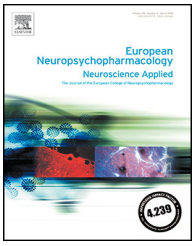
European Neuropsychopharmacology (2019) 29, 156–170

www.elsevier.com/locate/euroneuro



# Efficient region-based test strategy uncovers genetic risk factors for functional outcome in bipolar disorder

Monika Budde[a], Stefanie Friedrichs[b], Ney Alliey-Rodriguez[c], Seth Ament[d], Judith A. Badner[e], Wade H. Berrettini[f], Cinnamon S. Bloss[g], William Byerley[h], Sven Cichon[i,j,k], Ashley L. Comes[a,l], William Coryell[m], David W. Craig[n], Franziska Degenhardt[o,p], Howard J. Edenberg[q,r], Tatiana Foroud[r], Andreas J. Forstner[o,p,i,s], Josef Frank[t], Elliot S. Gershon[c], Fernando S. Goes[u], Tiffany A. Greenwood[v], Yiran Guo[w,x], Maria Hipolito[y], Leroy Hood[d], Brendan J. Keating[z,aa], Daniel L. Koller[r], William B. Lawson[ab], Chunyu Liu[ac], Pamela B. Mahon[u], Melvin G. McInnis[ad], Francis J. McMahon[ae], Sandra M. Meier[t,af], Thomas W. Mühleisen[k,i], Sarah S. Murray[ag,ah], Caroline M. Nievergelt[v], John I. Nurnberger Jr.[ai], Evaristus A. Nwulia[y], James B. Potash[aj], Danjuma Quarless[ak,g], John Rice[al], Jared C. Roach[d], William A. Scheftner[am], Nicholas J. Schork[ak,n,g], Tatyana Shekhtman[v], Paul D. Shilling[v], Erin N. Smith[ag,an], Fabian Streit[t], Jana Strohmaier[t], Szabolcs Szelinger[n], Jens Treutlein[t], Stephanie H. Witt[t], Peter P. Zandi[ao], Peng Zhang[ap], Sebastian Zöllner[ap,ad], Heike Bickeböller[b], Peter G. Falkai[aq], John R. Kelsoe[v], Markus M. Nöthen[o,p], Marcella Rietschel[t], Thomas G. Schulze[a,t,u,ae,2,*], Dörthe Malzahn[b,1,2]

* Corresponding author at: Institute of Psychiatric Phenomics and Genomics, University Hospital, LMU Munich, Nussbaumstr. 7, Munich D-80336, Germany.

*E-mail addresses:* Thomas.Schulze@med.uni-muenchen.de (T.G. Schulze), mz@mzbiostatistics.de (D. Malzahn).

[1]Present address: mzBiostatistics, Statistical Consultancy, Henri-Dunant-Str.10, D-37075 Göttingen, Germany.

[2]Both authors share the last author position.

https://doi.org/10.1016/j.euroneuro.2018.10.005

[a] *Institute of Psychiatric Phenomics and Genomics, University Hospital, LMU Munich, Nussbaumstr. 7, Munich 80336, Germany*
[b] *Department of Genetic Epidemiology, University Medical Center Göttingen, Georg-August-University, Göttingen 37099, Germany*
[c] *Department of Psychiatry and Behavioral Neuroscience, University of Chicago, Chicago, IL 60637, United States*
[d] *Institute for Systems Biology, Seattle, WA 98109, United States*
[e] *Department of Psychiatry, Rush University Medical Center, Chicago, IL 60612, United States*
[f] *Department of Psychiatry, University of Pennsylvania, Philadelphia, PA 19104, United States*
[g] *University of California San Diego, La Jolla, CA 92093, United States*
[h] *Department of Psychiatry, University of California at San Francisco, San Francisco, CA 94103, United States*
[i] *Human Genomics Research Group, Department of Biomedicine, University of Basel, Basel 4031, Switzerland*
[j] *Institute of Medical Genetics and Pathology, University Hospital Basel, Basel 4031, Switzerland*
[k] *Institute of Neuroscience and Medicine (INM-1), Research Centre Jülich, Jülich 52425, Germany*
[l] *International Max Planck Research School for Translational Psychiatry, Max Planck Institute of Psychiatry, Munich 80804, Germany*
[m] *University of Iowa Hospitals and Clinics, Iowa City, IA 52242, United States*
[n] *The Translational Genomics Research Institute, Phoenix, AZ 85004, United States*
[o] *Institute of Human Genetics, School of Medicine & University Hospital Bonn, University of Bonn, Bonn 53127, Germany*
[p] *Department of Genomics, Life & Brain Center, University of Bonn, Bonn 53127, Germany*
[q] *Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, IN 46202, United States*
[r] *Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN 46202, United States*
[s] *Department of Psychiatry (UPK), University of Basel, Basel 4012, Switzerland*
[t] *Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, Mannheim 68159, Germany*
[u] *Department of Psychiatry and Behavioral Sciences, Johns Hopkins School of Medicine, Baltimore, MD 21287, United States*
[v] *Department of Psychiatry, University of California San Diego, San Diego, CA 92093, United States*
[w] *Center for Applied Genomics, Children's Hospital of Philadelphia, Abramson Research Center, Philadelphia, PA 19104, United States*
[x] *Beijing Genomics Institute at Shenzhen, Shenzhen 518083, China*
[y] *Department of Psychiatry and Behavioral Sciences, Howard University Hospital, Washington, DC 20060, United States*
[z] *Cardiovascular Institute, University of Pennsylvania School of Medicine, Philadelphia, PA 19104-5159, United States*
[aa] *Institute for Translational Medicine and Therapeutics, School of Medicine, University of Pennsylvania, Philadelphia, PA 19104-5158, United States*
[ab] *Dell Medical School, University of Texas at Austin, Austin, TX 78723, United States*
[ac] *SUNY Upstate Medical University, Syracuse, NY 13210, United States*
[ad] *Department of Psychiatry, University of Michigan, Ann Arbor, MI 48105, United States*
[ae] *U.S. Department of Health & Human Services, Intramural Research Program, National Institute of Mental Health, National Institutes of Health, Bethesda, MD 20894, United States*
[af] *National Centre for Register-Based Research, Aarhus University, Aarhus V 8210, Denmark*
[ag] *Scripps Genomic Medicine & The Scripps Translational Sciences Institute (STSI), La Jolla, CA 92037, United States*
[ah] *Department of Pathology, University of California San Diego, La Jolla, CA 92093, United States*
[ai] *Department of Psychiatry, Indiana University School of Medicine, Indianapolis, IN 46202, United States*
[aj] *Department of Psychiatry, Carver College of Medicine, University of Iowa School of Medicine, Iowa City, IA 52242, United States*
[ak] *J. Craig Venter Institute, La Jolla, CA 92037, United States*
[al] *Department of Psychiatry, Washington University School of Medicine in St. Louis, St. Louis, MO 63110, United States*
[am] *Rush University Medical Center, Chicago, IL 60612, United States*
[an] *Department of Pediatrics and Rady's Children's Hospital, School of Medicine, University of California San Diego, La Jolla, CA 92037, United States*
[ao] *Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, United States*

[ap] *Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI 48109, United States*
[aq] *Department of Psychiatry and Psychotherapy, University Hospital, LMU Munich, Munich 80336, Germany*

Received 12 July 2018; received in revised form 16 October 2018; accepted 23 October 2018

**KEYWORDS**
Hypothesis-driven GWAS;
Psychiatric disorder;
Global Assessment of Functioning;
Kernel score test;
Linkage disequilibrium;
Functional annotation

**Abstract**
Genome-wide association studies of case-control status have advanced the understanding of the genetic basis of psychiatric disorders. Further progress may be gained by increasing sample size but also by new analysis strategies that advance the exploitation of existing data, especially for clinically important *quantitative* phenotypes. The *f*unctionally-*i*nformed *e*fficient *r*egion-based test strategy (FIERS) introduced herein uses *prior* knowledge on biological function and dependence of genotypes within a powerful statistical framework with improved sensitivity and specificity for detecting consistent genetic effects across studies. As proof of concept, FIERS was used for the first genome-wide single nucleotide polymorphism (SNP)-based investigation on bipolar disorder (BD) that focuses on an important aspect of disease course, the functional outcome. FIERS identified a significantly associated locus on chromosome 15 (hg38: chr15:48965004 – 49464789 bp) with consistent effect strength between two independent studies (*GAIN*/*TGen*: European Americans, *BOMA*: Germans; $n = 1592$ BD patients in total). Protective and risk haplotypes were found on the most strongly associated SNPs. They contain a *CTCF* binding site (rs586758); *CTCF* sites are known to regulate sets of genes within a chromatin domain. The rs586758 – rs2086256 – rs1904317 haplotype is located in the promoter flanking region of the *COPS2* gene, close to microRNA4716, and the *EID1, SHC4, DTWD1* genes as plausible biological candidates. While implication with BD is novel, *COPS2, EID1*, and *SHC4* are known to be relevant for neuronal differentiation and function and *DTWD1* for psychopharmacological side effects. The test strategy FIERS that enabled this discovery is equally applicable for tag SNPs and sequence data.


## 1. Introduction

For years, collaborative consortia have vastly increased sample sizes for genome-wide association studies (GWAS). However, worldwide sample size is finite, and data on clinically important *quantitative* phenotypes is currently limited, largely due to high costs of deep phenotyping and lacking harmonization of assessment scales and conditions across studies. Nevertheless, quantitative phenotypes are especially valuable for understanding underlying biological mechanisms and between-patient heterogeneity. Hence, complementary to increasing sample size, new approaches and strategies that advance the exploitation of existing genome-wide data are highly desirable.

To gain power and identify underlying mechanisms, recently single-marker tests have been replaced by joint statistics on biological units (Subramanian et al., 2005; Wang et al., 2007). Joint statistics greatly reduce the multiple-testing burden and may increase power by aggregating association signals from multiple functionally-related loci. Many pioneering approaches have aggregated single-SNP GWAS *p*-values into enrichment statistics for genes or pathways (Wang et al., 2010). However, unbiased scoring often necessitates time-consuming permutation procedures, since genes and pathways differ in numbers of SNPs, gene length, gene number and linkage disequilibrium (LD)-patterns. Alternatively, SNPs may be aggregated into polygenic risk scores that serve for association testing or trait prediction (Dudbridge, 2013). Risk scores reduce the model space: they collapse multiple SNPs into a single score with *a priori* assumptions on the selection and weighting of contributing SNPs (Dudbridge, 2013). A third set of methods provide actual joint tests of SNPs at the individual-data level. Among them, the kernel score test SKAT (Schaid, 2010) is very powerful for a broad range of genetic architectures, computationally convenient, and yields exact *p*-values.

