HHS Public Access

Author manuscript

Biol Psychiatry. Author manuscript; available in PMC 2020 July 15.

Published in final edited form as:

Biol Psychiatry. 2019 July 15; 86(2): 110–119. doi:10.1016/j.biopsych.2018.12.009.

Contribution of rare copy number variants to bipolar disorder risk is limited to schizoaffective cases

Alexander W. Charney[1,2,3,4], **Eli A. Stahl**[2], **Elaine K. Green**[5], **Chia-Yen Chen**[6,7], **Jennifer L. Moran**[8], **Kimberly Chambert**[6], **Jr. Richard A. Belliveau**[6], **Liz Forty**[9], **Katherine Gordon-Smith**[10], **Phil H. Lee**[6,11,12], **Evelyn J Bromet**[13], **Peter F Buckley**[14,15], **Michael A Escamilla**[16], **Ayman H. Fanous**[17,18], **Laura J Fochtmann**[13], **Douglas S. Lehrer**[19], **Dolores Malaspina**[1,2,3,20], **Stephen R. Marder**[21], **Christopher P. Morley**[22,23,24], **Humberto Nicolini**[25,26], **Diana O. Perkins**[27], **Jeffrey J. Rakofsky**[28], **Mark H. Rapaport**[28], **Helena Medeiros**[17], **Janet L. Sobell**[29], **Lena Backlund**[30], **Sarah E. Bergen**[31], **Anders Juréus**[31], **Martin Schalling**[32], **Paul Lichtenstein**[31], **James A. Knowles**[33], **Katherine E. Burdick**[1,11,34], **Ian Jones**[9], **Lisa A Jones**[10], **Christina M. Hultman**[1,31], **Roy Perlis**[35], **Shaun M. Purcell**[11,34], **Steven A. McCarroll**[6,36], **Carlos N. Pato**[17], **Michele T. Pato**[17], **Ariana Di Florio**[9,27], **Nick Craddock**[9], **Mikael Landén**[31,37], **Jordan W. Smoller**[6,7], **Douglas M. Ruderfer**[38], and **Pamela Sklar**[1,2,3,*]

[1]Department of Psychiatry, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, New York, NY, 10029, USA.

[2]Icahn Institute of Genomics and Multiscale Biology, Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, New York, NY, 10029, USA.

[3]Friedman Brain Institute, Department of Neuroscience, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, New York, NY, 10029, USA

[4]Department of Neurosurgery, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, New York, NY, 10029, USA.

[5]School of Biomedical and Health Sciences, Plymouth University Peninsula Schools of Medicine and Dentistry, Plymouth University, Portland Square, Plymouth, PL4 8AA, UK

[6]Stanley Center for Psychiatric Research, Broad Institute of Harvard and MIT, 75 Ames Street, Cambridge, MA, 02142, USA.

[7]Psychiatric and Neurodevelopmental Genetics Unit, Center for Genomic Medicine, Massachusetts General Hospital, 185 Cambridge St., Boston, MA, 02114, USA

Corresponding authors: Alexander W. Charney (alexander.charney@mssm.edu) and Douglas Ruderfer (douglas.ruderfer@vanderbilt.edu).

*Pamela Sklar, who led this work, passed away on November 20th, 2017. We dedicate this publication to her.

Conflicts of interest statement. The authors report no biomedical financial interests or potential conflicts of interest.

Disclosures. JWS is an unpaid member of the Bipolar/Depression Research Community Advisory Panel of 23andme.

[8]Department of Psychiatry, Massachusetts General Hospital, 185 Cambridge St, Boston, MA, 02114, USA.

[9]MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff Unviersity, Hadyn Ellis Building, Maindy Road, Cardiff, CF14 4HQ, UK.

[10]Department of Psychiatry, University of Birmingham, 25 Vincent Drive, Birmingham, B15 2FG, UK.

[11]Department of Psychiatry, Harvard Medical School, 77 Avenue Louis Pasteur, Boston MA 02115

[12]Center for Human Genetic Research, Massachusetts General Hospital, 185 Cambridge St, Boston, MA, 02114, USA.

[13]Department of Psychiatry, Stony Brook University, HSC, Level T-10, Room 020, Stony Brook, NY, 11794, USA.

[14]School of Medicine, Virginia Commonwealth University, 1201 E Marshall St., Richmond, VA, 23298, USA.

[15]Department of Psychiatry, Georgia Regents University Medical Center, 1120 15th Street, Augusta, GA, 30912, USA.

[16]Center of Excellence in Neuroscience, Department of Psychiatry, Texas Tech University Health Sciences Center at El Paso, 800 N. Mesa, Suite 200, El Paso, TX, 79902, USA.

[17]Department of Psychiatry and the Behavioral Sciences, State University of New York, Downstate Medical Center, 450 Clarkson Ave., Brooklyn, NY, 11203, USA.

[18]Department of Psychiatry, VA New York Harbor Healthcare System, 800 Poly Pl., Brooklyn, NY, 11209, USA.

