present study, we therefore tested whether common variants at these nine candidate miRNAs contribute to the variance in lithium response in BD. Furthermore, we systematically analyzed whether any other miRNA in the genome is implicated in the response to lithium. For this purpose, we performed gene-based tests for all known miRNA coding genes in the ConLiGen

GWAS dataset (n = 2,563 patients) using a set-based testing approach adapted from the versatile gene-based test for GWAS (VEGAS2). In the candidate approach, *miR-499a* showed a nominally significant association with lithium response, providing some evidence for involvement in both development and treatment of BD. In the genome-wide miRNA analysis, 71 miRNAs showed nominally significant associations with the dichotomous phenotype and 106 with the continuous trait for treatment response. A total of 15 miRNAs revealed nominal significance in both phenotypes with *miR-633* showing the strongest association with the continuous trait (p = 9.80E-04) and *miR-607* with the dichotomous phenotype (p = 5.79E-04). No association between miRNAs and treatment response to lithium in BD in either of the tested conditions withstood multiple testing correction. Given the limited power of our study, the investigation of miRNAs in larger GWAS samples of BD and lithium response is warranted.

Keywords: bipolar disorder, lithium response, microRNA, common variants, genome-wide association study

INTRODUCTION

Bipolar disorder (BD) is a severe neuropsychiatric condition categorized by recurrent episodes of depression and mania. BD is common, with a lifetime prevalence of around 1% in the general population (1). The elevated morbidity and mortality, the typically early age at onset in young adulthood and the chronic course of BD make it a major public health problem, and BD is classified as one of the top 25 leading causes of the global burden of disease (2). Epidemiological and molecular genetic data strongly suggest that BD is a complex disorder (3) which means that both genetic and environmental factors influence illness risk. Based on twin studies the overall heritability of BD has been estimated to be over 70% (4, 5), suggesting a substantial involvement of genetic factors in the development of the disease.

Mood stabilizers are used as the first-line mode of medication in the treatment of BD (6). Amongst these drugs, lithium is used as a preventive agent for manic and depressive episodes (7), suicide attempts, and death by suicide, and shows the greatest support for long-term relapse prevention (8, 9). Consequently, lithium is endorsed as a first-line and best-established long-term treatment for BD, even though individual response is highly variable (6, 8, 10). Evidence suggests that some of the variability in lithium response has a genetic basis (11, 12). This hypothesis is supported by the largest genome-wide association study (GWAS) of lithium response to date, which was conducted by the International Consortium on Lithium Genetics (ConLiGen) (13, 14). The study investigated genomic data of 2,563 BD patients, identifying a genome-wide significant locus on chromosome 21, which contains two long, non-coding RNA genes (lncRNAs) (14).

Non-coding RNAs (ncRNAs) are transcribed from DNA but do not encode protein, and are involved in complex mechanisms of gene regulation, particularly in fine regulation of the timing and level of expression of their target genes. Another class of ncRNAs whose role in the pathophysiology of psychiatric disorders is emerging, is that of microRNAs (miRNAs). miRNAs are short RNA molecules, which in the mature processed form are 21 to 25-nucleotides in length, that work as post-transcriptional regulators of gene expression (15). To create a mature miRNA, a primary miRNA (pri-miRNA, typically >1,000 nucleotides in length) is first transcribed, and forms a secondary structure through self-base pairing (16, 17). This is cleaved by the Drosha-DiGeorge syndrome critical region gene 8 (Drosha-DGCR8) complex to create a pre-miRNA of around 70 nucleotides (16). This double stranded RNA is exported from the nucleus, cleaved by Dicer-transactivation-responsive RNA-binding protein (Dicer-TRBP) to form the mature miRNA (16), which can then target complementary messenger RNA (mRNA) transcripts through the RNA-induced silencing complex (RISC) to regulate expression (e.g., via mRNA degradation or translational repression) (18). Several studies have reported that miRNAs are potential predictors of treatment response in complex genetic disorders (19–21) including lithium response in BD (22). Furthermore, miRNAs are implicated in biological pathways that regulate brain development and synaptic plasticity (23, 24). Indeed, *miR-137* has emerged as a key risk gene in schizophrenia, and is known to regulate the expression of several genes that are independently associated with schizophrenia (25, 26). This implies the potential involvement of miRNAs in the pathogenesis of psychiatric disorders including BD. This hypothesis is further supported by the results of a large GWAS of BD (27) where a single-nucleotide polymorphism (SNP) flanking *miR-2113* was amongst the strongest findings.

Our group performed the first genome-wide analysis of the involvement of miRNAs in the development of BD, in a sample of 9,747 patients and 14,278 controls (28) in which we identified nine BD-associated miRNAs that withstood stringent Bonferroni-correction for multiple testing. However, it is largely unknown whether these miRNAs are also associated with lithium response in BD.

Therefore, the aim of the present study was to determine whether common variants at any of the nine BD-associated miRNAs contribute to the variance in lithium response in BD. Furthermore, we systematically analyzed whether any other miRNA is implicated in the response to lithium. For this purpose, we performed window-based association testing for all known miRNA coding genes in the largest GWAS dataset of lithium response so far.

MATERIALS AND METHODS

Sample Description

Analyses were performed using summary statistics from the previously published GWAS of lithium response in BD patients ($n = 2{,}563$ patients) (14). These GWAS datasets were collected by ConLiGen and combine imputed genotype data from 22 contributing sites from four continents (Europe, America, Asia, and Australia). The study was approved by the respective local ethics committees. Written informed consent was obtained from all participants prior to inclusion. The Alda scale was used to create a dichotomous (good vs. poor response to lithium) and a continuous measure (range 0–10) for the evaluation of long-term treatment response to lithium. Briefly, the Alda scale measures symptom improvement in the course of lithium treatment (A score, range 0–10), which is then weighted against five criteria (B score) that assess confounding factors, each scored 0, 1, or 2. The total score is calculated by subtracting the total B score from the A score. Negative scores are set to 0 by default leading to a total score range from 0 to 10. For the purpose of the present analysis, subjects with a total score of 7 or higher were defined as showing "good response" to lithium treatment in the dichotomous phenotype. As continuous measurement, we used the A score, but excluded all individuals with a total B score greater than 4, as continuous measure (14, 29).

Definition of Candidate and Genome-Wide miRNAs

Information on the nine BD-associated miRNAs was obtained from our previously published genome-wide analysis of miRNAs in BD (28). The chromosomal positions of the miRNAs were obtained from the miRBase database (release 21) (30).

For the genome-wide miRNA association analysis chromosomal positions for all 1,871 remaining miRNAs were obtained from miRBase (release 21). miRNAs which were not located on autosomal chromosomes ($n = 120$) were removed from further analysis. Only miRNA genes which were covered in the summary statistics of lithium response were included, resulting in 1,692 miRNAs which were tested in the genome-wide analysis. For each gene, the entire preprocessed transcript ± 20 kilobase (kb) flanking sequence were analyzed, which would include the majority of the regulatory regions (17).

miRNA-Based Association Tests

For the gene-based tests, we applied a set-based testing approach adapted from the versatile gene-based test for GWAS (VEGAS2) (31, 32) with a minor correction for the top-0.1-test option (33). This algorithm is obtainable upon request. The top-0.1-test was used since it showed the highest sensitivity with less than 1% false positives across a variety of investigated gene-level methods (34). The applied statistical algorithm is described in more detail in the article by Mishra and Macgregor (32). Briefly, we grouped SNPs within the miRNA loci ± 20 kb flanking sequence together and calculated a set-based test statistic as the sum of the χ^2 one degree of freedom association *P*-values within the miRNA. The observed test statistic was compared with simulated test statistics from the multivariate normal distribution with correlation equal to the corresponding LD structure as derived from the 1,000 Genomes phase 3 European population genotypes (35, 36). We calculated an empirical miRNA-based *P*-value as the proportion of simulated test statistics above the observed test statistic. For the purposes of the present study, we used the top-0.1-test option which summarizes the 10% most significant SNPs for each miRNA.

Using the two summary statistics, miRNA-based *P*-values were calculated for all miRNAs. The calculated miRNA-based *P*-values were corrected for multiple testing according to Benjamini-Hochberg.

Enrichment Tests

To test whether nominally significant SNPs were enriched within miRNAs and their flanking regions, we conducted the Fisher's Exact Test for each summary statistic separately. Additionally, we tested whether the number of cis-miR-eQTL SNPs identified by Huan et al. (37) with a *p*-value of < 0.05 was higher than expected using the Fisher's Exact Test.

RESULTS

Of the nine tested BD-associated miRNAs, *miR-499a* showed nominally significant *P*-values in both datasets (dichotomous and continuous treatment response, **Table 1**). Of the remaining 1,692 miRNAs tested for the genome-wide miRNA analysis, 71 miRNAs showed nominally significant associations with the dichotomous and 106 with the continuous treatment response. Fifteen miRNAs revealed nominal significance with both phenotypes. *miR-633* showed the strongest association with the continuous phenotype ($p = 9.80\text{E-}04$). Regarding the dichotomous phenotype, *miR-607* showed the strongest association ($p = 5.79\text{E-}04$). No association between miRNAs and treatment response to lithium in BD in either of the tested conditions withstood multiple testing correction (**Tables 1**, **2**).

The number of nominally significant SNPs in both of our GWAS of lithium response located at miRNA loci ($n = 6{,}321$ and $n = 5{,}742$ for continuous and dichotomous measurement, respectively) was not significantly higher than expected ($p = 9.96\text{E-}01$ and $p = 1$ for continuous and dichotomous measurement, respectively, Fisher's Exact Test).

The number of cis-miR-eQTL SNPs [identified by Huan et al. (37)] in our summary statistics ($n = 341$ and $n = 318$ for continuous and dichotomous measurement, respectively) were significantly higher than expected ($p = 3.31\text{E-}05$ and $p = 6.23\text{E-}03$ for continuous and dichotomous measurement, respectively, Fisher's Exact Test).