Whereas LD is a nuisance for most statistics, SKAT can exploit LD to increase power compared to single-marker tests (Schifano et al., 2012) to the extent that testing LD-blocks with SKAT is especially powerful (Malzahn et al., 2016). Since SKAT is a joint test, power increases with cumulative association strength and the ratio between sample size and number of jointly tested SNPs. Therefore, tag SNPs may provide higher power than a denser common SNP panel of the same region (Malzahn et al., 2014). Whereas association strengths and available sample sizes depend on studied phenotypes, sizes of tested SNP sets are the

analysts' choice. Of all 284 human pathways listed in the KEGG (Kanehisa and Goto, 2000) database at the time of download, only 9.5% contained fewer than 500 SNPs of a typical GWAS marker panel, but 47% of the pathways contained more than 2000 SNPs, and the longest pathway contained around 14,500 SNPs. For clinically important phenotypes however, primary studies or even worldwide samples with comparable phenotyping may encompass only a few thousand subjects. In these instances, power likely differs profoundly between short and long pathways, whereas smaller biological units provide stable power. Note also that pathways may share genes and genes may share SNPs, thus yielding partially overlapping test sets. Herein, we leverage the observed enrichment of small *p*-values across GWAS among SNPs linked with specific functional elements (Schork et al., 2013). We *a priori* identify and test only LD-blocks containing specific functional SNPs, considering these regions putatively relevant in a hypothesis-driven GWAS. A variety of classes of putative functionality of SNPs may be used for selecting genomic regions of interest *a priori*. Herein, we chose to use non-synonymous coding SNPs (nsSNPs) and no other functional information, as currently nsSNPs can be most reliably predicted (Li and Wei, 2015; Saunders and Baker, 2002) and many genes implicated with BD susceptibility (Hou et al., 2016), the disorder of interest herein, are protein coding. Hence the presented analysis focused on LD-blocks that overlap with protein-coding sections of the genome, with the extension that exploiting LD putatively may include additional information from SNPs with other functionalities as well. The testing of LD-blocks fully capitalizes on SKAT's advantages. In addition, we improved sensitivity and specificity to detect consistent genetic effects across studies by employing an extension of SKAT (Malzahn et al., 2014) for cross-study analysis of individual-level data (mega-analysis).

As proof of concept, we demonstrate the success of this **f**unctionally-**i**nformed **e**fficient **r**egion-based test **s**trategy (FIERS) to uncover genetic risk factors for functional outcome in bipolar disorder (BD) in two independent studies, *Genetic Information Association Network* (*GAIN)* (Smith et al., 2009)/*Translational Genomics Research Institute (TGen)* study (Smith et al., 2011), United States, and the *Bonn-Mannheim (BOMA)* study (Cichon et al., 2011; Fangerau et al., 2005), Germany, comprising 1592 patients. BD is among the 20 leading causes of disability worldwide (Vos et al., 2012) and genetic factors contribute to BD susceptibility (Bienvenu et al., 2011; Charney et al., 2017). However, functional outcome of BD is highly variable. While some patients with a mild course of BD experience hardly any restrictions in work or personal relationships between illness episodes, an estimated 30–60% suffer from substantial impairment up to the point of disability (Sanchez-Moreno et al., 2009). Apart from severe socio-economic consequences, impaired functional outcome also implies a reduced perceived quality of life of patients (Sum et al., 2015). Several socio-demographic, clinical and cognitive factors associate with impaired functional outcome in BD (for an overview see Gade et al., 2015; Reinares et al., 2013; Solé et al., 2018). The knowledge of these factors and of their interplay is critical for optimizing individualized treatment (Reinares et al., 2013). Along the same line, it is of utmost importance to gain deeper insights into the biological underpinnings of between-patient heterogeneity of functional outcome of BD.

Heritability and familial clustering of reduced global (Savage et al., 2012; Vassos et al., 2008), social (Schulze et al., 2006), and occupational (Potash et al., 2007) functioning in families of patients with schizophrenia (Savage et al., 2012; Vassos et al., 2008) or BD (Potash et al., 2007; Schulze et al., 2006) suggest genetic influences. Furthermore, Global Assessment of Functioning (GAF; DSM-IV Axis V) was lower in healthy carriers of neuropsychiatric copy-number-variants compared to non-carriers (Stefansson et al., 2013). We present here the first genomic study of functional outcome of BD. While BD has an episodic character, most patients experience longer times outside of severe acute manic or depressive episodes than within. Consequently, FIERS was employed to analyze GAF assessed during outpatient treatment, as important cross-diagnostic indicator of overall course and severity of psychiatric disorder.

## 2. Experimental procedures

### 2.1. Study participants

Data were provided by the *GAIN*/*TGen* study, United States (Smith et al., 2009, 2011), and the *BOMA* study, Germany (Cichon et al., 2011; Fangerau et al., 2005). All participants gave written informed consent prior to study participation. Study protocols were approved by the respective institutional review boards and in accordance with the 1964 Declaration of Helsinki. For the *BOMA* sample, summary statistics can be accessed via the Psychiatric Genomic Consortium (http://www.med.unc.edu/pgc/) and individual data by contacting the Institute of Psychiatric Phenomics and Genomics, University Hospital, LMU Munich, Germany (Thomas G. Schulze). *GAIN*/*TGen* data can be obtained by contacting the Bipolar Genome Study (John R. Kelsoe). *GAIN* genotypes are also available at the *database of Genotypes and Phenotypes* (phs000017.v3.p1).

From *GAIN*/*TGen*, we analyzed 1081 adults of European American ancestry diagnosed with BD according to DSM-IV criteria who had GAF scores. *GAIN*/*TGen* provided imputed genome-wide genotypes (see Smith et al., 2009, 2011 for details). Patient age ranged from 17 to 77 years (mean $\pm$ sd: 43 $\pm$ 12 years), duration of illness from 0.5 to 64 years (mean $\pm$ sd: 24 $\pm$ 13 years), and 34.9% ($n = 377$) of participants were men. Diagnoses were obtained based on the Diagnostic Interview for Genetic Studies (DIGS) (Nurnberger et al., 1994) and review of available family history and medical records through a best estimate procedure.

*BOMA* participants had minimal illness duration of 6 months and were recruited for the purpose of genetic studies (Fangerau et al., 2005) from consecutive hospital admissions at the Central Institute of Mental Health, Mannheim and the Department of Psychiatry, University of Bonn, Germany. Diagnoses were established by the German version of the Structured Clinical interview for DSM-IV-TR Axis I Disorders (Diagnostic and Statistical Manual of Mental Disorders, 4th ed, text revision, SCID-I) (First et al., 2002; Wittchen and Fydrich, 1997). We analyzed 511 adult inpatients with a lifetime-diagnosis of BD according to DSM-IV criteria

and available pre-admission GAF scores and genome-wide genotypes (Illumina: HumanHap550v3, Human610, Human660w) (17). *BOMA* patient age was comparable to *GAIN/TGen* and ranged from 18 to 78 years (mean ± sd: 46 ± 13 years), duration of illness was on average shorter (ranging from 0.5 to 61 years, mean ± sd: 17 ± 12 years), and the proportion of men was higher (45.0%, $n = 230$).

## 2.2. Phenotype

Functional outcome was assessed by the GAF score (DSM-IV Axis V, American Psychiatric Association, 2002); details on scale development are described elsewhere (Endicott et al., 1976; Luborsky, 1962). GAF rates the overall psychological, social and occupational functioning of a subject on a continuum ranging from 1 to 100 (Luborsky, 1962). Poorer functioning is indicated by lower GAF scores.

We analyzed GAF in BD outpatients to target course of disorder outside acute illness episodes. GAF assessments were performed by board-certified psychiatrists or psychologists or psychiatry/psychology trainees at advanced stages in their postgraduate education. In *GAIN/TGen*, GAF was an average rating over the past (last) month assessed by direct interview of outpatients. Observed scores ranged from 5 to 100 with a median score of 61. In *BOMA*, the GAF score represents a pre-admission state right before the "current" episode for which the patient received clinical treatment at the time of study interview. Observed GAF scores in *BOMA* were higher compared to *GAIN/TGen* and ranged from 25 to 100 with a median score of 80.

## 2.3. Statistical methods - FIERS

FIERS applied a hypothesis-guided filter on the genome, combining two types of *prior* information: LD structure (from independent reference data) and functional knowledge (from bioinformatic annotation tools). The goal was to *a priori* identify LD-blocks that contain specific functional elements. In a second step, only these LD-blocks were tested for genotype-phenotype association by employing a generalization of SKAT (Malzahn et al., 2014) for cross-study analysis of individual-level data (mega-analysis, details below). This comprised the genotypic information of an LD-block into a single association test, yielding a single *p*-value per LD-block, respectively. The employed generalization of SKAT was especially powerful since it methodically optimally exploited all available information; specifically, genomic correlations and consistency (or lack thereof) of putative genetic effects across samples. Finally, to gain additional insight, the detected significant LD-block was examined in detail by single-SNP and haplotype association analyses.

### 2.3.1. FIERS - step I: hypothesis-guided LD-based selection of genomic regions

*Prior* information on nsSNPs and LD were obtained for independent population-based reference data from the International Haplotype Map Project (HapMap phase II CEU sample - northern and western European ancestry; 2591820 SNPs, Sabeti et al., 2003) and matched to *GAIN/TGen* and *BOMA* using hg38 SNP-positions obtained with biomaRt (bioconductor). A listing of nsSNPs was obtained based on SNP rs-identifier numbers from SNPnexus (Dayem Ullah et al., 2013) as predicted by at least one of the widely accepted SIFT (Kumar et al., 2009) or PolyPhen (Adzhubei et al., 2010) bioinformatics tools (Friedrichs et al., 2016), and irrespective of predicted nsSNP impact as this may vary across transcript isoforms. The hg38 start and end positions of LD-blocks that contain these functionally annotated SNPs were determined using the default algorithm of Haploview 4.2 (Barrett et al., 2005) such that within assigned LD-blocks at least 70% of all SNP pairs had D' estimates with lower 95% confidence limits above 0.5. The rationale was to detect reasonably strongly correlated SNP sets for subsequent combined evaluation (Malzahn et al., 2016).

*GAIN/TGen* and *BOMA* samples were genetically homogeneous and indistinguishable in the four most important principal components (multidimensional scaling analysis, PLINK, data not shown; see Table 1 and Fig. 2 [symbols in bottom panel] for high cross-study similarity of estimated variant frequencies and SNP correlations). Combining external LD information with nsSNP data identified 2957 LD-based blocks for association testing (containing 51,382 SNPs in total) from 410,943 common SNPs available in *GAIN/TGen* and *BOMA* after quality control. SNPs were directly typed (*BOMA*: Hardy-Weinberg equilibrium $p$-value $\geq 10^{-5}$, call rate $\geq 95\%$) or came from a larger imputed panel (*GAIN/TGen*, see Smith et al., 2009, 2011 for details on genotyping, quality control and imputation). By construction, GWAS marker panels are LD-pruned. Nevertheless, substantial amounts of LD remain and test strategy FIERS exploits this. The two largest tested LD-based blocks contained 430 and 186 SNPs; all other blocks contained fewer than 79 SNPs. With regards to nsSNP content, 72% of the tested LD-based blocks contained a single nsSNP, 28% contained at least two, with a maximum of 26 nsSNPs in a block.