[19]Department of Psychiatry, Wright State University, 3640 Colonel Gleen Hwy, Dayton, OH, 45435, USA.

[20]Department of Psychiatry, New York University, 550 First Ave., New York, NY, 10016, USA.

[21]Semel Institute for Neuroscience, University of California, Los Angeles, 11301 Wilshire Blvd, Los Angeles, CA, 90073, USA.

[22]Department of Psychiatry and Behavioral Science, State University of New York, Upstate Medical University, 750 East Adams St, MIMC 200, Syracuse, NY, 13210, USA.

[23]Department of Family Medicine, State University of New York, Upstate Medical University, 750 East Adams St, MIMC 200, Syracuse, NY, 13210, USA.

[24]Department of Public Health and Preventive Medicine, State University of New York, Upstate Medical University, 750 East Adams St, MIMC 200, Syracuse, NY, 13210, USA

[25]Center for Genomic Sciences, Universidad Autónoma de la Ciudad de México, Dr. García Diego # 168, Col. Doctores, Del., Mexico City, Mexico

[26]Department of Psychiatry, Carracci Medical Group, 107 Carracci Street, Mexico City, Mexico

[27]Department of Psychiatry, University of North Carolina at Chapel Hill, 321 S. Columbia Street, Chapel Hill, NC, 27516, USA.

[28]Department of Psychiatry and Behavioral Sciences, Emory University, 101 Woodruff Circle Suite 4000, Atlanta, GA, 30322, USA

[29]Department of Psychiatry and the Behavioral Sciences, University of Southern California, Keck School of Medicine, 2250 Alcazar Street, Los Angeles, CA, 90033, USA

[30]Department of Clinical Neuroscience, Karolinska Institutet, Vårdvägen 3, Stockholm, 11281, Sweden

[31]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Nobels väg 12 A, Stockholm, 17177, Sweden

[32]Department of Molecular Medicine and Surgery, Karolinska Institutet, Cmm (L8:00), Stockholm, 17176, Sweden

[33]Department of Cell Biology, State University of New York, Downstate Medical Center, 450 Clarkson Ave., Brooklyn, NY, 11203, USA.

[34]Department of Psychiatry, Brigham and Women's Hospital, 221 Longwood Ave Boston MA, 02115, USA.

[35]Center for Experimental Therapeutics, Massachusetts General Hospital, 185 Cambridge St, Boston, MA, 02114, USA

[36]Department of Genetics, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, MA, 02115, USA

[37]Institute of neuroscience and physiology, Sahlgenska academy at the Gothenburg university, Blå Sträket 15, Gothenburg, 41345, Sweden

[38]Division of Genetic Medicine, Departments of Medicine, Biomedical Informatics and Psychiatry, Vanderbilt Genetics Institute, Vanderbilt University Medical Center, Nashville, TN USA

Abstract

Background.—Genetic risk for bipolar disorder (BD) is conferred through many common alleles, while a role for rare copy number variants (CNVs) is less clear. BD subtypes schizoaffective disorder bipolar type (SAB), bipolar I disorder (BD I) and bipolar II disorder (BD II) differ according to the prominence and timing of psychosis, mania and depression. The factors contributing to the combination of symptoms within a given patient are poorly understood.

Methods.—Rare, large CNVs were analyzed in 6353 BD cases (3833 BD I [2676 with psychosis, 850 without psychosis], 1436 BD II, 579 SAB) and 8656 controls. Measures of CNV burden were integrated with polygenic risk scores (PRS) for schizophrenia (SCZ) to evaluate the relative contributions of rare and common variants to psychosis risk.

Results.—CNV burden did not differ relative to controls in BD when treated as a single diagnostic entity. Burden in SAB was increased relative to controls (p-value = 0.001), BD I (p-value = 0.0003) and BD II (p-value = 0.0007). Burden and SCZ PRS were increased in SAB compared to BD I with psychosis (CNV p-value = 0.0007, PRS p-value = 0.004) and BD I without

psychosis (CNV p-value = 0.0004, PRS p-value = 3.9×10^{-5}). Within BD I, psychosis was associated with increased SCZ PRS (p-value = 0.005) but not CNV burden.

Conclusions.—CNV burden in BD is limited to SAB. Rare and common genetic variants may contribute differently to risk for psychosis and perhaps other classes of psychiatric symptoms.

INTRODUCTION

Classically conceptualized as an episodic mood disorder with alternating periods of mania and depression, the diagnosis of bipolar disorder (BD) encompasses heterogeneous clinical presentations that vary with respect to symptomatology (1, 2), comorbidity (3) and longitudinal course (4). There are 3 diagnoses on the BD spectrum in current classifications of mental illness (5, 6): bipolar I disorder (BD I), bipolar II disorder (BD II) and schizoaffective disorder bipolar type (SAB). The criteria for these diagnoses differ from one another – and from clinically related diagnoses such as schizophrenia (SCZ) and major depressive disorder (MDD) – by nuances in the prominence and timing of manic, depressive and psychotic symptoms that are subject to change across versions of the same system of classification (5, 7, 8). The factors determining the combination of symptoms that occur in a given patient remain poorly understood.