DISCUSSION

The current study investigated whether common variants at BD-associated miRNA, or any other miRNA loci, contribute to the differences in lithium response in BD patients.

miR-499a showed a nominally significant association with lithium response in the candidate approach. Although the association did not withstand correction for multiple testing, this result provides some evidence that *miR-499a* might be involved

TABLE 1 | Results of the window-based tests for the nine BD-associated miRNAs.

miRNA [position]	n SNPs	p miRNA	p_{corr} miRNA	Top SNP [position]	p Top SNP
A) Dichotomous Treatment Response Measure					
miR-499a [chr20:34990376-34990497]	87	1.71E-02	9.31E-01	rs117616040 [chr20:34977725]	2.27E-03
miR-135a-1 [chr3:52294219-52294308]	41	9.29E-02	9.31E-01	rs699465 [chr3:52276426]	1.98E-02
let-7g [chr3:52268278-52268361]	44	1.04E-01	9.31E-01	rs699465 [chr3:52276426]	1.98E-02
miR-644a [chr20:34466325-34466418]	35	2.03E-01	9.37E-01	rs7266300 [chr20:34449603]	9.76E-03
miR-708 [chr11:79402022-79402109]	130	3.51E-01	9.50E-01	rs12275848 [chr11:79416285]	7.89E-02
miR-1908 [chr11:61815161-61815240]	39	4.33E-01	9.50E-01	rs61897792 [chr11:61819414]	3.46E-02
miR-640 [chr19:19435063-19435158]	44	6.08E-01	9.55E-01	rs79954596 [chr19:19437834]	1.74E-01
B) Continuous Treatment Response Measure					
miR-499a [chr20:34990376-34990497]	87	3.18E-02	7.52E-01	rs117616040 [chr20:34977725]	3.73E-03
miR-708 [chr11:79402022-79402109]	130	7.37E-02	8.02E-01	rs1355423 [chr11:79420556]	2.05E-02
miR-611 [chr11:61792495-61792561]	33	9.79E-02	8.22E-01	rs174532 [chr11:61781402]	3.30E-03
miR-644a [chr20:34466325-34466418]	35	1.95E-01	8.50E-01	rs7266300 [chr20:34449603]	6.47E-02
miR-1908 [chr11:61815161-61815240]	39	2.45E-01	8.64E-01	rs968567 [chr11:61828092]	3.20E-02
miR-640 [chr19:19435063-19435158]	45	2.69E-01	9.52E-01	rs79954596 [chr19:19437834]	1.15E-02
let-7g [chr3:52268278-52268361]	44	6.40E-01	9.81E-01	rs58315325 [chr3:52261812]	2.80E-02
miR-135a-1 [chr3:52294219-52294308]	41	7.81E-01	9.86E-01	rs34135146 [chr3:52279416]	6.28E-02
miR-581 [chr5:53951504-53951599]	55	9.22E-01	7.52E-01	rs697112 [chr5:53964849]	1.61E-01

microRNAs (miRNAs) are sorted according to their miRNA-based P-value. Genome build used is GRCh38 (hg38). Abbreviations: miRNA, microRNA; position, genomic position; nSNPs, number of investigated SNPs; p miRNA, miRNA-based P-value; p_{corr} miRNA, Benjamini-Hochberg corrected miRNA-based P-value; Top SNP, top single-nucleotide polymorphism within gene; p Top SNP, P-value of the Top SNP within gene.

in both development and treatment of BD. A previous study has shown an upregulation of this miRNA in the prefrontal cortex of patients with depression (38). In another study, *miR-499a* was differentially expressed in the postmortem brains of BD patients compared with controls (39). Furthermore, a recent study by Banach et al. (40) reported lower expression levels of *miR-499* in the peripheral blood of BD patients during depressive episodes in comparison to remission, suggesting *miR-499* as a potential new biomarker of illness state in BD.

Overall, the results of our candidate approach do not suggest that individual BD-associated miRNAs might have a strong influence on differential responses to lithium treatment in BD as no association withstood multiple testing correction. On the one hand, this might at least in part reflect that the power to detect associations between common variants and lithium response was limited in the present study, even though the ConLiGen GWAS comprised several thousand individuals (41). On the other hand, it might also indicate that the genetic factors that contribute to BD etiology are different from those contributing to treatment response or illness course. That there are such effects in multifactorial diseases is supported by a study in ulcerative colitis in which no SNPs from 163 inflammatory bowel disease susceptibility loci (42) were found to be associated with the disease course (43).

TABLE 2 | Results of the window-based tests for the top five genome-wide miRNAs.

miRNA [position]	n SNPs	p miRNA	p_{corr} miRNA	Top SNP [position]	p Top SNP
A) Dichotomous Treatment Response Measure					
miR-607 [chr10:96828669-96828764]	48	5.79E-04	9.31E-01	rs111682442 [chr10:96823685]	2.73E-04
miR-8085 [chr19:44758657-44758721]	44	2.54E-03	9.31E-01	rs7249244 [chr19:44742441]	1.11E-04
miR-1296 [chr10:63372957-63373048]	62	4.13E-03	9.31E-01	rs10995527 [chr10:63387659]	5.47E-03
B) Continuous Treatment Response Measure					
miR-633 [chr17:62944215-62944312]	29	9.80E-04	7.52E-01	rs1588368 [chr17:62938848]	1.21E-04
miR-6516 [chr17:77089417-77089497]	183	1.97E-03	8.02E-01	rs2411054 [chr17:77074245]	2.68E-05
miR-218-1 [chr4:20528275-20528384]	105	2.13E-03	8.22E-01	rs540146 [chr4:20544433]	6.08E-04
miR-7704 [chr2:176188843-176188901]	68	2.20E-03	8.50E-01	rs7589870 [chr2:176208720]	1.26E-03
miR-548e [chr10:110988926-110989013]	54	2.48E-03	8.64E-01	rs1327551 [chr10:111008438]	1.09E-03

microRNAs (miRNAs) are sorted according to their miRNA-based P-value. Genome build used is GRCh38 (hg38). Abbreviations: miRNA, microRNA; position, genomic position; nSNPs, number of investigated SNPs; p miRNA, miRNA-based P-value; p_{corr} miRNA, Benjamini-Hochberg corrected miRNA-based P-value; Top SNP, top single-nucleotide polymorphism within gene; p Top SNP, P-value of the Top SNP within gene.

In our systematic, genome-wide analysis of miRNAs, 106 miRNAs revealed nominally significant associations with the continuous and 71 with the dichotomous lithium treatment response.

The intergenic *miR-633* located on chromosome 17 showed the strongest association with the continuous phenotype ($p = 9.80E-04$). To date, few published studies have investigated the function of *miR-633*. Interestingly, one study reported that *miR-633* was differentially regulated in the cerebrospinal fluid of patients with multiple sclerosis compared to patients with other neurologic diseases. In addition, *miR-633* differentiated relapsing-remitting from secondary progressive multiple sclerosis courses suggesting this miRNA as a potential biomarker for disease course in multiple sclerosis (44).

miR-607, an intergenic miRNA located upstream of the ligand dependent nuclear receptor corepressor (*LCOR*) gene on chromosome 10, displayed the strongest association with the dichotomous treatment response measure ($p = 5.79E-04$). The function of this miRNA has been poorly characterized so far, so that we cannot currently speculate about possible disease- and treatment-relevant biological processes. Further research is needed to elucidate the potential role of *miR-607* in health or disease.

No association between miRNAs and BD treatment response to lithium in either of the tested conditions withstood multiple testing correction. In addition, we did not observe a significant enrichment for SNPs at all microRNA loci in the present study. Given the limited power of our study, future investigation of miRNAs in larger GWAS samples of BD and lithium response is warranted as better understanding of genetic factors contributing to disease etiology and treatment response might enable the individualization of treatment as well as the identification of novel therapeutic targets (45).

In the present study, we investigated all currently known miRNAs regardless of their tissue or developmental expression patterns. Approximately 70% of ncRNAs are thought to be brain expressed (23) and are dynamically regulated during development and over the lifespan. While the exact mechanisms by which lithium exerts its therapeutic effects remain unclear, pharmacokinetics and pharmacodynamics highlight the importance of specific tissues (e.g., brain and kidney) in treatment responsiveness (46, 47). Therefore, an analysis including miRNAs expressed specifically in these tissues would seem to be a rational follow-up step to reduce the multiple testing burden and to narrow-down the miRNAs to those that *a priori* may have a greater chance to be involved in lithium response. Unfortunately, a systematic enrichment analyses for miRNAs in particular tissues would be premature, since there are currently no comprehensive expression databases derived from normal tissue covering all known miRNAs investigated in the present study. Data on miRNA expression at various developmental stages is also still limited, as non-polyadenylated transcripts are typically not captured with standard library preparation for RNA sequencing. Furthermore, some miRNAs may only be expressed

during early developmental stages but can still have an influence on lithium response later on in life, particularly if expression is induced by pharmaceutical treatment. Nevertheless, these aspects remain important and should be considered in future analyses as soon as more comprehensive data on miRNA expression become available.

Using the present approach, we were not able to investigate SNPs with trans-expression quantitative trait loci (eQTL) effects on miRNAs. Previous studies suggest that a substantial proportion of the identified miR-eQTLs are trans-eQTLs (48). Therefore, future investigations into the molecular interactions underlying the association between miRNA trans-eQTLs and treatment response to lithium in BD are also warranted. Huan et al. (37) conducted a genome-wide miR-eQTL mapping study and found consistent evidence for 5,269 cis-miR-eQTLs for 76 mature microRNAs. The significant enrichment for cis-miR-eQTL SNPs found in our summary statistics provides some evidence for the importance of cis-miR-eQTLs in lithium response, although we were not able to identify cis-miR-eQTL SNPs in our top findings since those miRNAs were not among the 76 mature microRNAs reported by Huan et al. (37).

Moreover, miRNAs only represent one class of non-coding RNAs. In the ConLiGen GWAS a genome-wide significant locus containing two lncRNAs was identified (14). Further analyses on the contribution of lncRNAs to lithium response are therefore warranted. This was beyond the scope of the present analysis as the current understanding of the predicted structure of lncRNA molecules and their biological functions remains limited (49).

In conclusion, our analyses do not provide strong evidence that miRNAs are involved in individual response to lithium treatment in BD, as no association between miRNAs and lithium response withstood multiple testing correction. Our data should still be interesting for follow-up of independent studies, particularly when sufficient data is available to accurately define the tissue and temporal expression profile of all human miRNAs, which would allow a more targeted analysis of brain-expressed miRNAs, thereby reducing the search space to miRNAs with relevant expression profiles. We did not find any strong effect that could be useful in terms of a personalized treatment for individual patients. This does not exclude a possible (small) effect of miRNAs on lithium response, and further independent and even larger studies should be envisaged to clarify this question. In parallel, the investigation of other biological mechanisms possibly contributing to lithium treatment response may provide insights for individualizing future pharmacotherapy in BD.

AUTHOR CONTRIBUTIONS

CR, AF, MN, and SC contributed to the conception and design of the study. AF, LH, UH, FD, MA, KA, NA, RA, BA, LB, AB, SB, AKB, JB, ABi, CM-C, PC, G-BC, H-CC, CC, SRC, FC, DC, CCr, PMC, AD, BE, PF, LF, JF, SG, JG, FG, PG, OG, RH, JHa, SH, SJ, EJ, J-PK, LK, SK-S, SK, BK, IK, NL, GL, ML, CL, MLe, SL, CAL, GM, MM, LM, MMa, MJM, SM, MMi, FM, PM, CN, UÖ, NO, RP, AP, DR-E, GR, PRS, KOS, BS, FS, GS, TS, PDS, KS, CS, CMS, JS, AS, TSt, PS, ST, AT, GT, JV, SW, AW, LY, PZ, JP, JD, MB, ER, TN, J-MA, MMaj, BB, PM, EV, MF, JR, P-HK, TK, MG-S, AR, MD, FB, MS, NW, JK, MAl, FM, TGS, MR, and MN recruited the patients and contributed genotype data. CR and JH performed the statistical analysis. CR, AF, MR, MN, and SC prepared the manuscript, with feedback from the other authors. All authors contributed to manuscript revision, read and approved the submitted version.

FUNDING

The study was supported by the German Federal Ministry of Education and Research (BMBF) through the Integrated Network IntegraMent (Integrated Understanding of Causes and Mechanisms in Mental Disorders), under the auspices of the e:Med Programme (grant 01ZX1314A/01ZX1614A to MN and SC, grant 01ZX1314G/01ZX1614G to MR). The study was also supported by the German Research Foundation (DFG; grant FOR2107; RI908/11-1 to MR; NO246/10-1 to MN), and the Swiss National Science Foundation (SNSF, grant 310030_156791 to SC). MMN is a member of the DFG-funded Excellence-Cluster ImmunoSensation. AF received support from the BONFOR program of the Medical Faculty of the University of Bonn. The collection of the Geneva sample was supported by the Swiss National Foundation (grants Synapsy 51NF40-158776 and 32003B-125469). The Australian sample contribution was supported by the Australian National Health and Medical Research Council (grants 1037196, 1063960) and the Janette Mary O'Neil Research Fellowship to JF.