### 2.3.2. FIERS - step II: region-based cross-study analysis of individual-level data

PLINK and R (version 3.2.2) were used for statistical analyses. All p-values reported are two-sided. Genetic association screening was performed for the full sample (*quantitative GAF*). Additionally, subjects who had GAF values in the lowest versus highest sample quartile were compared (*GAF extremes*). All analyses were adjusted for fixed effects of sex and duration of illness (Gade et al., 2015). Putative functional LD-based blocks were tested with SKAT in each study (*GAIN/TGen*, *BOMA*) and in cross-study analyses (mega-analysis of individual-level data; *quantitative GAF*: linear model, adjusting for between-study differences of GAF values by a random effect (Malzahn et al., 2014); *GAF extremes*: logistic model, adjusting for between-study differences of GAF values by the additional covariate *study*). Mega-analysis of individual-level data within SKAT assumed common SNP effects across studies and a linear kernel on minor allele dosages (additive model) with *beta*-density SNP-weights *Beta*(MAF,0.5,0.5) that depend on the minor allele frequency (MAF) of SNPs. This choice of kernel and SNP-weights ensured robust power for detecting genetic main effects (Malzahn et al., 2016). Mega-analysis increased the power (*sensitivity*) for detecting reproducible genetic effects as it combined concordant effects across stud-

**Table 1** GAF in BD outpatients associates with an LD-block on chromosome 15 (hg38: chr15:48965004 - 49464789 bp).

| | | Position | Frequency | | Effect on *quantitative GAF* | | | | Effect on *GAF extremes* | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| REGION[a] | | | | | Single studies | | | Mega analysis | Single studies | | | Mega analysis |
| Chromosome 15 | | 48965004 - | 44 SNPs | | GER | $P = 4.9 \times 10^{-4}$ | | $P = 1.3 \times 10^{-5}$ | GER | $P = 5.4 \times 10^{-4}$ | | $P = 5.6 \times 10^{-6}$ |
| | | 49464789 | | | US | $P = 5.9 \times 10^{-3}$ | | | US | $P = 1.6 \times 10^{-3}$ | | |
| Top-ranked SNPs within this region[b] and negatively correlated nsSNP *rs11854184* | | | | | | | | | | | | |
| SNP | MA | Position | Frequency | | Effect: beta[b] | | 95%CI[b] | Meta-analysis | Effect: OR[b] | | 95% CI[b] | Meta-analysis |
| rs4474633 | A | 48968404 | GER | 0.316 | GER | −3.72 | [−5.63, −1.80] | $P = 4.4 \times 10^{-5}$ | GER | 2.21 | [1.47, 3.42] | $P = 1.3 \times 10^{-5}$ |
| | | | US | 0.329 | US | −1.59 | [−2.93, −0.25] | | US | 1.48 | [1.14, 1.93] | |
| *rs11854184* | A | 49000997 | GER | 0.199 | GER | 2.84 | [0.54, 5.13] | $P = 0.013$ | GER | 0.56 | [0.33, 0.91] | $P = 0.013$ |
| | | | US | 0.188 | US | 1.34 | [−0.32, 3.00] | | US | 0.75 | [0.54, 1.03] | |
| rs2413930 | T | 49083018 | GER | 0.230 | GER | −4.12 | [−6.28, −1.97] | $P = 1.5 \times 10^{-5}$ | GER | 2.26 | [1.44, 3.63] | $P = 5.8 \times 10^{-6}$ |
| | | | US | 0.285 | US | −1.99 | [−3.38, −0.61] | | US | 1.63 | [1.23, 2.16] | |
| rs586758 | A | 49216375 | GER | 0.287 | GER | −3.81 | [−5.80, −1.82] | $P = 2.5 \times 10^{-5}$ | GER | 2.31 | [1.51, 3.62] | $P = 2.0 \times 10^{-6}$ |
| | | | US | 0.297 | US | −1.84 | [−3.21, −0.47] | | US | 1.63 | [1.23, 2.17] | |
| rs2086256 | T | 49265829 | GER | 0.346 | GER | −3.23 | [−5.13, −1.33] | $P = 1.1 \times 10^{-5}$ | GER | 2.06 | [1.37, 3.16] | $P = 4.6 \times 10^{-6}$ |
| | | | US | 0.348 | US | −2.27 | [−3.60, −0.94] | | US | 1.62 | [1.24, 2.13] | |
| rs1904317 | T | 49270069 | GER | 0.289 | GER | −3.79 | [−5.77, −1.81] | $P = 2.5 \times 10^{-5}$ | GER | 2.28 | [1.49, 3.56] | $P = 2.2 \times 10^{-6}$ |
| | | | US | 0.297 | US | −1.84 | [−3.22, −0.47] | | US | 1.63 | [1.23, 2.17] | |
| Six-locus haplotype[c] rs4474633 - rs11854184 - rs2413930 - rs586758 - rs2086256 - rs1904317 | | | | | | | | | | | | |
| Haplotype group | | Position | Frequency | | Effect: beta[c] | | 95%CI[c] | Single studies | Effect: OR[c] | | 95% CI[c] | Single studies |
| ***GCC | | 48968404 - | GER | 0.654 | GER | 3.23 | [1.33, 5.13] | $P = 9.1 \times 10^{-4}$ | GER | 0.49 | [0.32, 0.73] | $P = 6.6 \times 10^{-4}$ |
| | | 49270069 | US | 0.652 | US | 2.27 | [0.94, 3.60] | $P = 8.3 \times 10^{-4}$ | US | 0.62 | [0.47, 0.81] | $P = 4.4 \times 10^{-4}$ |
| ***ATT | | 48968404 - | GER | 0.287 | GER | −3.81 | [−5.80, −1.82] | $P = 2.0 \times 10^{-4}$ | GER | 2.31 | [1.51, 3.62] | $P = 1.8 \times 10^{-4}$ |
| | | 49270069 | US | 0.297 | US | −1.84 | [−3.22, −0.47] | $P = 8.7 \times 10^{-3}$ | US | 1.63 | [1.23, 2.17] | $P = 6.7 \times 10^{-4}$ |
| ***GTC | | 48968404 - | GER | 0.058 | GER | 0.86 | [−2.99, 4.72] | $P = 0.661$ | GER | 0.76 | [0.29, 1.93] | $P = 0.567$ |
| | | 49270069 | US | 0.051 | US | −2.36 | [−5.14, 0.42] | $P = 0.097$ | US | 1.16 | [0.68, 1.97] | $P = 0.587$ |

CI, confidence interval; GER, German (*BOMA* study); MA, minor allele; nsSNP, non-synonymous coding SNP; OR, odds ratio; *P*, *p*-value; US, US American Europeans (*GAIN*/*TGen* study).

[a] SKAT cross-study mega-analysis of individual-level *BOMA* and *GAIN*/*TGen* data on chromosome 15, hg38: chr15:48965004 - 49464789 bp (*quantitative GAF*: linear model, *GAF extremes*: logistic model; adjusted for sex, illness duration, and study; see Methods for details). SKAT-derived *p*-values (*P*) summarize the joint influence of the available 44 SNPs in this LD-block on *quantitative GAF* and *GAF extremes*.

[b] Displayed are single-SNP analyses of additive minor allele effects on functional outcome for the five most strongly associated SNPs and the enclosed nsSNP *rs11854184* in the significant LD-block. Analyses within studies are adjusted for sex and illness duration. Studies were meta-analytically combined by Fisher's *p*-value pooling. A protective minor allele effect is indicated by a positive regression coefficient beta > 0 (*quantitative GAF*, linear model) and odds ratio OR < 1 (contrast between *GAF extremes*, logistic model); beta < 0, OR > 1 for risk minor alleles. Minor allele dosages of all five strongly associated SNPs are pairwise strongly positively correlated, and negatively correlated to the minor allele dosage of putative nsSNP *rs11854184*.

[c] Individual best-estimate haplotypes were nonambiguous on the last three positions and grouped accordingly. The three most frequent haplotype groups had the identifying nucleobase combinations GCC, ATT, GTC at rs586758, rs2086256, rs1904317; nucleotide base combinations at rs4474633, rs11854184, rs2413930 varied (indicated by ***). For haplotype groups, an additive haplotype effect on functional outcome was tested in each study, with adjustment for sex and illness duration. Effects (beta, OR) are specified per haplotype copy. A protective haplotype effect is indicated by a positive regression coefficient beta > 0 (*quantitative GAF*, linear model) and odds ratio OR < 1 (*GAF extremes*, logistic model); beta < 0, OR > 1 for risk haplotypes.

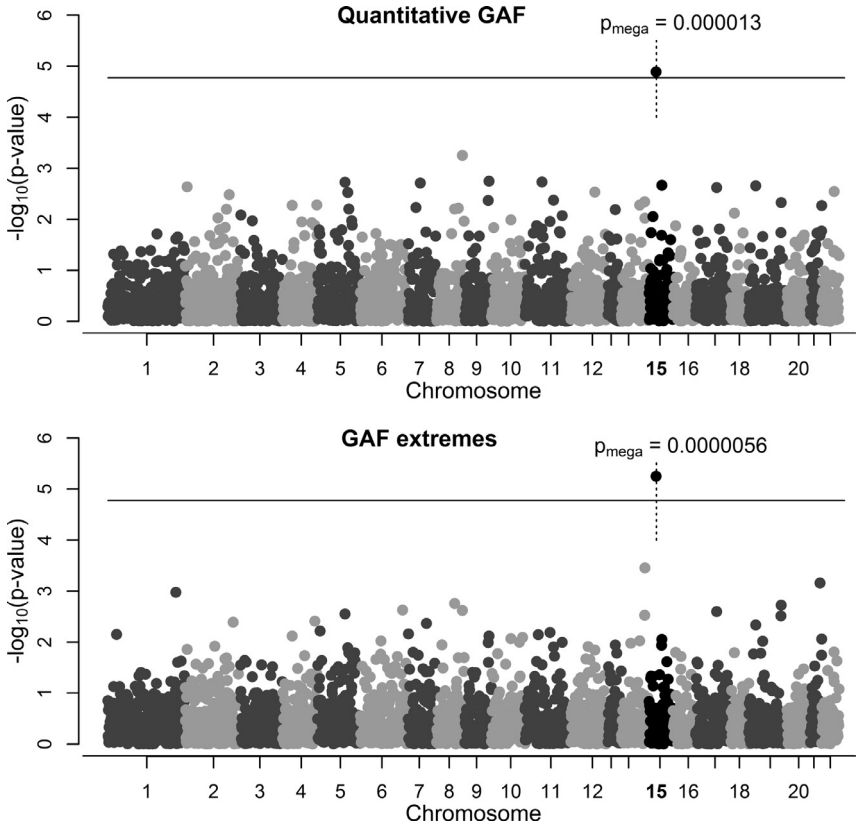



**Fig. 1** Mega-analysis of GAF in German and European American BD outpatients, adjusted for sex and duration of illness. Manhattan-like plots display SKAT-derived *p*-values on the 2957 tested putatively functionally relevant LD-based genomic regions (cross-study mega-analysis of individual-level *BOMA* and *GAIN/TGen* data). Significance (horizontal line, Bonferroni $\alpha = 0.05/2{,}957 = 1.7 \times 10^{-5}$) was reached for both *quantitative GAF* and *GAF extremes* for an LD-block on chromosome 15 (hg38: chr15:48965004 - 49464789 bp). The dashed vertical line highlights the coinciding location of significance.

ies into more powerful common effect estimates. Mega-analysis also increased *specificity* since discordant effect directions across studies will (partially) cancel into small(er) average effects, which suppresses their detection. SKAT exact *p*-values were obtained by Davies method (Davies, 1980). For the 2957 SKAT tests performed, the multiple-testing adjusted significance threshold was $\alpha = 1.7 \times 10^{-5}$ (Bonferroni).

#### 2.3.3. FIERS - step III: detailed insight into significant regions

For the 44 SNPs in the detected significant LD-block, single-SNP association tests were performed within studies and meta-analytically combined between studies by Fisher's *p*-value pooling (Fisher, 1925). Furthermore, individual best-estimate haplotypes on the five most strongly associated SNPs and an enclosed nsSNP were determined with PLINK. Haplotype association was analyzed within studies assuming an additive model of the effect of a haplotype or group of haplotypes, combining results meta-analytically across studies by Fisher's *p*-value pooling.