BD genetic risk is characterized by many common single-nucleotide polymorphisms (SNPs) of small effect across the genome (9), many of which also are implicated in clinically related psychiatric conditions (10, 11). The overlap between BD and SCZ is particularly high in this regard, with genetic correlation estimates between the two ($r_g = 0.6 – 0.7$) comparable to estimates between BD I and BD II ($r_g = 0.7 – 0.8$) (9–12). In contrast, rare variants – in particular, rare copy number variants (CNVs) – have not been consistently implicated in risk for BD (13, 14, 23–26, 15–22), unlike in SCZ where an increased burden of rare CNVs is well-established and recurrent risk CNVs have been identified (20, 22, 27, 28). The largest genome-wide study of rare CNVs in BD to date found no differences in burden between approximately 2,600 cases and 8,800 controls(13). Smaller studies have been inconsistent (21, 22, 29). For instance, CNV burden in early-onset BD – a focus of BD CNV studies due to the increased rare CNV burden in neurodevelopmental disorders (25) – has been found by some (15, 16, 20, 26) but not others (17, 21–23). Specific CNVs implicated in SCZ and neurodevelopmental disorders have been tested for association with BD, and a duplication of 16p11.2 implicated in SCZ (30) was recently reported to be enriched in BD (13). Tested as a set rather than individually, these psychiatric CNVs are not significantly enriched in BD (21, 22, 26), nor have CNVs in BD consistently been found enriched for particular biological pathways or gene sets (15–17, 26). In total, the evidence that rare CNVs contribute to BD risk broadly is limited.

There is mounting evidence suggesting that the common alleles conferring risk to BD and SCZ act at the symptom level (31, 32), rooting the clinical similarity of BD and SCZ at least partially in common genetic variation. In contrast, the relative absence of rare CNV burden in BD (13) raises the possibility that this class of variation confers risk to clinical phenomena more commonly associated with SCZ. Such phenomena could include both the nuances in the prominence and timing of psychotic symptoms that formally differentiate

SCZ and BD diagnostic criteria (5, 6), as well as non-diagnostic features such as differences in cognitive deficits (33) and clinical course that historically formed the basis for the dichotomization of BD and SCZ (34, 35). Profiling rare CNVs and common risk alleles in BD cases stratified by granular clinical data would provide the opportunity to more directly test whether these classes of genetic variation make differential contributions to particular psychiatric traits. To our knowledge, such studies are lacking.

Here, we present results on a genome-wide study of rare CNV burden in 6,353 BD cases and 8,656 controls. In addition, we compare the relative contribution of rare CNVs and common SCZ risk alleles to risk of psychosis, a clinical phenomenon that differentiates BD subtypes from one another and from SCZ.

METHODS AND MATERIALS

Sample Description

The International Cohort Collection for Bipolar Disorder (ICCBD) includes BD cases and unaffected controls from the Sweden Bipolar Disorder Cohort (SWEBIC), the Bipolar Disorder Research Network (BDRN) in the United Kingdom, and the Genomic Psychiatry Consortium (GPC) from the University of Southern California. Full ICCBD sample descriptions have been previously reported in a genome-wide association study (GWAS) (12). The BDRN controls were collected as part of the Wellcome Trust Case Control Consortium; half were utilized in a genome-wide CNV burden analysis with a set of BD cases not in the current study (22), and the other half in a separate genome-wide CNV analysis (13). The subset of the SWEBIC cases and controls genotyped on the Affymetrix platform were in a previous report of genome-wide CNV burden in BD (20). Genome-wide CNV burden has not been reported before for the GPC cohort or for the SWEBIC cases and controls genotyped on the Illumina platform (45% of ICCBD cases in this study).

Phenotyping methods

SWEBIC clinical data was derived from 3 primary sources, which utilized a mixture of semi-structured interviews, retrospective chart review, and standardized rating scales. BDRN cases were assessed using Schedules for Clinical Assessment in Neuropsychiatry. GPC cases were assessed through a combination of focused, direct interviews and data extraction from medical records. On the basis of these data, best-estimate lifetime diagnoses were made according to DSM-IV criteria, and key clinical variables were rated. The inter-site reliability of diagnoses was assessed using Fleiss' Kappa statistic for multiple raters ($\kappa = 0.72$ for the primary diagnostic variable). Full descriptions of the approaches utilized in the phenotyping of the ICCBD cohorts have been reported previously (12, 37) (Supplementary Text). For some analyses in this report, clinical variables beyond case-control status were included from all 3 ICCBD sites, including age of onset, history of psychosis and family history. Age of onset was defined as the age at which first symptoms, impairment or diagnosis occurred. Psychosis was defined as the lifetime presence of hallucinations or delusions. Family history was defined as having any family member with any psychiatric diagnosis. For each variable, a set of standardized numerical values were derived, and site investigators harmonized

datasets according to these metrics. This was necessary to facilitate analysis across sites that used different phenotyping approaches.