ACKNOWLEDGMENTS

We are grateful to all patients and control subjects for their participation. We thank the International Consortium on Lithium Genetics (ConLiGen) for providing access to the relevant data.

REFERENCES

1. CraddockN, Sklar P. Genetics of bipolar disorder. *Lancet* (2013) **381**:1654–62. doi: 10.1016/S0140-6736(13)60855-7
2. GBD 2016 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet* (2017) **390**:1211–59. doi: 10.1016/s0140-6736(17)32154-2
3. Lichtenstein P, Yip BH, Björk C, Pawitan Y, Cannon TD, Sullivan PF, et al. Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet* (2009) **373**:234–9. doi: 10.1016/S0140-6736(09)60072-6
4. McGuffin P, Rijsdijk F, Andrew M, Sham P, Katz R, Cardno A. The heritability of bipolar affective disorder and the genetic relationship to unipolar depression. *Arch Gen Psychiatry* (2003) **60**:497–502. doi: 10.1001/archpsyc.60.5.497
5. Edvardsen J, Torgersen S, Røysamb E, Lygren S, Skre I, Onstad S, et al. Heritability of bipolar spectrum disorders. Unity or heterogeneity? *J Affect Disord.* (2008) **106**:229–40. doi: 10.1016/j.jad.2007.07.001
6. Geddes JR, Miklowitz DJ. Treatment of bipolar disorder. *Lancet* (2013) **381**:1672–82. doi: 10.1016/S0140-6736(13)60857-0

7. Geddes JR, Burgess S, Hawton K, Jamison K, Goodwin GM. Long-term lithium therapy for bipolar disorder: systematic review and meta-analysis of randomized controlled trials. *Am. J. Psychiatry* (2004) **161**:217–22. doi: 10.1176/appi.ajp.161.2.217
8. Baldessarini RJ, Tondo L, Hennen J. Lithium treatment and suicide risk in major affective disorders: update and new findings. *J Clin Psychiatry* (2003) **64**:(Suppl. 5), 44–52.
9. Cipriani A, Pretty H, Hawton K, Geddes JR. Lithium in the prevention of suicidal behavior and all-cause mortality in patients with mood disorders: a systematic review of randomized trials. *Am J Psychiatry* (2005) **162**:1805–19. doi: 10.1176/appi.ajp.162.10.1805
10. Garnham J, Munro A, Slaney C, MacDougall M, Passmore M, Duffy A, et al. Prophylactic treatment response in bipolar disorder: results of a naturalistic observation study. *J Affect Disord.* (2007) **104**:185–90. doi: 10.1016/j.jad.2007.03.003
11. Grof P, Duffy A, Cavazzoni P, Grof E, Garnham J, MacDougall M, et al.. Is response to prophylactic lithium a familial trait? *J Clin Psychiatry* (2002) **63**:942–7. doi: 10.4088/JCP.v63n1013
12. Kleindienst N, Engel R, Greil W. Which clinical factors predict response to prophylactic lithium? A systematic review for bipolar disorders. *Bipolar Disord.* (2005) **7**:404–17. doi: 10.1111/j.1399-5618.2005.00244.x
13. Schulze TG, Alda M, Adli M, Akula N, Ardau R, Bui ET, et al. The International Consortium on Lithium Genetics (ConLiGen): an initiative by the NIMH and IGSLI to study the genetic basis of response to lithium treatment. *Neuropsychobiology* (2010) **62**:72–8. doi: 10.1159/000314708
14. 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**:1085–93. doi: 10.1016/S0140-6736(16)00143-4
15. Godnic I, Zorc M, Jevsinek Skok D, Calin GA, Horvat S, Dovc P, et al. Genome-wide and species-wide in silico screening for intragenic MicroRNAs in human, mouse and chicken. *PLoS ONE* (2013) **8**:e65165. doi: 10.1371/journal.pone.0065165
16. Krol J, Loedige I, Filipowicz W. The widespread regulation of microRNA biogenesis, function and decay. *Nat Rev Genet.* (2010) **11**:597–610. doi: 10.1038/nrg2843
17. Wang G, Wang Y, Shen C, Huang Y-W, Huang K, Huang TH, et al. RNA polymerase II binding patterns reveal genomic regions involved in MicroRNA gene regulation. *PLoS ONE* (2010) **5**:e13798. doi: 10.1371/journal.pone.0013798
18. Gregory RI, Chendrimada TP, Cooch N, Shiekhattar R. Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell* (2005) **123**:631–40. doi: 10.1016/j.cell.2005.10.022
19. Mahdian-Shakib A, Dorostkar R, Tat M, Hashemzadeh MS, Saidi N. Differential role of microRNAs in prognosis, diagnosis, and therapy of ovarian cancer. *Biomed Pharmacother.* (2016) **84**:592–600. doi: 10.1016/j.biopha.2016.09.087
20. Campos-Parra AD, Mitznahuatl GC, Pedroza-Torres A, Romo RV, Reyes FIP, Lopez-Urrutia E, et al. Micro-RNAs as potential predictors of response to breast cancer systemic therapy: future clinical implications. *Int J Mol Sci.* (2017) **18**:E1182. doi: 10.3390/ijms18061182
21. Liu Q, Wu DH, Han L, Deng JW, Zhou L, He R, et al. Roles of microRNAs in psoriasis: immunological functions and potential biomarkers. *Exp Dermatol.* (2017) **26**:359–67. doi: 10.1111/exd.13249
22. Hunsberger JG, Chibane FL, Elkahloun AG, Henderson R, Singh R, Lawson J, et al. Novel integrative genomic tool for interrogating lithium response in bipolar disorder. *Transl Psychiatry* (2015) **5**:e504. doi: 10.1038/tp.2014.139
23. Fineberg SK, Kosik KS, Davidson BL. MicroRNAs potentiate neural development. *Neuron* (2009) **64**:303–9. doi: 10.1016/j.neuron.2009.10.020
24. Schratt G. microRNAs at the synapse. *Nat Rev Neurosci.* (2009) **10**:842–9. doi: 10.1038/nrn2763
25. The Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium. Genome-wide association study identifies five new schizophrenia loci. *Nat Genet.* (2011) **43**:969. doi: 10.1038/ng.940
26. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* (2014) **511**:421–7. doi: 10.1038/nature13595
27. 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. doi: 10.1038/ncomms4339
28. Forstner AJ, Hofmann A, Maaser A, Sumer S, Khudayberdiev S, Mühleisen TW, et al. Genome-wide analysis implicates microRNAs and their target genes in the development of bipolar disorder. *Transl Psychiatry* (2015) **5**:e678. doi: 10.1038/tp.2015.159
29. Manchia M, Adli M, Akula N, Ardau R, Aubry J.-M, Backlund L, et al. Assessment of response to lithium maintenance treatment in bipolar disorder: a Consortium on Lithium Genetics (ConLiGen) Report. *PLoS ONE* (2013) **8**:e65636. doi: 10.1371/journal.pone.0065636
30. Kozomara A, Griffiths-Jones S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.* (2014) **42**:D68–73. doi: 10.1093/nar/gkt1181
31. Liu JZ, McRae AF, Nyholt DR, Medland SE, Wray NR, Brown KM, et al. A versatile gene-based test for genome-wide association studies. *Am J Hum Genet.* (2010) **87**:139–45. doi: 10.1016/j.ajhg.2010.06.009
32. Mishra A, Macgregor S. VEGAS2: Software for more flexible gene-based testing. *Twin Res Hum Genet.* (2015) **18**:86–91. doi: 10.1017/thg.2014.79
33. Hecker J, Maaser A, Prokopenko D, Fier HL, Lange C. Reporting correct p values in VEGAS analyses. *Twin Res Hum Genet.* (2017) **20**:257–9. doi: 10.1017/thg.2017.16
34. Wojcik GL, Kao WHL, Duggal P. Relative performance of gene- and pathway-level methods as secondary analyses for genome-wide association studies. *BMC Genet.* (2015) **16**:34. doi: 10.1186/s12863-015-0191-2
35. Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, et al. A global reference for human genetic variation. *Nature* (2015) **526**:68–74. doi: 10.1038/nature15393
36. Sudmant PH, Rausch T, Gardner EJ, Handsaker RE, Abyzov A, Huddleston J, et al. An integrated map of structural variation in 2,504 human genomes. *Nature* (2015) **526**:75. doi: 10.1038/nature15394
37. Huan T, Rong J, Liu C, Zhang X, Tanriverdi K, Joehanes R, et al. Genome-wide identification of microRNA expression quantitative trait loci. *Nat Commun.* (2015) **6**:6601. doi: 10.1038/ncomms7601
38. Smalheiser NR, Lugli G, Zhang H, Rizavi H, Cook EH, Dwivedi Y. Expression of microRNAs and other small RNAs in prefrontal cortex in schizophrenia, bipolar disorder and depressed subjects. *PLoS ONE* (2014) **9**:e86469. doi: 10.1371/journal.pone.0086469
39. Banigan MG, Kao PF, Kozubek JA, Winslow AR, Medina J, Costa J, et al. Differential expression of exosomal microRNAs in prefrontal cortices of schizophrenia and bipolar disorder patients. *PLoS ONE* (2013) **8**:e48814. doi: 10.1371/journal.pone.0048814
40. Banach E, Dmitrzak-Weglarz M, Pawlak J, Kapelski P, Szczepankiewicz A, Rajewska-Rager A, et al. (2017). Dysregulation of miR-499, miR-708 and miR-1908 during a depression episode in bipolar disorders. *Neurosci Lett.* **654**(Suppl. C):117–9. doi: 10.1016/j.neulet.2017.06.019
41. Sullivan PF, Daly MJ, O'Donovan M. Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nat Rev Genet.* (2012) **13**:537–51. doi: 10.1038/nrg3240
42. Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP, Hui KY, et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* (2012) **491**:119–24. doi: 10.1038/nature11582
43. Kopylov U, Boucher G, Waterman M, Rivers CR, Patel M, Cho JH, et al. Genetic predictors of benign course of ulcerative colitis-a north American inflammatory bowel disease genetics Consortium Study. *Inflamm Bowel Dis.* (2016) **22**:2311–6. doi: 10.1097/mib.0000000000000913
44. Haghikia A, Haghikia A, Hellwig K, Baraniskin A, Holzmann A, Decard BF, et al. Regulated microRNAs in the CSF of patients with multiple sclerosis: a case-control study. *Neurology* (2012) **79**:2166–70. doi: 10.1212/WNL.0b013e3182759621
45. Bevilacqua L, Goldman D. Genes and addictions. *Clin Pharmacol. Ther.* (2009) **85**:359–61. doi: 10.1038/clpt.2009.6
46. Mason RW, McQueen EG, Keary PJ, James NM. Pharmacokinetics of lithium: elimination half-time, renal clearance and apparent volume of distribution in schizophrenia. *Clin Pharmacokinet.* (1978) **3**:241–6.
47. Brown KM, Tracy DK. Lithium: the pharmacodynamic actions of the amazing ion. *Ther Adv. Psychopharmacol.* (2013) **3**:163–76. doi: 10.1177/2045125312471963