## 3. Results

FIERS tested 2957 putatively relevant LD-based regions. Fig. 1 displays Manhattan-like plots of SKAT-derived *p*-values from cross-study mega-analysis of individual-level *GAIN/TGen* and *BOMA* data. SNP correlations within LD-based blocks were subsumed by SKAT tests. Hence SKAT tests of LD-based blocks are largely independent of one another. The traits *quantitative GAF* (all subjects) and *GAF extremes* (lowest versus highest study quartile) both identified the same significant LD-block on chromosome 15 (hg38: chr15:48965004 - 49464789 bp; *quantitative GAF* $p_{mega} = 1.3 \times 10^{-5}$, *GAF extremes* $p_{mega} = 5.6 \times 10^{-6}$).

Of the 44 SNPs contained in this associated LD-block (see Supplement for summary statistics), 26 had consistent single-SNP effects across studies and meta-analysis *p*-values $p_{meta} < 0.05$ for both traits (Fisher's *p*-value pooling of studies). Eighteen of these SNPs even had *p*-values $p_{study} < 0.05$ in both studies and traits (Fig. 2, top and middle panel). The five top-ranked SNPs ($p_{meta} < 5 \times 10^{-5}$, Table 1) have strongly positively correlated minor allele dosages ($r > 0.67$); rs586758 and rs1904317 are nearly synonymous ($r = 0.998$). In the vicinity lies a nsSNP (rs11854184); its minor allele dosage is negatively correlated with that of the five top-ranked SNPs (range $= -0.25 > r > -0.34$). Individual best-estimate haplotypes on these six SNPs (estimated with PLINK) were nonambiguous on the last three positions and grouped accordingly. This revealed a protective haplotype group (***GCC, *quantitative GAF:* $p_{meta} = 1.1 \times 10^{-5}$, *GAF extremes:* $p_{meta} = 4.6 \times 10^{-6}$) and a risk haplotype group (***ATT, *quantitative GAF:* $p_{meta} = 2.4 \times 10^{-5}$, *GAF*

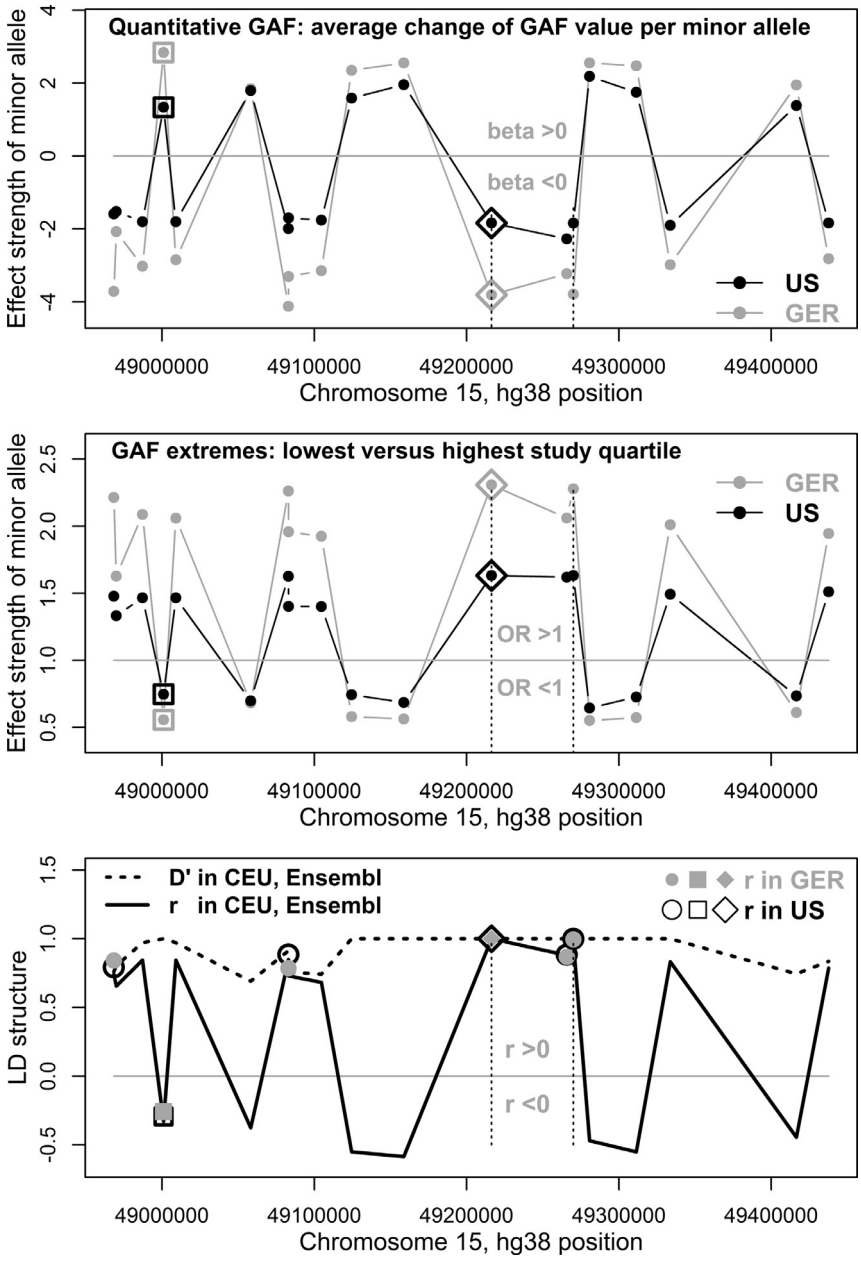



**Fig. 2** Consistent effect strength across BD outpatient samples.
In the associated LD-block hg38: chr15:48965004 – 49464789 bp, estimated minor allele effects were consistent across studies (US: *GAIN*/*TGen*, GER: *BOMA*) for *quantitative GAF* (top, additive effect per minor allele) and *GAF extremes* (middle, multiplicative effect per minor allele; within-study single-SNP analyses, adjusted for sex and illness duration). Results are displayed for 18 SNPs that had *p*-values $p_{study} < 0.05$ in each study and for nsSNP rs11854184 (square). Sign of effect estimates (risk: beta $< 0$, OR $> 1$; protective: beta $> 0$, OR $< 1$) corresponds to the correlation *r* of minor allele dosages (bottom panel, solid line) with *CTCF* binding site rs586758 (diamond). The latter is part of the discovered rs586758 – rs2086256 – rs1904317 haplotype (vertical lines, hg38: chr15:49216375 – 49270069 bp) and yielded the strongest association evidence among the 5 top-ranked SNPs ($p_{meta} < 5 \times 10^{-5}$, Table 1 middle panel). The 5 top-ranked SNPs have strongly positively correlated minor allele dosages ($r > 0.67$ bottom panel, *BOMA*: filled symbols, *GAIN*/*TGen*: open symbols; square: nsSNP rs11854184). Lines (*D'*: dashed, *r*: solid) display the highly similar LD structure of a CEU reference population (1000 Genomes phase 3, northern and western European ancestry, obtained from Ensembl (Cunningham et al., 2015, http://www.ensembl.org/).

*extremes:* $p_{meta} = 2.0 \times 10^{-6}$, Fisher's *p*-value pooling of studies). Between-study consistency of haplotype association is displayed in Table 1. A consistent reduction or increase of the risk of poor GAF, respectively, was also observed in all members of the two haplotype groups that were frequent enough for separate association testing (data not shown).

## 4. Discussion

Functional outcome in outpatient care is an important cross-diagnostic indicator for course of psychiatric disorder, clinically highly relevant, and highly variable in BD. To the best of our knowledge, this is the first SNP-based investigation into the genetic basis of GAF in outpatient care. Despite moderate sample size, we identified a significant genomic region by introducing the efficient test strategy FIERS. Plausibility of our finding on chromosome 15 (hg38: chr15:48965004 – 49464789 bp) is supported by consistency of effect strength between independent BD patient samples (Fig. 2) and between genotyped (*BOMA*) and imputed SNPs (*GAIN/TGen*). Moreover, further underlining plausibility, the association evidence centers at a functional SNP, and the associated region overlaps with and lies in the vicinity of genes that are relevant for neuronal differentiation and function (see below).

Single-SNP and haplotype analyses indicate that nsSNP rs11854184 is not likely responsible for the association. The reduced power of rs11854184 compared to the top-ranked SNPs (Table 1) cannot sufficiently be explained by its lower minor allele frequency. In contrast, rs586758 displayed the strongest single-SNP association, was part of the discovered haplotypes (G/A position) and is a *CTCF* binding site (Cunningham et al., 2015). *CTCF* sites regulate groups of genes within a chromatin domain.

GAF is not a measure of cognition and may deteriorate in psychiatric disorder for several reasons as it comprises social, occupational, and psychological functioning into a single score. Nevertheless, functional outcome and degree of cognitive impairment are significantly associated in BD and schizophrenia patients (Bowie et al., 2010). Furthermore, GAF scores and cognitive performance were lower in healthy carriers of neuropsychiatric copy-number-variants compared to non-carriers (Stefansson et al., 2013). The novel associated haplotype reported herein is located in the promoter flanking region of the *COPS2* gene (COP9 signalosome subunit 2, also known as *TRIP15*, *CSN2*, *ALIEN*), near microRNA4716 as plausible biological candidates. *COPS2* is involved in cell cycle regulation and DNA repair, mediates gene silencing, and participates in modulating hormone response and cell proliferation (Papaioannou, 2007). Moreover, functional studies demonstrated that *COPS2* plays crucial roles in neuronal differentiation and development as well as in maintaining neuronal functions (Akiyama et al., 2003; Chaerkady et al., 2011). Adjacent to the significantly associated LD-block, three additional genes are of interest: upstream *EID1* (EP300 interacting inhibitor of differentiation 1) which influences synaptic plasticity and memory function (Liu et al., 2012) and *SHC4* (Src homology 2 domain containing family member 4, also known as *ShcD*) which contributes to the regulation of neuronal function through mediation of the tyrosine kinase receptor TrkB downstream signaling pathway (You et al., 2010). Downstream, gene *DTWD1* (DTW domain containing 1) has previously been implicated in a pharmacogenomics study on side-effects of antidepressant treatment (Clark et al., 2012). Hence it is conceivable that *DTWD1* regulation through *CTCF* binding at rs586758 might alter GAF by altering side-effects of psychopharmacological medication and hence medication adherence.

A particular strength of this study is the test strategy. FIERS contributes to more powerful analyses of existing genome-wide data in general and even enables successful genomic analyses of moderately sized samples. Using general prior knowledge on putative function and LD, FIERS better focused association screening on relevant parts of the genome which greatly reduced the number of statistical tests performed. Across GWAS, small *p*-values are especially enriched among SNPs that are in LD with specific functional elements (Schork et al., 2013). FIERS exploits this by jointly testing SNPs within *LD-based regions* opposed to only testing functionally annotated SNPs. Among the variety of functional annotations that may be used to select LD-blocks *a priori,* nsSNPs have been most extensively validated so far. Currently, SIFT and PolyPhen provide one of the most widely accepted and accurate (Saunders and Baker, 2002) annotations (nsSNPs) whereas it is still difficult to annotate and predict non-coding SNPs (Li and Wei, 2015). Analyzing nsSNP-containing LD-blocks focused this analysis on protein-coding regions of the genome with the extension that exploiting LD putatively included additional information from SNPs with other functionalities as well.