Genotyping and ancestry covariates

Sample collection and genotyping procedures for the ICCBD have previously been reported (12). In brief, for all ICCBD sites DNA was extracted from peripheral blood samples that had been collected and stored at −20°C. Samples were then genotyped at the Broad Institute, and genotypes were called using either Birdsuite (Affymetrix) or BeadStudio (Illumina). Ancestry covariates were derived from the genotyping data through multidimensional scaling (MDS) analysis on genome-wide identity-by-descent distances calculated for all pairs of individuals. Quality control procedures implemented to derive the genotype calls utilized are detailed in an earlier genome-wide association study of this cohort (12).

CNV calling and quality control

Rare CNVs were identified using the Birdseye program in Birdsuite (38). Only subjects who passed quality control filters in an earlier GWAS of the same individuals(12) were considered for CNV analyses. CNVs were excluded if any of the following criteria were met: logarithm of the odds ratio score < 10, number of probes < 10, probe density of < 1 per 20 kilobases (KB), frequency in ICCBD > 1%, or location within a region known to contain common CNVs or large genomic gaps (e.g., centromeres). If in a given individual the distance between two CNVs was less than 20% of their combined size, they were considered artificially split by the calling algorithm and combined into a single event. For the BDRN cohort, only genomic regions covered in both cases and controls were retained in order to reduce batch effects resulting from cases and controls being genotyped on different Illumina arrays (Supplementary Text; Supplementary Figure 1). Subjects were removed for having total CNV number greater than two standard deviations different from the mean number of CNVs in the cohort (prior to applying filters for CNV frequency). Unless otherwise specified, burden analyses were restricted to autosomal CNVs > 100KB. Two events were considered equivalent for the purposes of defining frequency if one overlapped the other by at least 50%. Quality control checks were performed separately for the SWEBIC Affymetrix, SWEBIC Illumina, BDRN, and GPC cohorts (Table 1). In the context of burden analyses, we use the term “CNV” to refer to the combined set of deletions and duplications, and “singleton CNVs” were defined as any event that occurred once in the full ICCBD case-control cohort without consideration of whether the event was a deletion or a duplication. Singleton deletions and duplications were defined after first filtering the dataset for that type of event. As such, not all singleton deletions and duplications are in the singleton CNV group.

CNV burden tests

For our primary CNV burden tests, we defined CNV burden in 3 ways: the number of CNVs occurring per individual (the CNV number); the number of genes lying within CNVs per individual (the CNV gene count); the total distance covered by CNVs. We elected to focus on these 3 classes of burden because there is no clear class of burden most relevant to BD and these classes significantly differ between SCZ cases and controls (27). We stratified CNVs by 3 types: deletions only, duplications only, deletions and duplications (or “CNVs”);

by 2 sizes: over 100KB and over 500KB; and by 2 frequencies: singletons (a frequency of 6.7×10^{-5}) and those occurring in less than 1% in the ICCBD. This led to 36 tests between each of 7 pairs of phenotypes we compared: (1) BD cases to controls, (2) BD I cases to controls, (3) BD II cases to controls, (4) SAB cases to controls, (5) BD I cases to BD II cases, (6) BD I cases to SAB cases, and (7) BD II cases to SAB cases. Thus, a total of 252 tests comprised our primary assessment of CNV burden.

Previous studies of CNV burden in BD have reported significant results for tests where the definition of burden fell outside the scope of these 252 tests. Manual curation of the literature identified 34 nominal associations at a p-value of less than 0.05 in the original report. We were able to follow-up 27 of these in the ICCBD (for the other 7, the original study included either SCZ cases or BD parent-child trios), of which 21 were not in our primary 252 tests. For these tests, we excluded ICCBD samples overlapping those in the original report.

We also tested ICCBD CNVs (size > 100KB, frequency < 1 %) for enrichment of 3 sets of CNVs previously identified in studies of BD, SCZ or neurodevelopmental disorders. The BD CNV set (16 deletions, 14 duplications) was comprised of autosomal *de novo* CNVs reported in 3 previous studies of BD trios (16, 17, 24). The SCZ CNV set (11 deletions, 8 duplications) was comprised of autosomal CNVs with suggestive evidence for association in a meta-analysis of over 20,000 SCZ cases and 20,000 controls (27). The neurodevelopmental CNV set (27 deletions, 18 duplications) was from a list curated for a previous report (17) after removing those overlapping the SCZ set. In order for a CNV in the test set to be considered overlapping with an ICCBD CNV, the ICCBD CNV was required to cover at least 50% of the test CNV and be of the same CNV type (i.e., deletion or duplication).

All tests were performed using permutation in PLINK (39) controlling for genotyping platform and ICCBD site. Significance was evaluated using 10,000 permutations. The 252 tests in the primary assessment were 2-sided with the exception of 6 tests that had previously been reported as significant. A one-sided test in the direction of the association reported in the original paper was used for these 6 tests as well as for the additional 21 tests following up previous associations and the 3 tests of CNV sets.