48. Borel C, Deutsch S, Letourneau A, Migliavacca E, Montgomery SB, Dimas AS, et al. Identification of cis- and trans-regulatory variation modulating microRNA expression levels in human fibroblasts. *Genome Res.* (2011) **21**:68–73. doi: 10.1101/gr.109371.110
49. Zhao XY, Lin JD. Long noncoding RNAs: a new regulatory code in metabolic control. *Trends Biochem Sci.* (2015) **40**:586–96. doi: 10.1016/j.tibs.2015.08.002

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 09 February 2018; Accepted: 03 May 2018; Published: xx May 2018

Citation: Reinbold CS, Forstner AJ, Hecker J, Fullerton JM, Hoffmann P, Hou L, Heilbronner U, Degenhardt F, Adli M, Akiyama K, Akula N, Ardau R, Arias B, Backlund L, Benabarre A, Bengesser S, Bhattacharjee AK, Biernacka JM, Birner A, Marie-Claire C, Cervantes P, Chen G-B, Chen H-C, Chillotti C, Clark SR, Colom F, Cousins DA, Cruceanu C, Czerski PM, Dayer A, Étain B, Falkai P, Frisén L, Gard S, Garnham JS, Goes FS, Grof P, Gruber O, Hashimoto R, Hauser J, Herms S, Jamain S, Jiménez E, Kahn J-P, Kassem L, Kittel-Schneider S, Kliwicki S, König B, Kusumi I, Lackner N, Laje G, Landén M, Lavebratt C, Leboyer M, Leckband SG, López Jaramillo CA, MacQueen G, Manchia M, Martinsson L, Mattheisen M, McCarthy MJ, McElroy SL, Mitjans M, Mondimore FM, Monteleone P, Nievergelt CM, Ösby U, Ozaki N, Perlis RH, Pfennig A, Reich-Erkelenz D, Rouleau GA, Schofield PR, Schubert KO, Schweizer BW, Seemüller F, Severino G, Shekhtman T, Shilling PD, Shimoda K, Simhandl C, Slaney CM, Smoller JW, Squassina A, Stamm TJ, Stopkova P, Tighe SK, Tortorella A, Turecki G, Volkert J, Witt SH, Wright AJ, Young LT, Zandi PP, Potash JB, DePaulo JR, Bauer M, Reininghaus E, Novák T, Aubry J-M, Maj M, Baune BT, Mitchell PB, Vieta E, Frye MA, Rybakowski JK, Kuo P-H, Kato T, Grigoroiu-Serbanescu M, Reif A, Del Zompo M, Bellivier F, Schalling M, Wray NR, Kelsoe JR, Alda M, McMahon FJ, Schulze TG, Rietschel M, Nöthen MM and Cichon S (2018) Analysis of the Influence of microRNAs in Lithium Response in Bipolar Disorder. Front. Psychiatry 9:207. doi: 10.3389/fpsyt.2018.00207
This article was submitted to Behavioral and Psychiatric Genetics, a section of the journal Frontiers in Psychiatry.

5 General Discussion

Important advances in psychiatric genetics have been made in the recent years, with many replicated discoveries of common, rare and de novo variants that are converging on specific pathways and biological mechanisms. These successes predominantly result from the foundation of international consortia and their combined efforts in leveraging resources. Besides getting a better understanding of the genetic architecture these efforts also resulted in the development of novel bioinformatics tools. These improved tools were developed to cope with the limitations inherent to GWAS and ultimately unravel the complete molecular genetic basis of complex disorders. The studies described in this thesis aim to deepen our understanding of psychiatric disorders by the application of different bioinformatics tools and biological information to already existing GWAS data.

A complete portrait of the genetic architecture for any psychiatric disorder (or even any complex disease) does not yet exist. Gaining a more complete knowledge of the genetic contributors will therefore be of exceptional importance. To achieve this, respect must be paid to both the phenotypic and the genotypic heterogeneity. This is of particular relevance for psychiatric disorders, since symptoms are self-reported, differing assessment instruments are used, and comorbidities are complicating the clear definition of the phenotype (Breen et al., 2016). Furthermore, the misclassification of a phenotype, especially in case-control studies, has been shown to dramatically reduce the power to detect effects (Edwards, Haynes, Levenstien, Finch, & Gordon, 2005; Manchia, Cullis, et al., 2013). This is particularly true for BD, where the range of symptoms is diverse and overlapping with other disorders such as schizophrenia or major depressive disorder (American Psychiatric Association, 2013a). Despite the fact that most researchers disagree with the dichotomous concept of the established diagnostic and statistical manuals, still no consensus has been found on how to improve it (Angst, 2007). Promising approaches to circumvent the phenotypic heterogeneity in psychiatric disorder studies and to define more homogeneous etiological subgroups are to consider biology-derived phenotypic aspects, such as response to drug treatment or endophenotypes, deeply phenotyped samples or extreme group comparisons which all already yielded successes (Gershon et al., 2018; Gottesman & Gould, 2003; Ibrahim-Verbaas et al., 2016; Manchia, Adli, et al., 2013; Peloso et al., 2016; Riglin et al., 2016; Zabaneh et al., 2017).

The observed heterogeneity of any psychiatric disorder does not only manifest on a phenomenological level but also in the genome. Already early studies revealed that a familial overlap among different disorders was present (Kendler et al., 2011). Nowadays, with the aid of systematic and well-powered genetic studies, such as GWAS, the extensive cross-disorder heritability and high genetic correlations among some disorders has been established (Anttila et al., 2017; Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013; Witt et al., 2017; Wray et al., 2018). Furthermore, GWAS resulted in a plethora of SNPs associated with common disorders. However, the effect size of these SNPs is small, and the individual variant is neither necessary nor sufficient to cause the disorder. At this point it should be kept in mind that GWAS are by design best-powered to detect associations with variants that are common in the population (Hirschhorn & Daly, 2005). Due to the relatively small effects sizes of the associated variants, it soon became clear that large sample sizes are needed to reliably detect susceptibility genes. The combined efforts, the data-sharing mentality and the ever-increasing sample sizes are just some of the advantages of consortia (Psychiatric GWAS Consortium Steering Committee, 2009). Future studies investigating common and rare variants will be based on even larger sample sizes and as empirical evidence and simulations for GWAS have shown, after a certain samples size has been reached, the number of genome-wide significant loci will increase linearly (Levinson et al., 2014). Even if the approach of ever-growing sample sizes without deep phenotypic information has been widely criticized, it will definitely help unraveling the genetic architecture of disorder-specific and cross-disorder effects (Sullivan et al., 2017). The increase in sample sizes will be continued until all most important biological pathways involved in the respective trait/disorder will have been identified. However, some limitations are inherent to GWAS and cannot be conquered by increasing sample size alone. SNPs identified by GWAS usually don't identify the causal allele or gene itself, more likely the locus implicated several genes within the region. In fact, the functional effect of the GWAS hits is rarely understood, and the variant often have a regulatory effect on a gene outside the risk locus. The common SNPs found by GWAS have been estimated to explain only part, albeit a sometimes large part, of the phenotypic heritability for psychiatric disorders. Consequently, it could be argued that the so often discussed "missing heritability" is actually more likely to be hidden (Eichler et al., 2010; Manolio et al., 2009; Yang et al., 2010).

The discovered polygenicity of psychiatric disorders resulted in the assumption that the complex genetic architecture underlying mental disorders is based on sets of functionally related genes rather than single independent variants. It is hypothesized, that the investigation of these gene-sets will not only yield in a better understanding of the disorder but also in improved treatment options (Breen et al., 2016; Smoller et al., 2018; Sullivan et al., 2012). Based on these assumptions, a wealth of methods that leverage GWAS by implementation of biological information was published over the last years. Gene-set analysis methods are among the most frequently used novel methods since they not only allow the investigation of the joint effects of SNPs but also their biological interpretation. The last years have shown that the development and accurate application of reliable analysis methods can lead to an enormous increment of significant results and enable the in-silico investigation of functional mechanisms underlying complex disorders. Even though bioinformatics tools have become more user-friendly and consequently open to more researchers, it is indispensable to understand the limitations and prerequisites of the methods applied since the over-interpretation of their results may lead to deceptive results and waste of time, money and effort in (functional) follow-up attempts. However, it is not always easy to decide which method is best-suited as many, especially older tools are poorly explained.

The reported studies within this thesis aim to contribute to the field of psychiatric genetics by leveraging results from (publicly available) GWAS through applying biology informed methods. In the study by Forstner et al. (2017), we systematically investigated whether genome-wide significant loci associated with schizophrenia also contribute to the development of bipolar disorder. This study is an excellent example of the differences arising from comparing publicly available summary statistics. Firstly, even though imputation of summary statistics was done to circumvent the fact that different panels for the studies were used in the first place, still not all schizophrenia-associated loci could be investigated in the bipolar data set. Secondly, a complicated correction algorithm for the possible sample overlap (in particular for the control samples) was applied because neglecting this issue can result in inflated false positive rates (Zhu, Anttila, Smoller, & Lee, 2018). Since this study only investigated candidate SNPs, merely self-contained methods were appropriate to investigate the combined effect of the resulting SNPs associated with both disorders. Interestingly, results of both methods identified gene-sets described in earlier studies important in fundamental neuronal processes and human diseases such as calcium channel activity or

glutamate receptor signaling (Nurnberger et al., 2014; Ripke et al., 2013; Sklar et al., 2011). These findings are consistent with the previous reported genetic overlap between schizophrenia and bipolar disorder not just on the single SNP-level but also on the gene-set-level. Even though the combination of different GSA algorithms and databases is recommended, it is important to be aware of the specifics each method entails and interpret the results accordingly.

In the study by Witt et al. (2017), we took full advantage of a well-established genome-wide analysis pipeline (https://github.com/Nealelab/ricopili) not only to systematically investigate borderline personality disorder but also its genetic overlaps with other psychiatric disorders. This is of particular interest since borderline personality disorder and BD share some of the symptoms and the potential comorbidity between these two disorders is an ongoing debate (Fornaro et al., 2016). Since the estimated heritability of borderline personality disorder is limited and the genetic architecture complex, single marker analysis was unlikely to generate significant results with our sample size. Therefore, gene-level and gene-set analysis were conducted. To enhance the interpretability of the GSA results, we based our analysis on GO-terms solely but replicated the top finding with two independent methods. This resulted in a robust association with the gene-set called exocytosis. In neuronal synapses, exocytosis is triggered by an influx of calcium and critically underlies synaptic signaling. Dysregulated neuronal signaling and exocytosis are core features of psychiatric disorders (e.g. autism spectrum disorders, intellectual disability, BIP, SCZ and MDD) (Cupertino et al., 2016; Pescosolido, Gamsiz, Nagpal, & Morrow, 2013; Sullivan et al., 2008; Zhao et al., 2015). Furthermore, significant genetic correlation was found between borderline personality disorder and BD as well as a significant correlation between a genetic risk score for BD with borderline personality disorder. Since the investigation of single markers did not yield significant associations, this study further supports the idea that only the interpretation of the joint effect of SNPs will result in meaningful results. However, the results must be interpreted with caution since the sample size was small in relation to the estimated heritability.