A further strength of this investigation is cross-study mega-analysis, i.e. joint analysis of individual-level data across studies within SKAT. With appropriate covariate-adjustments, mega-analysis uses the data most efficiently and yields the greatest power. Mega-analysis within SKAT assumed concordant SNP effects across studies. This increased sensitivity for detecting replicable genetic effects and increased specificity by suppressing detection of discordant effects. In comparison, meta-analysis by Fisher's *p*-value pooling of separately analyzed covariate-adjusted studies was less sensitive and less specific but confirmed the reported significance on chromosome 15, albeit with lower power (data not shown). If mega-analysis should become infeasible (e.g., because studies have different covariates to accommodate or individual-level data cannot be shared), SKAT score statistics may also meta-analytically, i.e. on the level of summary statistics, account for between-study concordance of SNP effects (Lee et al., 2013).

Mandatory for power of any statistical method is that size of the unit of analysis (LD-block, gene, pathway) and model complexity (main effects, genetic interactions) are appropriate in relation to available sample size. Although large from a clinical perspective, available sample size was a study limitation. We mastered this challenge by testing putatively functionally relevant LD-blocks for main effects. For larger samples, natural extensions are analyzing genes or pathways and allowing for genetic interactions (Liu et al., 2007). In general, summary statistics on biological units either use select representative SNPs (Li et al., 2011, 2012) or aggregate association evidence from *all* contained SNPs. Aggregation is easily, exactly, and powerfully accomplished

by SKAT on individual-level data. In contrast, tedious corrections for SNP correlations are required when aggregating single-SNP *p*-values, e.g. by Fisher combination test (de Leeuw et al., 2016; Li et al., 2011). Other *p*-value based joint tests such as count-based (SNP-ratio or hypergeometric test) and rank-based (Kolmogorov–Smirnov) enrichment statistics suffer similar drawbacks as the Fisher combination test but with lower power (de Leeuw et al., 2016). Using representative SNPs instead of fully aggregating all evidence is more powerful only when causal SNPs are greatly outnumbered within tested SNP sets (Li et al., 2011) or when causal SNPs cannot share association signals well with other SNPs, e.g., due to low minor allele frequency (Li et al., 2012). Analogously, single-SNP tests may be more powerful than aggregate tests if associations are strong but involve very few SNPs only (Chen et al., 2014). Otherwise, SKAT is very often among the most powerful methods, and is robustly powerful for a broad range of genetic architectures (see below) (Chen et al., 2014; Li et al., 2012). Since SKAT tests combine individual-level information of multiple SNPs and their correlations, they exploit most of the information used in genotype imputation – without doing imputation (Howey and Cordell, 2014). SKAT can also analyze sequence and rare variants (Malzahn et al., 2016; Wu et al., 2011). However, LD should always be estimated on sufficiently frequent SNPs to avoid premature division of LD-blocks. As a self-contained test (de Leeuw et al., 2016), SKAT evaluates whether any of the jointly tested SNPs associates with a trait of interest. Aggregation of associations (multiple loci, pathway effects) but also relatively localized yet sufficiently strong associations such as polygenes or minor genes within pathways may make SKAT significant. This consideration hardly makes a difference for the LD-block analyses presented herein. However, it highlights that for correct *data interpretation*, genetic architectures underlying SKAT significances should be examined.

So far, the success of psychiatric genetics is largely based on the strategy of founding large consortia for case-control studies. However, data on clinically important *quantitative* phenotypes is still limited, largely due to high costs of deep phenotyping and lacking harmonization of assessment scales and conditions across studies. Owing to this, a potential limitation of the present investigation is that only two independent studies were available. That significance was reached in the total sample but not in single studies is typical for GWAS which commonly regard consistency of effect estimates across studies (Fig. 2) as additional conclusive evidence. Nevertheless, our consistent finding that rs586758 – rs2086256 – rs1904317 haplotype ATT carriers (*BOMA*: 49.2% ± 4.3%, *GAIN/TGen*: 50.6% ± 3.0%) have lower GAF values would require additional independent validation. Furthermore, no information on medication or medication adherence was available and GAF assessment differed to some extent between studies. GAF was assessed at a *time point* (*BOMA*: pre-admission), or *averaged* over a period (*GAIN/TGen*: past month) during which the state of illness, although sufficiently remitted for outpatient care, may have varied. Statistical analyses were adjusted for between-study differences of GAF values. Nevertheless, differences of GAF assessment might yield phenotypes with slightly different underlying biological mechanisms. This may explain why genetic effects in the associated region (Table 1 and Fig. 2), while consistent across studies, were slightly stronger in the putatively better remitted *BOMA* sample compared to *GAIN/TGen*.

GAF is an overall rating of a patient's psychological, social and occupational functioning. While clinically highly relevant and commonly used, a single overall score also presents some limitations. Specifically, GAF scores lack information regarding which of the three domains was most impaired and most decisive for individual overall rating. For example, a suicidal person with well-functioning relationships and good performance at his or her job would be assigned a very low GAF score. Hence future research may proceed by operationalizing functional outcome with a more differentiated measure like e.g. the functioning assessment short test (FAST; Rosa et al., 2007). Furthermore, when analyzing GAF scores, it would be of interest to stratify or adjust analyses with respect to concomitant symptom severity. Unfortunately, we did not have sufficient data for this, which is a study limitation.

While low GAF scores may occur in psychiatric patients for clinically different reasons, the generality of the GAF can also be seen as an advantage: GAF is applicable across different psychiatric diagnoses that share a common polygenic background (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013; Forstner et al., 2017; Purcell et al., 2009) and could be an indicator of a more general resilience/ vulnerability factor. Hence it would be of great interest to analyze GAF or other measures of functional outcome also in other psychiatric disorders, such as schizophrenia, and jointly in patients with different psychiatric disorders.

## Conflicts of interests

John I. Nurnberger Jr. is a consultant and investigator for Janssen and an investigator for Assurex. All other authors declare no conflicts of interests.

## Author contributions

DM and TGS designed the study. Data (sample collection, phenotyping, genotyping, quality control, linkage disequilibrium boundaries, pathway information) were provided by SF, NA-R, SA, JAB, WHB, CSB, WB, SC, WC, DWC, FD, HJE, TF, AJF, JF, ESG, FSG, TAG, YG, MH, LH, BJK, DLK, WBL, CL, PBM, MGM, FJM, SMM, TWM, SSM, CMN, JIN Jr, EAN, JBP, DQ, JR, JCR, WAS, NJS, TS, PDS, ENS, FS, JS, SS, JT, SHW, PPZ, PZ, SZ, JRK, MMN, MR and TGS. DM conducted the statistical analyses. Data were interpreted by all authors. DM and MB wrote the manuscript. All authors critically reviewed the manuscript and have approved the final version.

## Role of funding source

This research was funded by the Deutsche Forschungsgemeinschaft DFG (grants Klinische Forschergruppe (KFO) 241/ PsyCourse SCHU 1603/4-1, SCHU 1603/5-1, BI 576/5-1, FA 241/16-1 and SCHU 1603/7-1; Research Training Group "Scaling Problems in Statistics" RTG 1644; FOR2107:

RI908/11-1 and NO246/10-1; as well as HEIDE DFG RI908/8-1), and by the German Federal Ministry of Education and Research (BMBF, grants: IntegraMent 01ZX1314G, 01ZX1314A, 01ZX1314K, SysMedAlcoholism 01ZX1311A). Thomas G. Schulze was also supported by the Dr. Lisa-Oehler-Foundation (Kassel, Germany).

Sven Cichon acknowledges support by the Swiss National Science Foundation (SNSF grant 310030_156791 to SC) and from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 7202070 (HBP SGA1).

Genotype and phenotype data of the European American BD sample were provided by the *Collaborative Genomic Study of Bipolar Disorder* and the *Bipolar Genome Study*. These consortia acknowledge funding support by the National Institute of Mental Health (NIMH) and genotyping provided through the *Genetic Association Information Network* (*GAIN*). Data and biomaterials were collected in or as part of the following projects that participated in the NIMH Bipolar Disorder Genetics Initiative. From 1991 to 1998, the principal Investigators and co-investigators were: Indiana University, Indianapolis, IN, U01 MH46282, John Nurnberger, M.D., Ph.D., Marvin Miller, M.D., Howard J. Edenberg, Ph.D. and Elizabeth Bowman, M.D.; Washington University, St. Louis, MO, U01 MH46280, Theodore Reich, M.D., Allison Goate, Ph.D., and John Rice, Ph.D.; Johns Hopkins University, Baltimore, MD, U01 MH46274, J. Raymond DePaulo, Jr., M.D., Sylvia Simpson, M.D., MPH, and Colin Stine, Ph.D.; NIMH Intramural Research Program, Clinical Neurogenetics Branch, Bethesda, MD, Elliot Gershon, M.D., Diane Kazuba, B.A., and Elizabeth Maxwell, M.S.W. From 1999 to 2007 (NIMH studies 1 and 40), the principal investigators and co-investigators were: Indiana University, Indianapolis, IN, R01 MH59545, John Nurnberger, M.D., Ph.D., Marvin J. Miller, M.D., Elizabeth S. Bowman, M.D., N. Leela Rau, M.D., P. Ryan Moe, M.D., Nalini Samavedy, M.D., Rif El-Mallakh, M.D. (at University of Louisville), Husseini Manji, M.D. (at Wayne State University, Johnson and Johnson), Debra A. Glitz, M.D. (at Wayne State University), Eric T. Meyer, Ph.D., M.S. (at Oxford University, UK), Carrie Smiley, R.N., Tatiana Foroud, Ph.D., Leah Flury, M.S., Danielle M. Dick, Ph.D. (at Virginia Commonwealth University), Howard J. Edenberg, Ph.D.; Washington University, St. Louis, MO, R01 MH059534, John Rice, Ph.D, Theodore Reich, M.D., Allison Goate, Ph.D., Laura Bierut, M.D. K02 DA21237; Johns Hopkins University, Baltimore, MD, R01 MH59533, Melvin McInnis, M.D., J. Raymond DePaulo, Jr., M.D., Dean F. MacKinnon, M.D., Francis M. Mondimore, M.D., James B. Potash, M.D., Peter P. Zandi, Ph.D, Dimitrios Avramopoulos, and Jennifer Payne; University of Pennsylvania, PA, R01 MH59553, Wade Berrettini, M.D., Ph.D.; University of California at Irvine, University of California at San Francisco, CA, R01 MH60068, William Byerley, M.D., Mark Vawter, M.D., and Sophia Vinogradov, M.D.; University of Iowa, IA, R01 MH059548, William Coryell, M.D., and Raymond Crowe, M.D.; University of Chicago, IL, R01 MH59535, Elliot Gershon, M.D., Judith Badner, Ph.D., Francis McMahon, M.D., Chunyu Liu, Ph.D., Alan Sanders, M.D., Maria Caserta, Steven Dinwiddie, M.D., Tu Nguyen, Donna Harakal; University of California at San Diego, CA, R01 MH59567, John Kelsoe, M.D., Rebecca McKinney, B.A.; Rush University, IL, R01 MH059556, William Scheftner, M.D., Howard M. Kravitz, D.O., M.P.H., Diana Marta, B.S., Annette Vaughn-Brown, M.S.N., R.N., and Laurie Bederow, M.A.; NIMH Intramural Research Program, Bethesda, MD, 1Z01MH002810-01, Francis J. McMahon, M.D., Layla Kassem, PsyD, Sevilla Detera-Wadleigh, Ph.D, Lisa Austin, Ph.D, Dennis L. Murphy, M.D.; Howard University, MH070013, William B. Lawson, M.D., Ph.D., Evaristus Nwulia, M.D., and Maria Hipolito, M.D. Furthermore, the *Bipolar Genome Study* was supported by the National Institutes of Health grants P50CA89392 from the National Cancer Institute and 5K02DA021237 from the National Institute of Drug Abuse.