Multiple test correction for CNV burden tests

In the genome-wide CNV burden analyses described above there are a total 276 tests (252 in our primary assessment of CNV burden, 21 tests of previous associations and 3 tests of CNV sets). The empirical tests performed in PLINK as described above were controlled for multiple testing using the false discovery rate (FDR) estimation method of Benjamini-Hochberg (40) implemented in R using the p.adjust() function. Using an FDR of 5%, tests with empirical p-values below 0.002 were considered study-wide significant.

Contribution of CNV burden and SCZ PRS to psychosis

Following results from our primary burden analyses, we analyzed CNV burden and loading of common SCZ risk alleles in BD I and SAB cases. BD II was excluded from these analyses to remove effects resulting from known differences in polygenic loading of SCZ

alleles across BD subtypes (12). For these analyses, burden was defined as the number of CNVs greater than 500KB and present in less than 1 % of the study sample. We focused on this particular burden class because it was the only class in our primary 252 tests where an increase was seen in SAB compared to controls, BD I and BD II (see Results). For these analyses, burden was tested using logistic regression, which returned similar results to permutation but allowed us to include in the model continuously-distributed ancestry covariates and facilitated the calculation of odds-ratios (ORs) for CNV burden (27). In the regression model, we used phenotype status as the dependent variable and CNV burden as an independent predictor variable. The OR was calculated as the exponential of the logistic regression coefficient, and OR > 1 represents increased risk for the "affected" phenotype in the model, which was designated to be the phenotype more clinically similar to SCZ. Using a similar regression model, we carried out polygenic scoring analyses (41). Quantitative polygenic risk scores (PRS) were computed for each case subject based on the set of SNPs with p-values less than 0.5 in the second SCZ GWAS from the Psychiatric Genomics Consortium (PGC) (42). PRS analyses excluded ICCBD samples present in the PGC studies. We calculated the proportion of variance explained (Nagelkerke's R^2) by SCZ PRS by subtracting the Nagelkerke's R^2 attributable to covariates alone from the Nagelkerke's R^2 for PRS plus covariates. Effect sizes for both CNV burden and SCZ PRS were calculated as a t-statistic that is the ratio of the coefficient of the burden or PRS variable and its standard error from a generalized linear regression model equation. As studies of SCZ have consistently demonstrated higher CNV burden in cases compared to controls (27, 28), cases were stratified by clinical dimensions related to SCZ (i.e., psychosis) and 1-sided statistical tests were used evaluating for higher rates in groups with the more SCZ-like phenotype.

Power calculations

We calculated power for tests of CNV burden in BD compared to controls, as well as between subtypes of BD stratified by psychosis. Specifically, calculations were performed for the 3 primary classes of burden assessed in BD compared to controls, and the 1 class of burden assessed in our analyses of psychosis. Effect sizes ranging from 1 to 2.5 (by increments of 0.01) were utilized in the power calculations. To account for the possibility of allele frequency differences between cohorts, the effect size in the power calculation was divided by the standard error from the burden test.

RESULTS

CNV burden in BD

We assessed genome-wide differences in rare CNV burden between 6,353 BD cases and 8,656 controls (Table 1). After initial filters for size (> 100KB) and frequency (occurring in < 1% of ICCBD), we observed 10,515 CNVs (3,970 deletions and 6,545 duplications). No difference in the CNV number was found between cases and controls (case rate = 0.698, control rate = 0.702, p-value = 0.86). This was true both for deletions (case rate = 0.266, control rate = 0.264, p-value = 0.78) and duplications (case rate = 0.433, control rate = 0.439, p-value = 0.72). Similarly, no differences were observed between cases and controls with respect to the number of genes hit or the total distance covered by CNVs (Table 2; Supplementary Table 1). We calculated power to detect differences in these 3 burden classes

across a range of effect sizes (Methods). Assuming effect sizes reported for SCZ (27), power of 100% was attained to identify differences between BD and controls for the number of genes hit and total distance covered by CNVs per individual, while 33% power was attained for the number of CNVs per individual. Following previous literature showing that rarer and larger CNVs carry increased burden for neuropsychiatric illness (28), we further filtered CNVs by size (> 500KB) and frequency (those that occur once in the 15,009 ICCBD individuals, a frequency of 6.7×10^{-5}). No burden in these classes was observed below our study-wide p-value threshold (Table 2; Supplementary Table 1). Similarly, manual curation of the literature identified 21 additional associations of BD and CNV burden (p-value < 0.05 in the initial report) that we followed-up here (Methods), none of which withstood correction for multiple tests (Figure 1; Table 3). Sets of CNVs previously implicated in neuropsychiatric disorders (Methods) were also not enriched for deletions or duplications in BD compared to controls.