The study by Mühleisen, Reinbold et al. (2017) was based on the largest sample of BD patients at that time and aimed to extract more biological information by applying GSA tools and explored the implicated genes for expression. The implicated gene-sets themselves revealed novel insights into the etiology of BD, for example, we found a pathway involved in the promotion of cell proliferation, survival, and differentiation, not

only in the brain. But it is interesting to mention that plenty of the underlying genes have been previously reported to be associated with psychiatric disorders (Cichon et al., 2011; Hou et al., 2016; Sklar et al., 2008). However, this study again is a good example why it is important to understand not only the algorithms applied but also the databases used since one of the most replicated findings in BD (*CACNA1C*) was not even present in the chosen database and therefore had no chance of being found. However, these results further support the hypothesis that the genetic underpinnings of psychiatric disorders are more likely to function as a set than on a single variant basis.

The last study reported within the framework of this thesis systematically investigated the influence of microRNAs in lithium response in BD. Lithium is the best-established long-term treatment for BD, even though individual response is highly variable (Baldessarini, Tondo, & Hennen, 2003; Garnham et al., 2007; Geddes & Miklowitz, 2013). The main aim of the study was to investigate whether common variants associated with BD also influence the treatment response to lithium. Despite the hypothesis-driven approach, no BD-associated microRNA revealed a statistically significant association with lithium response. Furthermore, no association between any microRNA and treatment response to lithium withstood multiple testing correction. This was surprising since evidence from literature pointed not only to a potential link of microRNAs and treatment response in various complex disorders but also to an involvement in brain development and psychiatric disorders (Campos-Parra et al., 2017; Fineberg, Kosik, & Davidson, 2009; Hunsberger et al., 2015; Q. Liu et al., 2017; Mühleisen et al., 2014; Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014). These results suggest that the genetic factors that contribute to BD are different from those contributing to treatment response or illness course. Furthermore, the self-contained analysis confirmed that no significant enrichment for SNPs at all microRNA loci was observed. However, an important limiting factor in the accomplishment and interpretation of this study represents the scarcity of microRNA-specific information such as expression profiles.

All the studies reported within this framework aimed to shed additional light on the complex genetic architecture underlying psychiatric disorders, and bipolar disorder specifically. When comparing the outcomes of the four studies described herein, it becomes apparent that each bioinformatics tool has its merits but the shear amount of methods available and the often scarcely described parameters applied make it difficult

to reliably compare results between studies. Further, it becomes clear that only the integrative investigation of all genetic variants together will help to discover the disease pathomechanisms. Therefore, not only robust methods and computational resources but also large and deeply phenotyped samples will be required for future studies.

6 References

Alkan, C., Coe, B. P., & Eichler, E. E. (2011). Genome structural variation discovery and genotyping. *Nature Reviews Genetics, 12*, 363. doi:10.1038/nrg2958

Ament, S. A., Szelinger, S., Glusman, G., Ashworth, J., Hou, L., Akula, N., . . . Roach, J. C. (2015). Rare variants in neuronal excitability genes influence risk for bipolar disorder. *Proceedings of the National Academy of Sciences, 112*(11), 3576-3581. doi:10.1073/pnas.1424958112

American Psychiatric Association. (2013a). Bipolar and Related Disorders *Diagnostic and Statistical Manual of Mental Disorders*: American Psychiatric Association.

American Psychiatric Association. (2013b). *Diagnostic and statistical manual of mental disorders: DSM-5* (5 ed.). Washington, DC: American Psychiatric Association.

Angst, J. (2007). Psychiatric diagnoses: the weak component of modern research. *World Psychiatry, 6*(2), 94-95.

Anttila, V., Bulik-Sullivan, B., Finucane, H. K., Walters, R., Bras, J., Duncan, L., . . . Neale, B. (2017). Analysis of shared heritability in common disorders of the brain. *bioRxiv*. doi:10.1101/048991

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., . . . Sherlock, G. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet, 25*(1), 25-29. doi:10.1038/75556

Baldessarini, R. J., Tondo, L., & Hennen, J. (2003). Lithium treatment and suicide risk in major affective disorders: update and new findings. *J Clin Psychiatry, 64 Suppl 5*, 44-52.

Balding, D. J. (2006). A tutorial on statistical methods for population association studies. *Nature Reviews Genetics, 7*, 781. doi:10.1038/nrg1916

Bassett, A. S., Marshall, C. R., Lionel, A. C., Chow, E. W. C., & Scherer, S. W. (2008). Copy number variations and risk for schizophrenia in 22q11.2 deletion syndrome. *Human Molecular Genetics, 17*(24), 4045-4053. doi:10.1093/hmg/ddn307

Benner, C., Spencer, C. C., Havulinna, A. S., Salomaa, V., Ripatti, S., & Pirinen, M. (2016). FINEMAP: efficient variable selection using summary data from genome-wide association studies. *Bioinformatics, 32*(10), 1493-1501. doi:10.1093/bioinformatics/btw018

Biernacka, J. M., Tang, R., Li, J., McDonnell, S. K., Rabe, K. G., Sinnwell, J. P., . . . Fridley, B. L. (2009). Assessment of genotype imputation methods. *BMC Proceedings, 3*(Suppl 7), S5-S5.

Bowden, J., Davey Smith, G., & Burgess, S. (2015). Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *International Journal of Epidemiology, 44*(2), 512-525. doi:10.1093/ije/dyv080

Breen, G., Li, Q., Roth, B. L., O'Donnell, P., Didriksen, M., Dolmetsch, R., . . . Edenberg, H. J. (2016). Translating genome-wide association findings into new therapeutics for psychiatry. *Nature Neuroscience, 19*, 1392. doi:10.1038/nn.4411

Bulik-Sullivan, B. K., Finucane, H. K., Anttila, V., Gusev, A., Day, F. R., Loh, P.-R., . . . Neale, B. M. (2015). An atlas of genetic correlations across human diseases and traits. *Nat Genet, 47*(11), 1236-1241. doi:10.1038/ng.3406

Bulik-Sullivan, B. K., Loh, P.-R., Finucane, H. K., Ripke, S., Yang, J., Schizophrenia Working Group of the Psychiatric Genomics Consortium, . . . Neale, B. M. (2015). LD Score regression distinguishes confounding from polygenicity in

genome-wide association studies. *Nature Genetics, 47*, 291. doi:10.1038/ng.3211

Campos-Parra, A. D., Mitznahuatl, G. C., Pedroza-Torres, A., Romo, R. V., Reyes, F. I. P., Lopez-Urrutia, E., & Perez-Plasencia, C. (2017). Micro-RNAs as Potential Predictors of Response to Breast Cancer Systemic Therapy: Future Clinical Implications. *Int J Mol Sci, 18*(6). doi:10.3390/ijms18061182

Carvalho, B., Bengtsson, H., Speed, T. P., & Irizarry, R. A. (2007). Exploration, normalization, and genotype calls of high-density oligonucleotide SNP array data. *Biostatistics, 8*(2), 485-499. doi:10.1093/biostatistics/kxl042

Chanock, S. J., Manolio, T., Boehnke, M., Boerwinkle, E., Hunter, D. J., Thomas, G., . . . Collins, F. S. (2007). Replicating genotype-phenotype associations. *Nature, 447*(7145), 655-660. doi:10.1038/447655a

Cichon, S., Mühleisen, T. W., Degenhardt, F. A., Mattheisen, M., Miró, X., Strohmaier, J., . . . Nöthen, M. M. (2011). Genome-wide Association Study Identifies Genetic Variation in Neurocan as a Susceptibility Factor for Bipolar Disorder. *The American Journal of Human Genetics, 88*(3), 372-381. doi:10.1016/j.ajhg.2011.01.017

Croft, D., Mundo, A. F., Haw, R., Milacic, M., Weiser, J., Wu, G., . . . D'Eustachio, P. (2014). The Reactome pathway knowledgebase. *Nucleic Acids Res, 42*(Database issue), D472-477. doi:10.1093/nar/gkt1102

Cross-Disorder Group of the Psychiatric Genomics Consortium. (2013). Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *The Lancet, 381*(9875), 1371-1379. doi:10.1016/S0140-6736(12)62129-1

Cruceanu, C., Schmouth, J. F., Torres-Platas, S. G., Lopez, J. P., Ambalavanan, A., Darcq, E., . . . Rouleau, G. A. (2017). Rare susceptibility variants for bipolar disorder suggest a role for G protein-coupled receptors. *Molecular psychiatry*. doi:10.1038/mp.2017.223

Cupertino, R. B., Kappel, D. B., Bandeira, C. E., Schuch, J. B., da Silva, B. S., Muller, D., . . . Mota, N. R. (2016). SNARE complex in developmental psychiatry: neurotransmitter exocytosis and beyond. *J Neural Transm (Vienna), 123*(8), 867-883. doi:10.1007/s00702-016-1514-9

de Leeuw, C. A., Mooij, J. M., Heskes, T., & Posthuma, D. (2015). MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLoS Comput Biol, 11*(4), e1004219. doi:10.1371/journal.pcbi.1004219

de Leeuw, C. A., Neale, B. M., Heskes, T., & Posthuma, D. (2016). The statistical properties of gene-set analysis. *Nat Rev Genet, advance online publication*. doi:10.1038/nrg.2016.29

Delaneau, O., Marchini, J., & Zagury, J.-F. (2011). A linear complexity phasing method for thousands of genomes. *Nature Methods, 9*, 179. doi:10.1038/nmeth.1785

Devlin, B., & Roeder, K. (1999). Genomic control for association studies. *Biometrics, 55*(4), 997-1004.

Dixson, L., Walter, H., Schneider, M., Erk, S., Schäfer, A., Haddad, L., . . . Meyer-Lindenberg, A. (2014). Retraction for Dixson et al., Identification of gene ontologies linked to prefrontal–hippocampal functional coupling in the human brain. *Proceedings of the National Academy of Sciences of the United States of America, 111*(37), 13582-13582. doi:10.1073/pnas.1414905111

Duncan, L. E., Ratanatharathorn, A., Aiello, A. E., Almli, L. M., Amstadter, A. B., Ashley-Koch, A. E., . . . Koenen, K. C. (2017). Largest GWAS of PTSD (N=20 070) yields genetic overlap with schizophrenia and sex differences in heritability. *Molecular psychiatry, 23*, 666. doi:10.1038/mp.2017.77

Edwards, B. J., Haynes, C., Levenstien, M. A., Finch, S. J., & Gordon, D. (2005). Power and sample size calculations in the presence of phenotype errors for case/control genetic association studies. *BMC Genet, 6*, 18. doi:10.1186/1471-2156-6-18

Eichler, E. E., Flint, J., Gibson, G., Kong, A., Leal, S. M., Moore, J. H., & Nadeau, J. H. (2010). Missing heritability and strategies for finding the underlying causes of complex disease. *Nat Rev Genet, 11*(6), 446-450. doi:10.1038/nrg2809

Fabregat, A., Jupe, S., Matthews, L., Sidiropoulos, K., Gillespie, M., Garapati, P., . . . D'Eustachio, P. (2018). The Reactome Pathway Knowledgebase. *Nucleic Acids Res, 46*(D1), D649-d655. doi:10.1093/nar/gkx1132

Fibiger, H. C. (2012). Psychiatry, The Pharmaceutical Industry, and The Road to Better Therapeutics. *Schizophrenia Bulletin, 38*(4), 649-650. doi:10.1093/schbul/sbs073

Fineberg, S. K., Kosik, K. S., & Davidson, B. L. (2009). MicroRNAs potentiate neural development. *Neuron, 64*(3), 303-309. doi:10.1016/j.neuron.2009.10.020