None of the funding sponsors had any further role in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication.

## Acknowledgments

We would like to express our wholehearted thanks to everyone who participated in the *BOMA* and *GAIN*/*TGen* studies.

## Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.euroneuro.2018.10.005.

## References

Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., Sunyaev, S.R., 2010. A method and server for predicting damaging missense mutations. Nat. Methods 7, 248–249. https://doi.org/10.1038/nmeth0410-248.

Akiyama, H., Sugiyama, A., Uzawa, K., Fujisawa, N., Tashiro, Y., Tashiro, F., 2003. Implication of Trip15/CSN2 in early stage of neuronal differentiation of P19 embryonal carcinoma cells. Brain Res. Dev. Brain Res. 140, 45–56.

American Psychiatric Association, 2002. Diagnostic and Statistical Manual of Mental Disorders, 4th ed. American Psychiatric Association, US, Washington, DC text revision.

Barrett, J.C., Fry, B., Maller, J., Daly, M.J., 2005. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21, 263–265. https://doi.org/10.1093/bioinformatics/bth457.

Bienvenu, O.J., Davydow, D.S., Kendler, K.S., 2011. Psychiatric "diseases" versus behavioral disorders and degree of genetic influence. Psychol. Med. 41, 33–40. https://doi.org/10.1017/S003329171000084X.

Bowie, C.R., Depp, C., McGrath, J.A., Wolyniec, P., Mausbach, B.T., Thornquist, M.H., Luke, J., Patterson, T.L., Harvey, P.D., Pulver, A.E., 2010. Prediction of real-world functional disability in chronic mental disorders: a comparison of schizophrenia and bipolar disorder. Am. J. Psychiatry 167, 1116–1124. https://doi.org/10.1176/appi.ajp.2010.09101406.

Chaerkady, R., Letzen, B., Renuse, S., Sahasrabuddhe, N.A., Kumar, P., All, A.H., Thakor, N.V., Delanghe, B., Gearhart, J.D., Pandey, A., Kerr, C.L., 2011. Quantitative temporal proteomic analysis of human embryonic stem cell differentiation into oligodendrocyte progenitor cells. Proteomics 11, 4007–4020. https://doi.org/10.1002/pmic.201100107.

Charney, A.W., Ruderfer, D.M., Stahl, E.A., Moran, J.L., Chambert, K., Belliveau, R.A., Forty, L., Gordon-Smith, K.,

Di Florio, A., Lee, P.H., Bromet, E.J., Buckley, P.F., Escamilla, M.A., Fanous, A.H., Fochtmann, L.J., Lehrer, D.S., Malaspina, D., Marder, S.R., Morley, C.P., Nicolini, H., Perkins, D.O., Rakofsky, J.J., Rapaport, M.H., Medeiros, H., Sobell, J.L., Green, E.K., Backlund, L., Bergen, S.E., Juréus, A., Schalling, M., Lichtenstein, P., Roussos, P., Knowles, J.A., Jones, I., Jones, L.A., Hultman, C.M., Perlis, R.H., Purcell, S.M., McCarroll, S.A., Pato, C.N., Pato, M.T., Craddock, N., Landén, M., Smoller, J.W., Sklar, P., 2017. Evidence for genetic heterogeneity between clinical subtypes of bipolar disorder. Transl. Psychiatry 7, e993. https://doi.org/10.1038/tp.2016.242.

Chen, H., Malzahn, D., Balliu, B., Li, C., Bailey, J.N., 2014. Testing genetic association with rare and common variants in family data: rare and common variant tests in families. Genet. Epidemiol. 38, S37–S43. https://doi.org/10.1002/gepi.21823.

Cichon, S., Mühleisen, T.W., Degenhardt, F.A., Mattheisen, M., Miró, X., Strohmaier, J., Steffens, M., Meesters, C., Herms, S., Weingarten, M., Priebe, L., Haenisch, B., Alexander, M., Vollmer, J., Breuer, R., Schmäl, C., Tessmann, P., Moebus, S., Wichmann, H.-E., Schreiber, S., Müller-Myhsok, B., Lucae, S., Jamain, S., Leboyer, M., Bellivier, F., Etain, B., Henry, C., Kahn, J.-P., Heath, S., Hamshere, M., O'Donovan, M.C., Owen, M.J., Craddock, N., Schwarz, M., Vedder, H., Kammerer-Ciernioch, J., Reif, A., Sasse, J., Bauer, M., Hautzinger, M., Wright, A., Mitchell, P.B., Schofield, P.R., Montgomery, G.W., Medland, S.E., Gordon, S.D., Martin, N.G., Gustafsson, O., Andreassen, O., Djurovic, S., Sigurdsson, E., Steinberg, S., Stefansson, H., Stefansson, K., Kapur-Pojskic, L., Oruc, L., Rivas, F., Mayoral, F., Chuchalin, A., Babadjanova, G., Tiganov, A.S., Pantelejeva, G., Abramova, L.I., Grigoroiu-Serbanescu, M., Diaconu, C.C., Czerski, P.M., Hauser, J., Zimmer, A., Lathrop, M., Schulze, T.G., Wienker, T.F., Schumacher, J., Maier, W., Propping, P., Rietschel, M., Nöthen, M.M., 2011. Genome-wide association study identifies genetic variation in neurocan as a susceptibility factor for bipolar disorder. Am. J. Hum. Genet. 88, 372–381. https://doi.org/10.1016/j.ajhg.2011.01.017.

Clark, S.L., Adkins, D.E., Aberg, K., Hettema, J.M., McClay, J.L., Souza, R.P., van den Oord, E.J.C.G., 2012. Pharmacogenomic study of side-effects for antidepressant treatment options in STAR*D. Psychol. Med. 42, 1151–1162. https://doi.org/10.1017/S003329171100239X.

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, 1371–1379. https://doi.org/10.1016/S0140-6736(12)62129-1.

Cunningham, F., Amode, M.R., Barrell, D., Beal, K., Billis, K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fitzgerald, S., Gil, L., Giron, C.G., Gordon, L., Hourlier, T., Hunt, S.E., Janacek, S.H., Johnson, N., Juettemann, T., Kahari, A.K., Keenan, S., Martin, F.J., Maurel, T., McLaren, W., Murphy, D.N., Nag, R., Overduin, B., Parker, A., Patricio, M., Perry, E., Pignatelli, M., Riat, H.S., Sheppard, D., Taylor, K., Thormann, A., Vullo, A., Wilder, S.P., Zadissa, A., Aken, B.L., Birney, E., Harrow, J., Kinsella, R., Muffato, M., Ruffier, M., Searle, S.M.J., Spudich, G., Trevanion, S.J., Yates, A., Zerbino, D.R., Flicek, P., 2015. Ensembl 2015. Nucleic Acids Res. 43, D662–D669. https://doi.org/10.1093/nar/gku1010.

Davies, R.B., 1980. Algorithm AS 155: the distribution of a linear combination of $\chi^2$ random variables. Appl. Stat. 29, 323–333. https://doi.org/10.2307/2346911.

Dayem Ullah, A.Z., Lemoine, N.R., Chelala, C., 2013. A practical guide for the functional annotation of genetic variations using SNPnexus. Brief. Bioinform. 14, 437–447. https://doi.org/10.1093/bib/bbt004.

de Leeuw, C.A., Neale, B.M., Heskes, T., Posthuma, D., 2016. The statistical properties of gene-set analysis. Nat. Rev. Genet 17, 353–364. https://doi.org/10.1038/nrg.2016.29.

Dudbridge, F., 2013. Power and predictive accuracy of polygenic risk scores. PLoS Genet. 9, e1003348. https://doi.org/10.1371/journal.pgen.1003348.

Endicott, J., Spitzer, R.L., Fleiss, J.L., Cohen, J., 1976. The global assessment scale. A procedure for measuring overall severity of psychiatric disturbance. Arch. Gen. Psychiatry 33, 766–771.

Fangerau, H., Ohlraun, S., Granath, R.O., Nöthen, M.M., Rietschel, M., Schulze, T.G., 2005. Computer-assisted phenotype characterization for genetic research in psychiatry. Hum. Hered. 58, 122–130. https://doi.org/10.1159/000083538.

First, M.B., Spitzer, R.L., Gibbon, M., Williams, J.B.W., 2002. Structured clinical interview for DSM-IV-TR axis I disorders. Research version, Patient edition. (SCID-I/P). Biometrics Research. New York State Psychiatric Institute, US, New York.

Fisher, R.A., 1925. Statistical Methods for Research Workers. Oliver and Boyd, Edinburgh, Scotland.

Forstner, A.J., Hecker, J., Hofmann, A., Maaser, A., Reinbold, C.S., Mühleisen, T.W., Leber, M., Strohmaier, J., Degenhardt, F., Treutlein, J., Mattheisen, M., Schumacher, J., Streit, F., Meier, S., Herms, S., Hoffmann, P., Lacour, A., Witt, S.H., Reif, A., Müller-Myhsok, B., Lucae, S., Maier, W., Schwarz, M., Vedder, H., Kammerer-Ciernioch, J., Pfennig, A., Bauer, M., Hautzinger, M., Moebus, S., Schenk, L.M., Fischer, S.B., Sivalingam, S., Czerski, P.M., Hauser, J., Lissowska, J., Szeszenia-Dabrowska, N., Brennan, P., McKay, J.D., Wright, A., Mitchell, P.B., Fullerton, J.M., Schofield, P.R., Montgomery, G.W., Medland, S.E., Gordon, S.D., Martin, N.G., Krasnov, V., Chuchalin, A., Babadjanova, G., Pantelejeva, G., Abramova, L.I., Tiganov, A.S., Polonikov, A., Khusnutdinova, E., Alda, M., Cruceanu, C., Rouleau, G.A., Turecki, G., Laprise, C., Rivas, F., Mayoral, F., Kogevinas, M., Grigoroiu-Serbanescu, M., Becker, T., Schulze, T.G., Rietschel, M., Cichon, S., Fier, H., Nöthen, M.M., 2017. Identification of shared risk loci and pathways for bipolar disorder and schizophrenia. PLOS One 12, e0171595. https://doi.org/10.1371/journal.pone.0171595.

Friedrichs, S., Malzahn, D., Pugh, E.W., Almeida, M., Liu, X.Q., Bailey, J.N., 2016. Filtering genetic variants and placing informative priors based on putative biological function. BMC Genet. 17 (Suppl 2), 8. https://doi.org/10.1186/s12863-015-0313-x.

Gade, K., Malzahn, D., Anderson-Schmidt, H., Strohmaier, J., Meier, S., Frank, J., Falkai, P.G., Rietschel, M., Schulze, T.G., 2015. Functional outcome in major psychiatric disorders and associated clinical and psychosocial variables: a potential cross-diagnostic phenotype for further genetic investigations? World J. Biol. Psychiatry 16, 237–248. https://doi.org/10.3109/15622975.2014.995221.