BD is a heterogeneous disorder clinically, and a previous report of common variation in this cohort found evidence for genetic heterogeneity between clinical subtypes of BD (12). This information, combined with CNV burden being a well-established component of SCZ genetic architecture (27), led us to hypothesize that increased CNV burden may be present in the BD subtypes most clinically similar to SCZ. To test this hypothesis, we first sought to determine if CNV burden differed between BD subtypes (BD I n = 3,833, BD II n = 1,436, SAB n = 579) and controls (n = 6,383), as well as between BD subtypes and one another. Increased burden was seen in SAB compared to controls in all 3 of the primary burden classes evaluated, as well as compared to both BD I and BD II (Table 2). For one burden class, number of CNVs with size over 500KB and frequency < 1%, SAB had higher burden compared to controls (p-value = 0.001), BD I (p-value = 3×10^{-4}; Figure 2a) and BD II (p-value = 7×10^{-4}). We therefore elected to focus downstream CNV analyses on this class of burden.

Contribution of CNV burden and SCZ PRS to psychosis in BD

SCZ is the archetypal psychotic illness in current psychiatric classification systems (5) and increased CNV burden is a well-established component of its genetic architecture (27, 28). Psychosis is also a prominent component of BD, and the diagnostic criteria differentiating BD subtypes (e.g., BD I, SAB) from one another and from SCZ relate to the co-occurrence of psychosis with mania (5, 6). The observed CNV burden in SAB, a diagnosis that requires most of the criteria of SCZ be met, being absent in BD as a whole prompted inquiry into whether CNV burden contributes to psychosis or to non-diagnostic clinical phenomena that differentiate SAB from other BD subtypes, and whether the same pattern is seen for common SCZ risk alleles. We stratified the ICCBD cases by the prominence of psychotic symptoms, correlating psychosis risk with both the CNV burden and SCZ PRS (12, 32). Cases were stratified into SAB (n = 579), BD I with psychosis (n = 2,676) and BD I without psychosis (n = 850). CNV burden was increased in SAB compared to BD I with and without psychosis (SAB rate = 0.116; BD I with psychosis rate = 0.069, p-value = 7.21×10^{-4}; BD I without psychosis rate = 0.067, p-value = 4.42×10^{-4}), but no difference was observed between BD I with and without psychosis (p-value = 0.88; Figure 2b; Supplementary Figure 2). SCZ PRS were higher in SAB compared to BD I with psychosis (Nagelkerke's R^2 =

0.004; p-value = 0.004) and in BD I with psychosis compared to BD I without psychosis (Nagelkerke's R^2 = 0.003; p-value = 0.005; Figure 2b). We calculated the power to detect differences in CNV burden between these cohorts across a range of effect sizes (Supplementary Figure 3). At the effect size observed in the comparison of SAB to controls (OR = 1.58), a nominally significant difference could be detected with 81% power between SAB and BD I with psychosis, 53% power between SAB and BD I without psychosis, and 84% between BD I with and without psychosis.

DISCUSSION

We observed no differences in the genome-wide burden of rare, large CNVs in a cohort of 6,353 BD cases and 8,656 controls. Furthermore, we did not find strong support for any previously reported BD CNV burden associations despite reproducing original analyses with respect to phenotypes compared and the cutoffs for CNV size and frequency used in quality control procedures. Taken together, the case-control analyses presented here confirm in a well-powered cohort that rare CNV burden is not a feature of BD when treated as a single diagnostic entity.

Individuals with a diagnosis of BD comprise a clinically heterogeneous group, and the lack of CNV burden when BD is treated as a single diagnostic entity does not preclude a role of CNV burden in the pathogenesis of subsets of cases. Specifically, we hypothesized this may the case for individuals who present with psychotic symptoms in the absence of a major mood episode, given the known CNV burden in SCZ (27, 28) and the clinical overlap between SCZ and BD. Indeed, we found that cases with SAB, who by definition experience psychosis both in the presence and absence of mania, have higher rates of large, rare CNVs compared to controls and other BD subtypes. The class of burden with the strongest signal genome-wide in SCZ compared to controls is the number of genes hit by deletions per individual (27). We observed this to also be the case in SAB compared to controls (Table 2).

The diagnostic criteria differentiating BD I with psychosis, SAB and SCZ from one another relate to the prominence and timing of psychotic symptoms. Through deeper analyses comparing SAB and BD I, however, we found that CNV burden was unrelated to the presence of psychosis. This was in contrast to SCZ PRS, which were increased in the phenotypes characterized by more prominent psychosis. Taken together, these results suggest that common variants may contribute to psychotic symptoms whereas rare CNVs may contribute to dimensions of illness that differentiate psychotic illnesses from one another. One possibility in this regard is that CNVs may influence risk for cognitive deficits, which are more prominent in SCZ compared to BD and affect cognition in the general population (33). Another possibility is that CNV burden increases risk for spontaneous psychosis (i.e., the psychoses of SCZ and SAB) but not psychosis secondary to severe mental stress, which some argue is the mechanism underlying psychosis during mania. Alternatively, it is possible that compared to the persistent psychosis seen in SAB the psychosis of mania and/or depression are rated less reliably. Future studies with deeper phenotyping should aim to test these and other hypotheses.