Finucane, H. K., Bulik-Sullivan, B., Gusev, A., Trynka, G., Reshef, Y., Loh, P.-R., . . . Price, A. L. (2015). Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nature Genetics, 47*, 1228. doi:10.1038/ng.3404

Fiorentino, A., O'Brien Niamh, L., Locke Devin, P., McQuillin, A., Jarram, A., Anjorin, A., . . . Gurling Hugh Malcolm, D. (2014). Analysis of ANK3 and CACNA1C variants identified in bipolar disorder whole genome sequence data. *Bipolar Disord, 16*(6), 583-591. doi:10.1111/bdi.12203

Fornaro, M., Orsolini, L., Marini, S., De Berardis, D., Perna, G., Valchera, A., . . . Stubbs, B. (2016). The prevalence and predictors of bipolar and borderline personality disorders comorbidity: Systematic review and meta-analysis. *J Affect Disord, 195*, 105-118. doi:10.1016/j.jad.2016.01.040

Forstner, A. J., Hecker, J., Hofmann, A., Maaser, A., Reinbold, C. S., Mühleisen, T. W., . . . Nöthen, M. M. (2017). Identification of shared risk loci and pathways for bipolar disorder and schizophrenia. *PLoS ONE, 12*(2), e0171595. doi:10.1371/journal.pone.0171595

Forstner, A. J., Hofmann, A., Maaser, A., Sumer, S., Khudayberdiev, S., Mühleisen, T. W., . . . Nöthen, M. M. (2015). Genome-wide analysis implicates microRNAs and their target genes in the development of bipolar disorder. *Transl Psychiatry, 5*, e678. doi:10.1038/tp.2015.159

Fromer, M., Pocklington, A. J., Kavanagh, D. H., Williams, H. J., Dwyer, S., Gormley, P., . . . O'Donovan, M. C. (2014). De novo mutations in schizophrenia implicate synaptic networks. *Nature, 506*(7487), 179-184. doi:10.1038/nature12929

Gandal, M. J., Haney, J. R., Parikshak, N. N., Leppa, V., Ramaswami, G., Hartl, C., . . . Geschwind, D. H. (2018). Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. *Science, 359*(6376), 693. doi:10.1126/science.aad6469

Garnham, J., Munro, A., Slaney, C., MacDougall, M., Passmore, M., Duffy, A., . . . Alda, M. (2007). Prophylactic treatment response in bipolar disorder: Results of a naturalistic observation study. *Journal of Affective Disorders, 104*(1), 185-190. doi:10.1016/j.jad.2007.03.003

GBD 2016 Disease and Injury Incidence and Prevalence Collaborators. (2017). Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet, 390*(10100), 1211-1259. doi:10.1016/s0140-6736(17)32154-2

Geddes, J. R., & Miklowitz, D. J. (2013). Treatment of bipolar disorder. *The Lancet, 381*(9878), 1672-1682. doi:10.1016/S0140-6736(13)60857-0

Georgi, B., Craig, D., Kember, R. L., Liu, W., Lindquist, I., Nasser, S., . . . Bućan, M. (2014). Genomic View of Bipolar Disorder Revealed by Whole Genome Sequencing in a Genetic Isolate. *PLoS Genet, 10*(3), e1004229. doi:10.1371/journal.pgen.1004229

Gershon, E. S., Pearlson, G., Keshavan, M. S., Tamminga, C., Clementz, B., Buckley, P. F., . . . Ivleva, E. I. (2018). Genetic analysis of deep phenotyping projects in common disorders. *Schizophrenia Research, 195*, 51-57. doi:10.1016/j.schres.2017.09.031

Goeman, J. J., & Bühlmann, P. (2007). Analyzing gene expression data in terms of gene sets: methodological issues. *Bioinformatics, 23*(8), 980-987. doi:10.1093/bioinformatics/btm051

Goes, F. S. (2016). Genetics of Bipolar Disorder: Recent Update and Future Directions. *Psychiatric Clinics of North America, 39*(1), 139-155. doi:10.1016/j.psc.2015.10.004

Goes, F. S., Pirooznia, M., Parla, J. S., Kramer, M., Ghiban, E., Mavruk, S., . . . Potash, J. B. (2016). Exome Sequencing of Familial Bipolar Disorder. *JAMA Psychiatry, 73*(6), 590-597. doi:10.1001/jamapsychiatry.2016.0251

Gottesman, I. I., & Gould, T. D. (2003). The Endophenotype Concept in Psychiatry: Etymology and Strategic Intentions. *American Journal of Psychiatry, 160*(4), 636-645. doi:10.1176/appi.ajp.160.4.636

Gratten, J., Wray, N. R., Keller, M. C., & Visscher, P. M. (2014). Large-scale genomics unveils the genetic architecture of psychiatric disorders. *Nature Neuroscience, 17*, 782. doi:10.1038/nn.3708

Green, E. K., Rees, E., Walters, J. T. R., Smith, K. G., Forty, L., Grozeva, D., . . . Kirov, G. (2016). Copy number variation in bipolar disorder. *Mol Psychiatry, 21*(1), 89-93. doi:10.1038/mp.2014.174

Gui, H., Li, M., Sham, P. C., & Cherny, S. S. (2011). Comparisons of seven algorithms for pathway analysis using the WTCCC Crohn's Disease dataset. *BMC Research Notes, 4*(1), 386. doi:10.1186/1756-0500-4-386

Hägg, S., Ganna, A., Van Der Laan, S. W., Esko, T., Pers, T. H., Locke, A. E., . . . Ingelsson, E. (2015). Gene-based meta-analysis of genome-wide association studies implicates new loci involved in obesity. *Human Molecular Genetics, 24*(23), 6849-6860. doi:10.1093/hmg/ddv379

Hammerschlag, A. R., Stringer, S., de Leeuw, C. A., Sniekers, S., Taskesen, E., Watanabe, K., . . . Posthuma, D. (2017). Genome-wide association analysis of insomnia complaints identifies risk genes and genetic overlap with psychiatric and metabolic traits. *Nature Genetics, 49*, 1584. doi:10.1038/ng.3888

Hatzimanolis, A., Bhatnagar, P., Moes, A., Wang, R., Roussos, P., Bitsios, P., . . . Avramopoulos, D. (2015). Common genetic variation and schizophrenia polygenic risk influence neurocognitive performance in young adulthood. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 168*(5), 392-401. doi:10.1002/ajmg.b.32323

Hirschhorn, J. N., & Daly, M. J. (2005). Genome-wide association studies for common diseases and complex traits. *Nature Reviews Genetics, 6*, 95. doi:10.1038/nrg1521

Holmans, P. (2010). Statistical Methods for Pathway Analysis of Genome-Wide Data for Association with Complex Genetic Traits. In C. D. Jay & H. M. Jason (Eds.), *Advances in Genetics* (Vol. Volume 72, pp. 141-179): Academic Press.

Hou, L., Bergen, S. E., Akula, N., Song, J., Hultman, C. M., Landén, M., . . . McMahon, F. J. (2016). Genome-wide association study of 40,000 individuals identifies two novel loci associated with bipolar disorder. *Human Molecular Genetics, 25*, 3383-3394. doi:10.1093/hmg/ddw181

Howard, D. M., Adams, M. J., Shirali, M., Clarke, T.-K., Marioni, R. E., Davies, G., . . . McIntosh, A. M. (2018). Genome-wide association study of depression phenotypes in UK Biobank identifies variants in excitatory synaptic pathways. *Nature Communications, 9*(1), 1470. doi:10.1038/s41467-018-03819-3

Howie, B. N., Donnelly, P., & Marchini, J. (2009). A Flexible and Accurate Genotype Imputation Method for the Next Generation of Genome-Wide Association Studies. *PLOS Genetics, 5*(6), e1000529. doi:10.1371/journal.pgen.1000529

Howie, B. N., Fuchsberger, C., Stephens, M., Marchini, J., & Abecasis, G. R. (2012). Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nature Genetics, 44*, 955. doi:10.1038/ng.2354

Hunsberger, J. G., Chibane, F. L., Elkahloun, A. G., Henderson, R., Singh, R., Lawson, J., . . . Chuang, D. M. (2015). Novel integrative genomic tool for interrogating lithium response in bipolar disorder. *Transl Psychiatry, 5*, e504. doi:10.1038/tp.2014.139

Ibrahim-Verbaas, C. A., Bressler, J., Debette, S., Schuur, M., Smith, A. V., Bis, J. C., . . . Mosley, T. H. (2016). GWAS for executive function and processing speed suggests involvement of the CADM2 gene. *Molecular psychiatry, 21*(2), 189-197. doi:10.1038/mp.2015.37

Jansen, A., Dieleman, G. C., Smit, A. B., Verhage, M., Verhulst, F. C., Polderman, T. J. C., & Posthuma, D. (2017). Gene-set analysis shows association between FMRP targets and autism spectrum disorder. *European Journal Of Human Genetics, 25*, 863. doi:10.1038/ejhg.2017.55

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., & Morishima, K. (2017). KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Research, 45*(D1), D353-D361. doi:10.1093/nar/gkw1092

Kanehisa, M., & Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research, 28*(1), 27-30. doi:10.1093/nar/28.1.27

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., & Tanabe, M. (2016). KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Research, 44*(D1), D457-D462. doi:10.1093/nar/gkv1070

Kang, G., Jiang, B., & Cui, Y. (2013). Gene-based Genomewide Association Analysis: A Comparison Study. *Current Genomics, 14*(4), 250-255. doi:10.2174/13892029113149990001

Kauppi, K., Westlye, L. T., Tesli, M., Bettella, F., Brandt, C. L., Mattingsdal, M., . . . Andreassen, O. A. (2015). Polygenic Risk for Schizophrenia Associated With Working Memory-related Prefrontal Brain Activation in Patients With Schizophrenia and Healthy Controls. *Schizophrenia Bulletin, 41*(3), 736-743. doi:10.1093/schbul/sbu152

Kendler, K. S., Aggen, S. H., Knudsen, G. P., Røysamb, E., Neale, M. C., & Reichborn-Kjennerud, T. (2011). The structure of genetic and environmental risk factors for syndromal and subsyndromal common DSM-IV axis I and all axis II disorders. *American Journal of Psychiatry, 168*(1), 29-39.

Kendler, K. S., & Eaves, L. J. (2005). *Psychiatric genetics, review of psychiatry* (Vol. 24): American Psychiatric Publishing, Inc.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of dsm-iv disorders

in the national comorbidity survey replication. *Archives of General Psychiatry, 62*(6), 593-602. doi:10.1001/archpsyc.62.6.593

Kirov, G. (2015). CNVs in neuropsychiatric disorders. *Human Molecular Genetics, 24*(R1), R45-R49. doi:10.1093/hmg/ddv253

Lander, E. S., & Schork, N. J. (1994). Genetic dissection of complex traits. *Science, 265*(5181), 2037.