Hou, L., Bergen, S.E., Akula, N., Song, J., Hultman, C.M., Landén, M., Adli, M., Alda, M., Ardau, R., Arias, B., Aubry, J.-M., Backlund, L., Badner, J.A., Barrett, T.B., Bauer, M., Baune, B.T., Bellivier, F., Benabarre, A., Bengesser, S., Berrettini, W.H., Bhattacharjee, A.K., Biernacka, J.M., Birner, A., Bloss, C.S., Brichant-Petitjean, C., Bui, E.T., Byerley, W., Cervantes, P., Chillotti, C., Cichon, S., Colom, F., Coryell, W., Craig, D.W., Cruceanu, C., Czerski, P.M., Davis, T., Dayer, A., Degenhardt, F., Del Zompo, M., DePaulo, J.R., Edenberg, H.J., Étain, B., Falkai, P., Foroud, T., Forstner, A.J., Frisén, L., Frye, M.A., Fullerton, J.M., Gard, S., Garnham, J.S., Gershon, E.S., Goes, F.S., Greenwood, T.A., Grigoroiu-Serbanescu, M., Hauser, J., Heilbronner, U., Heilmann-Heimbach, S., Herms, S., Hipolito, M., Hitturlingappa, S., Hoffmann, P., Hofmann, A., Jamain, S., Jiménez, E., Kahn, J.-P., Kassem, L., Kelsoe, J.R., Kittel-Schneider, S., Kliwicki, S., Koller, D.L., König, B., Lackner, N., Laje, G., Lang, M., Lavebratt, C., Lawson, W.B., Leboyer, M., Leckband, S.G., Liu, C., Maaser, A., Mahon, P.B., Maier, W., Maj, M., Manchia, M., Martinsson, L., McCarthy, M.J., McElroy, S.L., McInnis, M.G., McKinney, R., Mitchell, P.B., Mitjans, M., Mondimore, F.M., Monteleone, P., Mühleisen, T.W., Nievergelt, C.M., Nöthen, M.M., Novák, T., Nurnberger, J.I.,

Nwulia, E.A., Ösby, U., Pfennig, A., Potash, J.B., Propping, P., Reif, A., Reininghaus, E., Rice, J., Rietschel, M., Rouleau, G.A., Rybakowski, J.K., Schalling, M., Scheftner, W.A., Schofield, P.R., Schork, N.J., Schulze, T.G., Schumacher, J., Schweizer, B.W., Severino, G., Shekhtman, T., Shilling, P.D., Simhandl, C., Slaney, C.M., Smith, E.N., Squassina, A., Stamm, T., Stopkova, P., Streit, F., Strohmaier, J., Szelinger, S., Tighe, S.K., Tortorella, A., Turecki, G., Vieta, E., Volkert, J., Witt, S.H., Wright, A., Zandi, P.P., Zhang, P., Zollner, S., McMahon, F.J., 2016. Genome-wide association study of 40,000 individuals identifies two novel loci associated with bipolar disorder. Hum. Mol. Genet. 25, 3383–3394. https://doi.org/10.1093/hmg/ddw181.

Howey, R., Cordell, H.J., 2014. Imputation Without Doing imputation: a new method for the detection of non-genotyped causal variants. Genet. Epidemiol. 38, 173–190. https://doi.org/10.1002/gepi.21792.

Kanehisa, M., Goto, S., 2000. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 28, 27–30.

Kumar, P., Henikoff, S., Ng, P.C., 2009. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat. Protoc. 4, 1073–1081. https://doi.org/10.1038/nprot.2009.86.

Lee, S., Teslovich, T.M., Boehnke, M., Lin, X., 2013. General framework for meta-analysis of rare variants in sequencing association studies. Am. J. Hum. Genet. 93, 42–53. https://doi.org/10.1016/j.ajhg.2013.05.010.

Li, L., Wei, D., 2015. Bioinformatics tools for discovery and functional analysis of single nucleotide polymorphisms. Adv. Exp. Med. Biol. 827, 287–310. https://doi.org/10.1007/978-94-017-9245-5_17.

Li, M.-X., Gui, H.-S., Kwan, J.S.H., Sham, P.C., 2011. GATES: a rapid and powerful gene-based association test using extended simes procedure. Am. J. Hum. Genet. 88, 283–293. https://doi.org/10.1016/j.ajhg.2011.01.019.

Li, M.-X., Kwan, J.S.H., Sham, P.C., 2012. HYST: a hybrid set-based test for genome-wide association studies, with application to protein-protein interaction-based association analysis. Am. J. Hum. Genet. 91, 478–488. https://doi.org/10.1016/j.ajhg.2012.08.004.

Liu, D., Lin, X., Ghosh, D., 2007. Semiparametric regression of multidimensional genetic pathway data: least-squares kernel machines and linear mixed models. Biometrics 63, 1079–1088. https://doi.org/10.1111/j.1541-0420.2007.00799.x.

Liu, R., Lei, J.X., Luo, C., Lan, X., Chi, L., Deng, P., Lei, S., Ghribi, O., Liu, Q.Y., 2012. Increased EID1 nuclear translocation impairs synaptic plasticity and memory function associated with pathogenesis of Alzheimer's disease. Neurobiol. Dis. 45, 902–912. https://doi.org/10.1016/j.nbd.2011.12.007.

Luborsky, L., 1962. Clinicians' judgments of mental health: a proposed scale. Arch. Gen. Psychiatry 7, 407–417. https://doi.org/10.1001/archpsyc.1962.01720060019002.

Malzahn, D., Friedrichs, S., Bickeböller, H., 2016. Comparing strategies for combined testing of rare and common variants in whole sequence and genome-wide genotype data. BMC Proc. 10 (Suppl 7), 269–273. https://doi.org/10.1186/s12919-016-0042-9.

Malzahn, D., Friedrichs, S., Rosenberger, A., Bickeböller, H., 2014. Kernel score statistic for dependent data. BMC Proc. 8 (Suppl 1), S41. https://doi.org/10.1186/1753-6561-8-S1-S41.

Nurnberger, J.I., Blehar, M.C., Kaufmann, C.A., York-Cooler, C., Simpson, S.G., Harkavy-Friedman, J., Severe, J.B., Malaspina, D., Reich, T., 1994. Diagnostic interview for genetic studies. Rationale, unique features, and training. NIMH Genetics Initiative. Arch. Gen. Psychiatry 51, 849–859 discussion 863–864.

Papaioannou, 2007. The coregulator Alien. Nucl. Recept. Signal 5, e008. https://doi.org/10.1621/nrs.05008.

Potash, J.B., Toolan, J., Steele, J., Miller, E.B., Pearl, J., Zandi, P.P., Schulze, T.G., Kassem, L., Simpson, S.G., Lopez, V., MacKinnon, D.F., McMahon, F.J.NIMH Genetics Initiative Bipolar Disorder Consortium, 2007. The bipolar disorder phenome database: a resource for genetic studies. Am. J. Psychiatry 164, 1229–1237. https://doi.org/10.1176/appi.ajp.2007.06122045.

Purcell, S.M., Wray, N.R., Stone, J.L., Visscher, P.M., O'Donovan, M.C., Sullivan, P.F., Sklar, P.International Schizophrenia Consortium, 2009. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature.. 10.1038/nature08185.

Reinares, M., Papachristou, E., Harvey, P., Mar Bonnín, C., Sánchez-Moreno, J., Torrent, C., Ayuso-Mateos, J.L., Ploubidis, G.B., Vieta, E., Frangou, S., 2013. Towards a clinical staging for bipolar disorder: defining patient subtypes based on functional outcome. J. Affect. Disord. 144, 65–71. https://doi.org/10.1016/j.jad.2012.06.005.

Rosa, A.R., Sánchez-Moreno, J., Martínez-Aran, A., Salamero, M., Torrent, C., Reinares, M., Comes, M., Colom, F., Van Riel, W., Ayuso-Mateos, J.L., Kapczinski, F., Vieta, E., 2007. Validity and reliability of the Functioning Assessment Short Test (FAST) in bipolar disorder. Clin. Pract. Epidemiol. Ment. Health CP EMH 3, 5. https://doi.org/10.1186/1745-0179-3-5.

Sabeti, P.C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., Xie, X., Byrne, E.H., McCarroll, S.A., Gaudet, R., Schaffner, S.F., Lander, E.S.The International HapMap Consortium, 2003. The International HapMap Project. Nature 426, 789–796. https://doi.org/10.1038/nature02168.

Sanchez-Moreno, J., Martinez-Aran, A., Tabarés-Seisdedos, R., Torrent, C., Vieta, E., Ayuso-Mateos, J.L., 2009. Functioning and disability in bipolar disorder: an extensive review. Psychother. Psychosom. 78, 285–297. https://doi.org/10.1159/000228249.

Saunders, C.T., Baker, D., 2002. Evaluation of structural and evolutionary contributions to deleterious mutation prediction. J. Mol. Biol. 322, 891–901.

Savage, R.M., Wiener, H.W., Nimgaonkar, V., Devlin, B., Calkins, M.E., Gur, R.E., O'Jile, J., Bradford, L.D., Edwards, N., Kwentus, J., Allen, T., McEvoy, J.P., Nasrallah, H., Santos, A.B., Aduroja, T., Lahti, A., May, R.S., Montgomery-Barefield, L., Go, R.C.P., 2012. Heritability of functioning in families with schizophrenia in relation to neurocognition. Schizophr. Res. 139, 105–109. https://doi.org/10.1016/j.schres.2012.04.015.

Schaid, D.J., 2010. Genomic Similarity and Kernel Methods I: advancements by building on mathematical and statistical foundations. Hum. Hered. 70, 109–131. https://doi.org/10.1159/000312641.

Schifano, E.D., Epstein, M.P., Bielak, L.F., Jhun, M.A., Kardia, S.L.R., Peyser, P.A., Lin, X., 2012. SNP set association analysis for familial data: SNP set analysis for familial data. Genet. Epidemiol. 36, 797–810. https://doi.org/10.1002/gepi.21676.

Schork, A.J., Thompson, W.K., Pham, P., Torkamani, A., Roddey, J.C., Sullivan, P.F., Kelsoe, J.R., O'Donovan, M.C., Furberg, H., Schork, N.J., Andreassen, O.A., Dale, A.M.The Schizophrenia Psychiatric Genomics Consortium, 2013. All SNPs are not created equal: genome-wide association studies reveal a consistent pattern of enrichment among functionally annotated SNPs. PLoS Genet. 9, e1003449. https://doi.org/10.1371/journal.pgen.1003449.

Schulze, T.G., Hedeker, D., Zandi, P., Rietschel, M., McMahon, F.J., 2006. What is familial about familial bipolar disorder? resemblance among relatives across a broad spectrum of phenotypic characteristics. Arch. Gen. Psychiatry 63 63, 1368–1376. https://doi.org/10.1001/archpsyc.63.12.1368.

Smith, E.N., Bloss, C.S., Badner, J.A., Barrett, T., Belmonte, P.L., Berrettini, W., Byerley, W., Coryell, W., Craig, D., Edenberg, H.J., Eskin, E., Foroud, T., Gershon, E., Greenwood, T.A., Hipolito, M., Koller, D.L., Lawson, W.B., Liu, C., Lohoff, F., McInnis, M.G., McMahon, F.J., Mirel, D.B., Murray, S.S., Nievergelt, C., Nurnberger, J., Nwulia, E.A., Paschall, J.,

Potash, J.B., Rice, J., Schulze, T.G., Scheftner, W., Panganiban, C., Zaitlen, N., Zandi, P.P., Zöllner, S., Schork, N.J., Kelsoe, J.R., 2009. Genome-wide association study of bipolar disorder in European American and African American individuals. Mol. Psychiatry 14, 755–763. https://doi.org/10.1038/mp.2009.43.