This study has important limitations. Diagnostic misclassification of SCZ cases with SAB is possible, and while unlikely could account for the observed PRS and CNV results. For some of these analyses, sample size is an important consideration, and we emphasize that these findings must be followed up in larger cohorts. If replicated, they would provide support for the notion that different classes of genetic variants contribute to different classes of symptomatology in mood and psychotic syndromes. It might then be fair to inquire whether the higher CNV burden in SCZ compared to BD may be evidence not that they comprise two biologically distinct disease entities, but rather that clinicians are more likely to diagnose SCZ when a particular clinical phenomenon is present (e.g., cognitive deficits, spontaneous psychosis). These unresolved questions highlight the need for a multiscale approach to the study of mental illness, whereby integrating high-dimensional molecular and clinical data from each patient at the scale that GWAS has shown can be achieved may facilitate the development of a data-driven taxonomy.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments.

We are grateful for the participation of all subjects contributing to this research, and to the collection team that worked to recruit them. We acknowledge funding support from by National Institutes of Health (NIH)/National Institute of Mental Health (NIMH) grant R01MH085542 (AWC, JWS and PS), NIH/NIMH grant R01 MH106547 (JWS, PS, SAM), NIH/NIMH grant R01MH085548 (AHF, CNP, CPM, DM, DOP, DSL, ELB, HM, HN, JAK, JJR, JLS, LJF, MAE, MHR, MTP, PFB, SRM), NIH/NIMH grant K99MH101367 (PHL), the Stanley Medical Research Institute (JLM, KC, RAB, SAM), philanthropic gifts from Kent and Elizabeth Dauten and Ted and Vada Stanley (JLM, KC, RAB, SAM), the Swedish Research Council 2013–3196 (CMH), the Swedish Medical Research Council grants K2014–62X-14647–12-51 and K2010–61P-21568–01-4 (ML), the Swedish foundation for Strategic Research grant KF10–0039 (ML), the Swedish Federal Government under the LUA/ALF agreement grants ALF 20130032 and ALFGbG-142041 (ML), European Commission-Marie Curie Fellowship (AD), Wellcome Trust (IJ, KG, LAJ, LF, NC). JwS is an MGH Tepper Family Research Scholar. Work at the Icahn School of Medicine at Mount Sinai was also supported by the Institute for Genomics and Multiscale Biology and the the computational resources and staff expertise provided by Scientific Computing at the Icahn School of Medicine at Mount Sinai. The funders had no role in study design, execution, analysis or manuscript preparation.

REFERENCES

1. Dunayevich E, Keck PE (2000): Prevalence and description of psychotic features in bipolar mania. Curr Psychiatry Rep. 2: 286–290. [PubMed: 11122970]
2. Burdick KE, Ketter TA, Goldberg JF, Calabrese JR (2015): Assessing cognitive function in bipolar disorder: challenges and recommendations for clinical trial design. J Clin Psychiatry. 76: e342–50. [PubMed: 25830456]
3. Merikangas KR, Akiskal HS, Angst J, Greenberg PE, Hirschfeld RMA, Petukhova M, Kessler RC (2007): Lifetime and 12-Month Prevalence of Bipolar Spectrum Disorder in the National Comorbidity Survey Replication. Arch Gen Psychiatry. 64: 543–552. [PubMed: 17485606]
4. Perlis RH, Ostacher MJ, Patel JK, Marangell LB, Zhang H, Wisniewski SR, et al. (2006): Predictors of Recurrence in Bipolar Disorder: Primary Outcomes From the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD). Am J Psychiatry. 163: 217–224. [PubMed: 16449474]
5. American Psychiatric Association. (2013): Diagnostic and statistical manual of mental disorders, Fifth Edition American Psychiatric Pub. doi: 10.1176/appi.books.9780890425596.744053.
6. World Health Organization (1992): The ICD-10 Classification of Mental and Behavioural Disorders. Int Classif. 10: 1–267.