Larsson, H., Anckarsater, H., Råstam, M., Chang, Z., & Lichtenstein, P. (2011). Childhood attention-deficit hyperactivity disorder as an extreme of a continuous trait: a quantitative genetic study of 8,500 twin pairs. *Journal of Child Psychology and Psychiatry, 53*(1), 73-80. doi:10.1111/j.1469-7610.2011.02467.x

Lee, J. J., McGue, M., Iacono, W. G., & Chow, C. C. (2018). The accuracy of LD Score regression as an estimator of confounding and genetic correlations in genome-wide association studies. *bioRxiv*. doi:10.1101/234815

Lee, P. H., O'Dushlaine, C., Thomas, B., & Purcell, S. M. (2012). INRICH: interval-based enrichment analysis for genome-wide association studies. *Bioinformatics, 28*(13), 1797-1799. doi:10.1093/bioinformatics/bts191

Lee, Sang H., Wray, Naomi R., Goddard, Michael E., & Visscher, Peter M. (2011). Estimating Missing Heritability for Disease from Genome-wide Association Studies. *American Journal of Human Genetics, 88*(3), 294-305. doi:10.1016/j.ajhg.2011.02.002

Levinson, D. F., Mostafavi, S., Milaneschi, Y., Rivera, M., Ripke, S., Wray, N. R., & Sullivan, P. F. (2014). Genetic Studies of Major Depressive Disorder: Why Are There No Genome-wide Association Study Findings and What Can We Do About It? *Biological Psychiatry, 76*, 510-512. doi:10.1016/j.biopsych.2014.07.029

Lin, D. Y., & Zeng, D. (2010). Meta-Analysis of Genome-Wide Association Studies: No Efficiency Gain in Using Individual Participant Data. *Genetic Epidemiology, 34*(1), 10.1002/gepi.20435. doi:10.1002/gepi.20435

Liu, L., Zhang, D., Liu, H., & Arendt, C. (2013). Robust methods for population stratification in genome wide association studies. *BMC Bioinformatics, 14*(1), 132. doi:10.1186/1471-2105-14-132

Liu, Q., Wu, D.-H., Han, L., Deng, J.-W., Zhou, L., He, R., . . . Mi, Q.-S. (2017). Roles of microRNAs in psoriasis: Immunological functions and potential biomarkers. *Exp Dermatol, 26*(4), 359-367. doi:10.1111/exd.13249

Loh, P.-R., Tucker, G., Bulik-Sullivan, B. K., Vilhjálmsson, B. J., Finucane, H. K., Salem, R. M., . . . Price, A. L. (2015). Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nature Genetics, 47*, 284. doi:10.1038/ng.3190

Mackay, T. F. C. (2001). The Genetic Architecture of Quantitative Traits. *Annual Review of Genetics, 35*(1), 303-339. doi:10.1146/annurev.genet.35.102401.090633

Manchia, M., Adli, M., Akula, N., Ardau, R., Aubry, J.-M., Backlund, L., . . . Alda, M. (2013). Assessment of Response to Lithium Maintenance Treatment in Bipolar Disorder: A Consortium on Lithium Genetics (ConLiGen) Report. *PLoS ONE, 8*(6), e65636. doi:10.1371/journal.pone.0065636

Manchia, M., Cullis, J., Turecki, G., Rouleau, G. A., Uher, R., & Alda, M. (2013). The Impact of Phenotypic and Genetic Heterogeneity on Results of Genome Wide Association Studies of Complex Diseases. *PLoS ONE, 8*(10), e76295. doi:10.1371/journal.pone.0076295

Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J., . . . Visscher, P. M. (2009). Finding the missing heritability of complex diseases. *Nature, 461*(7265), 747-753.

Marangell, L. B., Bauer, M. S., Dennehy, E. B., Wisniewski, S. R., Allen, M. H., Miklowitz, D. J., . . . Thase, M. E. (2006). Prospective predictors of suicide and suicide attempts in 1,556 patients with bipolar disorders followed for up to 2 years. *Bipolar Disord, 8*(5 Pt 2), 566-575. doi:10.1111/j.1399-5618.2006.00369.x

Maston, G. A., Evans, S. K., & Green, M. R. (2006). Transcriptional regulatory elements in the human genome. *Annu Rev Genomics Hum Genet, 7*, 29-59. doi:10.1146/annurev.genom.7.080505.115623

Merikangas, K. R., Jin, R., He, J. P., Kessler, R. C., Lee, S., Sampson, N. A., . . . Zarkov, Z. (2011). Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. *Arch Gen Psychiatry, 68*(3), 241-251. doi:10.1001/archgenpsychiatry.2011.12

Mishra, A., & Macgregor, S. (2015). VEGAS2: Software for More Flexible Gene-Based Testing. *Twin Research and Human Genetics, 18*(01), 86-91. doi:10.1017/thg.2014.79

Mooney, M. A., & Wilmot, B. (2015). Gene set analysis: A step-by-step guide. *Am J Med Genet B Neuropsychiatr Genet, 168*(7), 517-527. doi:10.1002/ajmg.b.32328

Moskvina, V., Craddock, N., Holmans, P., Owen, M. J., & O'Donovan, M. C. (2006). Effects of Differential Genotyping Error Rate on the Type I Error Probability of Case-Control Studies. *Human Heredity, 61*(1), 55-64.

Mühleisen, T. W., Leber, M., Schulze, T. G., Strohmaier, J., Degenhardt, F., Treutlein, J., . . . Cichon, S. (2014). Genome-wide association study reveals two new risk loci for bipolar disorder. *Nature Communications, 5*, 3339. doi:10.1038/ncomms4339

Mühleisen, T. W., Reinbold, C. S., Forstner, A. J., Abramova, L. I., Alda, M., Babadjanova, G., . . . Cichon, S. (2017). Gene set enrichment analysis and expression pattern exploration implicate an involvement of neurodevelopmental processes in bipolar disorder. *Journal of Affective Disorders*. doi:10.1016/j.jad.2017.11.068

Nam, D., Kim, J., Kim, S.-Y., & Kim, S. (2010). GSA-SNP: a general approach for gene set analysis of polymorphisms. *Nucleic Acids Research, 38*(suppl 2), W749-W754. doi:10.1093/nar/gkq428

Neale, B. M., & Sham, P. C. (2004). The Future of Association Studies: Gene-Based Analysis and Replication. *The American Journal of Human Genetics, 75*(3), 353-362. doi:10.1086/423901

Nurnberger, J. I., Jr., Koller, D. L., Jung, J., Edenberg, H. J., Foroud, T., Guella, I., . . . Psychiatric Genomics Consortium Bipolar Group. (2014). Identification of pathways for bipolar disorder: a meta-analysis. *JAMA Psychiatry, 71*(6), 657-664. doi:10.1001/jamapsychiatry.2014.176

O'Donovan, M. C., Craddock, N., Norton, N., Williams, H., Peirce, T., Moskvina, V., . . . Owen, M. J. (2008). Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nat Genet, 40*(9), 1053-1055. doi:10.1038/ng.201

Papassotiropoulos, A., & de Quervain, D. J. F. (2015). Failed drug discovery in psychiatry: time for human genome-guided solutions. *Trends in Cognitive Sciences, 19*(4), 183-187. doi:10.1016/j.tics.2015.02.002

Pardiñas, A. F., Holmans, P., Pocklington, A. J., Escott-Price, V., Ripke, S., Carrera, N., . . . Walters, J. T. R. (2018). Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nature Genetics, 50*(3), 381-389. doi:10.1038/s41588-018-0059-2

Pe'er, I., Yelensky, R., Altshuler, D., & Daly, M. J. (2008). Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genet Epidemiol, 32*(4), 381-385. doi:10.1002/gepi.20303

Peloso, G. M., Rader, D. J., Gabriel, S., Kathiresan, S., Daly, M. J., & Neale, B. M. (2016). Phenotypic extremes in rare variant study designs. *European Journal Of Human Genetics, 24*(6), 924-930. doi:10.1038/ejhg.2015.197

Pescosolido, M. F., Gamsiz, E. D., Nagpal, S., & Morrow, E. M. (2013). Distribution of disease-associated copy number variants across distinct disorders of cognitive development. *J Am Acad Child Adolesc Psychiatry, 52*(4), 414-430.e414. doi:10.1016/j.jaac.2013.01.003

Pickrell, J. K., Berisa, T., Liu, J. Z., Segurel, L., Tung, J. Y., & Hinds, D. A. (2016). Detection and interpretation of shared genetic influences on 42 human traits. *Nat Genet, 48*. doi:10.1038/ng.3570

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics, 38*, 904. doi:10.1038/ng1847

Priebe, L., Degenhardt, F. A., Herms, S., Haenisch, B., Mattheisen, M., Nieratschker, V., . . . Mühleisen, T. W. (2012). Genome-wide survey implicates the influence of copy number variants (CNVs) in the development of early-onset bipolar disorder. *Molecular psychiatry, 17*, 421-432. doi:10.1038/mp.2011.8

Pritchard, J. K. (2001). Are Rare Variants Responsible for Susceptibility to Complex Diseases? *The American Journal of Human Genetics, 69*(1), 124-137. doi:10.1086/321272

Psychiatric GWAS Consortium Steering Committee. (2009). A framework for interpreting genome-wide association studies of psychiatric disorders. *Mol Psychiatry, 14*(1), 10-17. doi:10.1038/mp.2008.126

Purcell, S. M., Moran, J. L., Fromer, M., Ruderfer, D., Solovieff, N., Roussos, P., . . . Sklar, P. (2014). A polygenic burden of rare disruptive mutations in schizophrenia. *Nature, 506*(7487), 185-190. doi:10.1038/nature12975

Purcell, S. M., Wray, N. R., Stone, J. L., Visscher, P. M., O'Donovan, M. C., Sullivan, P. F., . . . Morris, D. W. (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature, 460*(7256), 748-752. doi:10.1038/nature08185

Ramanan, V. K., Shen, L., Moore, J. H., & Saykin, A. J. (2012). Pathway analysis of genomic data: concepts, methods, and prospects for future development. *Trends in Genetics, 28*(7), 323-332. doi:10.1016/j.tig.2012.03.004

Reich, D. E., & Lander, E. S. (2001). On the allelic spectrum of human disease. *Trends in Genetics, 17*(9), 502-510. doi:10.1016/S0168-9525(01)02410-6

Riglin, L., Collishaw, S., Thapar, A. K., Dalsgaard, S., Langley, K., Smith, G. D., . . . Thapar, A. (2016). Association of Genetic Risk Variants With Attention-Deficit/Hyperactivity Disorder Trajectories in the General Population. *JAMA Psychiatry, 73*(12), 1285-1292. doi:10.1001/jamapsychiatry.2016.2817

Ripke, S., O'Dushlaine, C., Chambert, K., Moran, J. L., Kähler, A. K., Akterin, S., . . . Sullivan, P. F. (2013). Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nature Genetics, 45*, 1150. doi:10.1038/ng.2742

Robinson, E. B., St Pourcain, B., Anttila, V., Kosmicki, J. A., Bulik-Sullivan, B., Grove, J., . . . Daly, M. J. (2016). Genetic risk for autism spectrum disorders and

neuropsychiatric variation in the general population. *Nature Genetics, 48*, 552. doi:10.1038/ng.3529

Sanders, S. J., Neale, B. M., Huang, H., Werling, D. M., An, J.-Y., Dong, S., . . . Freimer, N. B. (2017). Whole genome sequencing in psychiatric disorders: the WGSPD consortium. *Nature Neuroscience, 20*(12), 1661-1668. doi:10.1038/s41593-017-0017-9

Schizophrenia Working Group of the Psychiatric Genomics Consortium. (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature, 511*(7510), 421-427. doi:10.1038/nature13595

Sedeño-Cortés, A. E., & Pavlidis, P. (2014). Pitfalls in the application of gene-set analysis to genetics studies. *Trends in Genetics, 30*(12), 513-514. doi:10.1016/j.tig.2014.10.001