Smith, E.N., Koller, D.L., Panganiban, C., Szelinger, S., Zhang, P., Badner, J.A., Barrett, T.B., Berrettini, W.H., Bloss, C.S., Byerley, W., Coryell, W., Edenberg, H.J., Foroud, T., Gershon, E.S., Greenwood, T.A., Guo, Y., Hipolito, M., Keating, B.J., Lawson, W.B., Liu, C., Mahon, P.B., McInnis, M.G., McMahon, F.J., McKinney, R., Murray, S.S., Nievergelt, C.M., Nurnberger, J.I., Nwulia, E.A., Potash, J.B., Rice, J., Schulze, T.G., Scheftner, W.A., Shilling, P.D., Zandi, P.P., Zöllner, S., Craig, D.W., Schork, N.J., Kelsoe, J.R., 2011. Genome-wide association of bipolar disorder suggests an enrichment of replicable associations in regions near genes. PLoS Genet. 7, e1002134. https://doi.org/10.1371/journal.pgen.1002134.

Solé, B., Bonnin, C.M., Jiménez, E., Torrent, C., Torres, I., Varo, C., Valls, E., Montejo, L., Gómez-Ocaña, C., Tomioka, Y., Vieta, E., Martinez-Aran, A., Reinares, M., 2018. Heterogeneity of functional outcomes in patients with bipolar disorder: a cluster-analytic approach. Acta Psychiatr. Scand. 137, 516–527. https://doi.org/10.1111/acps.12871.

Stefansson, H., Meyer-Lindenberg, A., Steinberg, S., Magnusdottir, B., Morgen, K., Arnarsdottir, S., Bjornsdottir, G., Walters, G.B., Jonsdottir, G.A., Doyle, O.M., Tost, H., Grimm, O., Kristjansdottir, S., Snorrason, H., Davidsdottir, S.R., Gudmundsson, L.J., Jonsson, G.F., Stefansdottir, B., Helgadottir, I., Haraldsson, M., Jonsdottir, B., Thygesen, J.H., Schwarz, A.J., Didriksen, M., Stensbøl, T.B., Brammer, M., Kapur, S., Halldorsson, J.G., Hreidarsson, S., Saemundsen, E., Sigurdsson, E., Stefansson, K., 2013. CNVs conferring risk of autism or schizophrenia affect cognition in controls. Nature 505, 361–366. https://doi.org/10.1038/nature12818.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., Mesirov, J.P., 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. 102, 15545–15550. https://doi.org/10.1073/pnas.0506580102.

Sum, M.Y., Ho, N.F., Sim, K., 2015. Cross diagnostic comparisons of quality of life deficits in remitted and unremitted patients with schizophrenia and bipolar disorder. Schizophr. Res. 168, 191–196. https://doi.org/10.1016/j.schres.2015.08.030.

Vassos, E., Sham, P.C., Cai, G., Deng, H., Liu, X., Sun, X., Zhao, J., Murray, R.M., Collier, D.A., Li, T., 2008. Correlation and familial aggregation of dimensions of psychosis in affected sibling pairs from China. Br. J. Psychiatry 193, 305–310. https://doi.org/10.1192/bjp.bp.107.046037.

Vos, T., Flaxman, A.D., Naghavi, M., Lozano, R., Michaud, C., Ezzati, M., Shibuya, K., Salomon, J.A., Abdalla, S., Aboyans, V., Abraham, J., Ackerman, I., Aggarwal, R., Ahn, S.Y., Ali, M.K., AlMazroa, M.A., Alvarado, M., Anderson, H.R., Anderson, L.M., Andrews, K.G., Atkinson, C., Baddour, L.M., Bahalim, A.N., Barker-Collo, S., Barrero, L.H., Bartels, D.H., Basáñez, M.-G., Baxter, A., Bell, M.L., Benjamin, E.J., Bennett, D., Bernabé, E., Bhalla, K., Bhandari, B., Bikbov, B., Abdulhak, A.B., Birbeck, G., Black, J.A., Blencowe, H., Blore, J.D., Blyth, F., Bolliger, I., Bonaventure, A., Boufous, S., Bourne, R., Boussinesq, M., Braithwaite, T., Brayne, C., Bridgett, L., Brooker, S., Brooks, P., Brugha, T.S., Bryan-Hancock, C., Bucello, C., Buchbinder, R., Buckle, G., Budke, C.M., Burch, M., Burney, P., Burstein, R., Calabria, B., Campbell, B., Canter, C.E., Carabin, H., Carapetis, J., Carmona, L., Cella, C., Charlson, F., Chen, H., Cheng, A.T.-A., Chou, D., Chugh, S.S., Coffeng, L.E., Colan, S.D., Colquhoun, S., Colson, K.E., Condon, J., Connor, M.D., Cooper, L.T., Corriere, M., Cortinovis, M., de Vaccaro, K.C., Couser, W., Cowie, B.C., Criqui, M.H., Cross, M., Dabhadkar, K.C., Dahiya, M., Dahodwala, N., Damsere-Derry, J., Danaei, G., Davis, A., De Leo, D., Degenhardt, L., Dellavalle, R., Delossantos, A., Denenberg, J., Derrett, S., Des Jarlais, D.C., Dharmaratne, S.D., Dherani, M., Diaz-Torne, C., Dolk, H., Dorsey, E.R., Driscoll, T., Duber, H., Ebel, B., Edmond, K., Elbaz, A., Ali, S.E., Erskine, H., Erwin, P.J., Espindola, P., Ewoigbokhan, S.E., Farzadfar, F., Feigin, V., Felson, D.T., Ferrari, A., Ferri, C.P., Fèvre, E.M., Finucane, M.M., Flaxman, S., Flood, L., Foreman, K., Forouzanfar, M.H., Fowkes, F.G.R., Franklin, R., Fransen, M., Freeman, M.K., Gabbe, B.J., Gabriel, S.E., Gakidou, E., Ganatra, H.A., Garcia, B., Gaspari, F., Gillum, R.F., Gmel, G., Gosselin, R., Grainger, R., Groeger, J., Guillemin, F., Gunnell, D., Gupta, R., Haagsma, J., Hagan, H., Halasa, Y.A., Hall, W., Haring, D., Haro, J.M., Harrison, J.E., Havmoeller, R., Hay, R.J., Higashi, H., Hill, C., Hoen, B., Hoffman, H., Hotez, P.J., Hoy, D., Huang, J.J., Ibeanusi, S.E., Jacobsen, K.H., James, S.L., Jarvis, D., Jasrasaria, R., Jayaraman, S., Johns, N., Jonas, J.B., Karthikeyan, G., Kassebaum, N., Kawakami, N., Keren, A., Khoo, J.-P., King, C.H., Knowlton, L.M., Kobusingye, O., Koranteng, A., Krishnamurthi, R., Lalloo, R., Laslett, L.L., Lathlean, T., Leasher, J.L., Lee, Y.Y., Leigh, J., Lim, S.S., Limb, E., Lin, J.K., Lipnick, M., Lipshultz, S.E., Liu, W., Loane, M., Ohno, S.L., Lyons, R., Ma, J., Mabweijano, J., MacIntyre, M.F., Malekzadeh, R., Mallinger, L., Manivannan, S., Marcenes, W., March, L., Margolis, D.J., Marks, G.B., Marks, R., Matsumori, A., Matzopoulos, R., Mayosi, B.M., McAnulty, J.H., McDermott, M.M., McGill, N., McGrath, J., Medina-Mora, M.E., Meltzer, M., Memish, Z.A., Mensah, G.A., Merriman, T.R., Meyer, A.-C., Miglioli, V., Miller, M., Miller, T.R., Mitchell, P.B., Mocumbi, A.O., Moffitt, T.E., Mokdad, A.A., Monasta, L., Montico, M., Moradi-Lakeh, M., Moran, A., Morawska, L., Mori, R., Murdoch, M.E., Mwaniki, M.K., Naidoo, K., Nair, M.N., Naldi, L., Narayan, K.V., Nelson, P.K., Nelson, R.G., Nevitt, M.C., Newton, C.R., Nolte, S., Norman, P., Norman, R., O'Donnell, M., O'Hanlon, S., Olives, C., Omer, S.B., Ortblad, K., Osborne, R., Ozgediz, D., Page, A., Pahari, B., Pandian, J.D., Rivero, A.P., Patten, S.B., Pearce, N., Padilla, R.P., Perez-Ruiz, F., Perico, N., Pesudovs, K., Phillips, D., Phillips, M.R., Pierce, K., Pion, S., Polanczyk, G.V., Polinder, S., Pope, C.A., Popova, S., Porrini, E., Pourmalek, F., Prince, M., Pullan, R.L., Ramaiah, K.D., Ranganathan, D., Razavi, H., Regan, M., Rehm, J.T., Rein, D.B., Remuzzi, G., Richardson, K., Rivara, F.P., Roberts, T., Robinson, C., De Leòn, F.R., Ronfani, L., Room, R., Rosenfeld, L.C., Rushton, L., Sacco, R.L., Saha, S., Sampson, U., Sanchez-Riera, L., Sanman, E., Schwebel, D.C., Scott, J.G., Segui-Gomez, M., Shahraz, S., Shepard, D.S., Shin, H., Shivakoti, R., Silberberg, D., Singh, D., Singh, G.M., Singh, J.A., Singleton, J., Sleet, D.A., Sliwa, K., Smith, E., Smith, J.L., Stapelberg, N.J., Steer, A., Steiner, T., Stolk, W.A., Stovner, L.J., Sudfeld, C., Syed, S., Tamburlini, G., Tavakkoli, M., Taylor, H.R., Taylor, J.A., Taylor, W.J., Thomas, B., Thomson, W.M., Thurston, G.D., Tleyjeh, I.M., Tonelli, M., Towbin, J.A., Truelsen, T., Tsilimbaris, M.K., Ubeda, C., Undurraga, E.A., van der Werf, M.J., van Os, J., Vavilala, M.S., Venketasubramanian, N., Wang, M., Wang, W., Watt, K., Weatherall, D.J., Weinstock, M.A., Weintraub, R., Weisskopf, M.G., Weissman, M.M., White, R.A., Whiteford, H., Wiersma, S.T., Wilkinson, J.D., Williams, H.C., Williams, S.R., Witt, E., Wolfe, F., Woolf, A.D., Wulf, S., Yeh, P.-H., Zaidi, A.K., Zheng, Z.-J., Zonies, D., Lopez, A.D., Murray, C.J., 2012. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the global burden of disease study 2010. The Lancet 380, 2163–2196. https://doi.org/10.1016/S0140-6736(12)61729-2.

Wang, K., Li, M., Bucan, M., 2007. Pathway-based Approaches for analysis of genomewide association studies. Am. J. Hum. Genet. 81, 1278–1283. https://doi.org/10.1086/522374.

Wang, K., Li, M., Hakonarson, H., 2010. Analysing biological pathways in genome-wide association studies. Nat. Rev. Genet. 11, 843–854. https://doi.org/10.1038/nrg2884.

Wittchen, H.-U., Fydrich, T., 1997. Strukturiertes Klinisches Interview Für DSM-IV (SKID-I Und SKID-II). Hogrefe, Göttingen, Germany.

Wu, M.C., Lee, S., Cai, T., Li, Y., Boehnke, M., Lin, X., 2011. Rare-variant association testing for sequencing data with the sequence kernel association test. Am. J. Hum. Genet. 89, 82–93. https://doi.org/10.1016/j.ajhg.2011.05.029.

You, Y., Li, W., Gong, Y., Yin, B., Qiang, B., Yuan, J., Peng, X., 2010. ShcD interacts with TrkB via its PTB and SH2 domains and regulates BDNF-induced MAPK activation. BMB Rep. 43, 485–490.