Segrè, A. V., DIAGRAM Consortium, MAGIC investigators, Groop, L., Mootha, V. K., Daly, M. J., & Altshuler, D. (2010). Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. *PLOS Genetics, 6*, e1001058. doi:10.1371/journal.pgen.1001058

Sklar, P., Ripke, S., Scott, L. J., Andreassen, O. A., Cichon, S., Craddock, N., . . . Purcell, S. M. (2011). Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nature Genetics, 43*, 977-983. doi:10.1038/ng.943

Sklar, P., Smoller, J. W., Fan, J., Ferreira, M. A. R., Perlis, R. H., Chambert, K., . . . Purcell, S. M. (2008). Whole-genome association study of bipolar disorder. *Molecular psychiatry, 13*, 558-569. doi:10.1038/sj.mp.4002151

Skol, A. D., Scott, L. J., Abecasis, G. R., & Boehnke, M. (2006). Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nature Genetics, 38*, 209. doi:10.1038/ng1706

Smoller, J. W., Andreassen, O. A., Edenberg, H. J., Faraone, S. V., Glatt, S. J., & Kendler, K. S. (2018). Psychiatric genetics and the structure of psychopathology. *Molecular psychiatry*. doi:10.1038/s41380-017-0010-4

Stahl, E., Breen, G., Forstner, A., McQuillin, A., Ripke, S., Bipolar Disorder Working Group of the Psychiatric Genomics Consortium, . . . Sklar, P. (2018). Genome-wide association study identifies 30 loci associated with bipolar disorder. *bioRxiv*. doi:10.1101/173062

Strauss, K. A., Markx, S., Georgi, B., Paul, S. M., Jinks, R. N., Hoshi, T., . . . Puffenberger, E. G. (2014). A population-based study of KCNH7 p.Arg394His and bipolar spectrum disorder. *Human Molecular Genetics, 23*(23), 6395-6406. doi:10.1093/hmg/ddu335

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., . . . Mesirov, J. P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences, 102*(43), 15545-15550. doi:10.1073/pnas.0506580102

Sudmant, P. H., Rausch, T., Gardner, E. J., Handsaker, R. E., Abyzov, A., Huddleston, J., . . . Korbel, J. O. (2015). An integrated map of structural variation in 2,504 human genomes. *Nature, 526*, 75. doi:10.1038/nature15394

Sullivan, P. F., Agrawal, A., Bulik, C. M., Andreassen, O. A., Børglum, A. D., Breen, G., . . . O'Donovan, M. C. (2017). Psychiatric Genomics: An Update and an Agenda. *American Journal of Psychiatry, 175*(1), 15-27. doi:10.1176/appi.ajp.2017.17030283

Sullivan, P. F., Daly, M. J., & O'Donovan, M. (2012). Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nature Reviews Genetics, 13*, 537-551. doi:10.1038/nrg3240

Sullivan, P. F., de Geus, E. J. C., Willemsen, G., James, M. R., Smit, J. H., Zandbelt, T., . . . Penninx, B. W. J. H. (2008). Genome-wide association for major depressive disorder: a possible role for the presynaptic protein piccolo. *Molecular psychiatry, 14*, 359. doi:10.1038/mp.2008.125

Svishcheva, G. R., Axenovich, T. I., Belonogova, N. M., van Duijn, C. M., & Aulchenko, Y. S. (2012). Rapid variance components–based method for whole-genome association analysis. *Nature Genetics, 44*, 1166. doi:10.1038/ng.2410

Teo, Y. Y., Inouye, M., Small, K. S., Gwilliam, R., Deloukas, P., Kwiatkowski, D. P., & Clark, T. G. (2007). A genotype calling algorithm for the Illumina BeadArray platform. *Bioinformatics, 23*(20), 2741-2746. doi:10.1093/bioinformatics/btm443

Tesli, M., Espeseth, T., Bettella, F., Mattingsdal, M., Aas, M., Melle, I., . . . Andreassen, O. A. (2014). Polygenic risk score and the psychosis continuum model. *Acta Psychiatrica Scandinavica, 130*(4), 311-317. doi:10.1111/acps.12307

Thapar, A., Martin, J., Mick, E., Arias Vásquez, A., Langley, K., Scherer, S. W., . . . Holmans, P. (2015). Psychiatric gene discoveries shape evidence on ADHD's biology. *Molecular psychiatry, 21*, 1202. doi:10.1038/mp.2015.163

The Gene Ontology Consortium. (2017). Expansion of the Gene Ontology knowledgebase and resources. *Nucleic Acids Res, 45*(D1), D331-d338. doi:10.1093/nar/gkw1108

The Network Pathway Analysis Subgroup of the Psychiatric Genomics Consortium. (2015). Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. *Nat Neurosci, 18*(2), 199-209. doi:10.1038/nn.3922

Timpson, N. J., Greenwood, C. M. T., Soranzo, N., Lawson, D. J., & Richards, J. B. (2017). Genetic architecture: the shape of the genetic contribution to human traits and disease. *Nature Reviews Genetics, 19*, 110. doi:10.1038/nrg.2017.101

Veyrieras, J.-B., Kudaravalli, S., Kim, S. Y., Dermitzakis, E. T., Gilad, Y., Stephens, M., & Pritchard, J. K. (2008). High-Resolution Mapping of Expression-QTLs Yields Insight into Human Gene Regulation. *PLoS Genet, 4*(10), e1000214. doi:10.1371/journal.pgen.1000214

Visscher, P. M. (2008). Sizing up human height variation. *Nature Genetics, 40*, 489. doi:10.1038/ng0508-489

Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M. I., Brown, M. A., & Yang, J. (2017). 10 Years of GWAS Discovery: Biology, Function, and Translation. *The American Journal of Human Genetics, 101*(1), 5-22. doi:10.1016/j.ajhg.2017.06.005

Wang, K., Li, M., & Bucan, M. (2007). Pathway-Based Approaches for Analysis of Genomewide Association Studies. *American Journal of Human Genetics, 81*(6), 1278-1283.

Wang, L., Jia, P., Wolfinger, R. D., Chen, X., & Zhao, Z. (2011). Gene set analysis of genome-wide association studies: Methodological issues and perspectives. *Genomics, 98*(1), 1-8. doi:10.1016/j.ygeno.2011.04.006

Welter, D., MacArthur, J., Morales, J., Burdett, T., Hall, P., Junkins, H., . . . Parkinson, H. (2014). The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Research, 42*(D1), D1001-D1006. doi:10.1093/nar/gkt1229

Witt, S. H., Streit, F., Jungkunz, M., Frank, J., Awasthi, S., Reinbold, C. S., . . . Rietschel, M. (2017). Genome-wide association study of borderline personality disorder reveals genetic overlap with bipolar disorder, major depression and schizophrenia. *Transl Psychiatry, 7*(6), e1155. doi:10.1038/tp.2017.115

Wray, N. R., Lee, S. H., Mehta, D., Vinkhuyzen, A. A. E., Dudbridge, F., & Middeldorp, C. M. (2014). Research Review: Polygenic methods and their application to psychiatric traits. *Journal of Child Psychology and Psychiatry, 55*(10), 1068-1087. doi:10.1111/jcpp.12295

Wray, N. R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E. M., Abdellaoui, A., . . . Sullivan, P. F. (2018). Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature Genetics, 50*(5), 668-681. doi:10.1038/s41588-018-0090-3

Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., . . . Visscher, P. M. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics, 42*, 565. doi:10.1038/ng.608

Yang, J., Ferreira, T., Morris, A. P., Medland, S. E., Genetic Investigation of ANthropometric Traits Consortium, DIAbetes Genetics Replication Meta-analysis (DIAGRAM) Consortium, . . . Visscher, P. M. (2012). Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nature Genetics, 44*, 369. doi:10.1038/ng.2213

Yang, J., Lee, T., Kim, J., Cho, M.-C., Han, B.-G., Lee, J.-Y., . . . Kim, H. (2013). Ubiquitous Polygenicity of Human Complex Traits: Genome-Wide Analysis of 49 Traits in Koreans. *PLOS Genetics, 9*(3), e1003355. doi:10.1371/journal.pgen.1003355

Zabaneh, D., Krapohl, E., Gaspar, H. A., Curtis, C., Lee, S. H., Patel, H., . . . Breen, G. (2017). A genome-wide association study for extremely high intelligence. *Molecular psychiatry, 23*, 1226. doi:10.1038/mp.2017.121

Zarrei, M., MacDonald, J. R., Merico, D., & Scherer, S. W. (2015). A copy number variation map of the human genome. *Nature Reviews Genetics, 16*, 172. doi:10.1038/nrg3871

Zhang, K., Chang, S., Guo, L., & Wang, J. (2015). I-GSEA4GWAS v2: a web server for functional analysis of SNPs in trait-associated pathways identified from genome-wide association study. *Protein & Cell, 6*(3), 221-224. doi:10.1007/s13238-014-0114-4

Zhao, Z., Xu, J., Chen, J., Kim, S., Reimers, M., Bacanu, S. A., . . . Chen, X. (2015). Transcriptome sequencing and genome-wide association analyses reveal lysosomal function and actin cytoskeleton remodeling in schizophrenia and bipolar disorder. *Mol Psychiatry, 20*(5), 563-572. doi:10.1038/mp.2014.82

Zhu, Z., Anttila, V., Smoller, J. W., & Lee, P. H. (2018). Statistical power and utility of meta-analysis methods for cross-phenotype genome-wide association studies. *PLoS ONE, 13*(3), e0193256. doi:10.1371/journal.pone.0193256

7 Declaration by candidate

I declare herewith that I have independently carried out the doctoral thesis entitled „Application of bioinformatics tools for better interpretation of psychiatric GWAS through integration of biological information“. This thesis consists of original research articles that have been written in cooperation with the enlisted co-authors and have been published in peer-reviewed scientific journals. Only allowed resources were used and all references used were cited accordingly.

For the purpose of the cumulative dissertation, the following studies have been submitted for publication in various journals:

Study I:

Forstner, A. J., Hecker, J., Hofmann, A., Maaser, A., Reinbold, C. S., Mühleisen, T. W., . . . Nöthen, M. M. (2017). Identification of shared risk loci and pathways for bipolar disorder and schizophrenia. PLoS ONE, 12(2), e0171595. doi:10.1371/journal.pone.0171595

Study II:

Witt, S. H., Streit, F., Jungkunz, M., Frank, J., Awasthi, S., Reinbold, C. S., . . . Rietschel, M. (2017). Genome-wide association study of borderline personality disorder reveals genetic overlap with bipolar disorder, major depression and schizophrenia. Transl Psychiatry, 7(6), e1155. doi:10.1038/tp.2017.115

Study III:

Mühleisen, T. W., Reinbold, C. S., Forstner, A. J., Abramova, L. I., Alda, M., Babadjanova, G., . . . Cichon, S. (2017). Gene set enrichment analysis and expression pattern exploration implicate an involvement of neurodevelopmental processes in bipolar disorder. Journal of Affective Disorders. doi:10.1016/j.jad.2017.11.068

Study IV:

Reinbold, C. S., Forstner, A. J., Hecker, J., Fullerton, J. M., Hoffmann, P., Hou, L., . . . Cichon, S. (2018). Analysis of the Influence of microRNAs in Lithium Response in Bipolar Disorder. Accepted for publication in Frontiers in Psychiatry. doi:10.3389/fpsyt.2018.00207

With my signature, I testify that all statements are true and complete.

Basel, _____ Signature: _____