7. American Psychiatric Association (1980): Diagnostic and Statistical Manual of Mental Disorders, Third Edition. .
8. American Psychiatric Association. (2000): Diagnostic and statistical manual of mental disorders, Fourth Edition, Text Revision. Washington DC: Fourth edition Washington, DC : American Psychiatric Association, [1994] ©1994 Retrieved from https://search.library.wisc.edu/catalog/999733358502121.
9. Stahl E, Forstner A, McQuillin A, Ripke S, Ophoff R, Scott L, et al. (2017): Genomewide association study identifies 30 loci associated with bipolar disorder. bioRxiv. 173062.
10. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh P-RR, et al. (2015): An atlas of genetic correlations across human diseases and traits. Nat Genet. 47: 1236–1241. [PubMed: 26414676]
11. Lee SH, Ripke S, Neale BM, Faraone SV, Purcell SM, Perlis RH, et al. (2013): Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. Nat Genet. 45: 984–994. [PubMed: 23933821]
12. Charney AW, Ruderfer DM, Stahl EA, Moran JL, Chambert K, Belliveau RA, et al. (2016): Evidence for genetic heterogeneity between clinical subtypes of bipolar disorder. Transl Psychiatry. .
13. Green EK, Rees E, Walters JTR, Smith K-G, Forty L, Grozeva D, et al. (2016): Copy number variation in bipolar disorder. Mol Psychiatry. 21: 89–93. [PubMed: 25560756]
14. Yang S, Wang K, Gregory B, Berrettini W, Wang LS, Hakonarson H, Bucan M (2009): Genomic landscape of a three-generation pedigree segregating affective disorder. PLoS One. 4: 1–10.
15. Zhang D, Cheng L, Qian Y, Alliey-Rodriguez N, Kelsoe JR, Greenwood T, et al. (2009): Singleton deletions throughout the genome increase risk of bipolar disorder. Mol Psychiatry. 14: 376–380. [PubMed: 19114987]
16. Malhotra D, McCarthy S, Michaelson JJ, Vacic V, Burdick KE, Yoon S, et al. (2011): High frequencies of de novo CNVs in bipolar disorder and schizophrenia. Neuron. 72: 951–963. [PubMed: 22196331]
17. Georgieva L, Rees E, Moran JL, Chambert KD, Milanova V, Craddock N, et al. (2014): De novo CNVs in bipolar affective disorder and schizophrenia. Hum Mol Genet. 23: 6677–6683. [PubMed: 25055870]
18. Chen Y-H, Ru-B Lu, Hung H, Kuo P-H, Lu R-B (2014): Identifying Potential Regions of Copy Number Variation for Bipolar Disorder. Microarrays. 3: 52–71. [PubMed: 27605030]
19. Chen J, Calhoun VD, Perrone-Bizzozero NI, Pearlson GD, Sui J, Du Y, Liu J (2016): A pilot study on commonality and specificity of copy number variants in schizophrenia and bipolar disorder. Transl Psychiatry. 6: e824. [PubMed: 27244233]
20. Bergen SE, O'Dushlaine CT, Ripke S, Lee PH, Ruderfer DM, Akterin S, et al. (2012): Genome-wide association study in a Swedish population yields support for greater CNV and MHC involvement in schizophrenia compared with bipolar disorder. Mol Psychiatry. 17: 880. [PubMed: 22688191]
21. McQuillin A, Bass N, Anjorin A, Lawrence J, Kandaswamy R, Lydall G, et al. (2011): Analysis of genetic deletions and duplications in the University College London Bipolar Disorder case control sample. Eur J Hum Genet. 19: 588–592. [PubMed: 21206513]
22. Grozeva D, Kirov G, Ivanov D, Jones IR, Jones L, Green EK, et al. (2010): Rare copy number variants: a point of rarity in genetic risk for bipolar disorder and schizophrenia. Arch Gen Psychiatry. 67: 318–327. [PubMed: 20368508]
23. Grozeva D, Kirov G, Conrad DF, Barnes CP, Hurles M, Owen MJ, et al. (2013): Reduced burden of very large and rare CNVs in bipolar affective disorder. Bipolar Disord. 15: 893–898. [PubMed: 24127788]
24. Noor A, Lionel AC, Cohen-Woods S, Moghimi N, Rucker J, Fennell A, et al. (2014): Copy number variant study of bipolar disorder in Canadian and UK populations implicates synaptic genes. Am J Med Genet Part B Neuropsychiatr Genet. 165: 303–313.
25. Malhotra D, Sebat J (2012): CNVs: harbingers of a rare variant revolution in psychiatric genetics. Cell, 2012/3/20 148: 1223–1241. [PubMed: 22424231]

26. Priebe L, Degenhardt F a, Herms S, Haenisch B, Mattheisen M, Nieratschker V, et al. (2012): Genome-wide survey implicates the influence of copy number variants (CNVs) in the development of early-onset bipolar disorder. Mol Psychiatry. 17: 421–432. [PubMed: 21358712]
27. Marshall CR, Howrigan DP, Merico D, Thiruvahindrapuram B, Wu W, Greer DS, et al. (2016): Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. Nat Genet. 49: 27–35. [PubMed: 27869829]
28. International Schizophrenia Consortium (2008): Rare chromosomal deletions and duplications increase risk of schizophrenia. Nature. 455: 237–241. [PubMed: 18668038]
29. Green EK, Hamshere M, Forty L, Gordon-Smith K, Fraser C, Russell E, et al. (2013): Replication of bipolar disorder susceptibility alleles and identification of two novel genome-wide significant associations in a new bipolar disorder case-control sample. Mol Psychiatry. 18: 1302–1307. [PubMed: 23070075]
30. McCarthy SE, Makarov V, Kirov G, Addington AM, McClellan J, Yoon S, et al. (2009): Microduplications of 16p11.2 are associated with schizophrenia. Nat Genet. 41: 1223–7. [PubMed: 19855392]
31. Ruderfer DM, Fanous AH, Ripke S, McQuillin A, Amdur RL, Schizophrenia Working Group of Psychiatric Genomics C, et al. (2014): Polygenic dissection of diagnosis and clinical dimensions of bipolar disorder and schizophrenia. Mol Psychiatry, 2013/11/28 19: 1017–1024. [PubMed: 24280982]
32. Bipolar Disorder and Schizophrenia Working Group of the Psychiatric Genomics Consortium (2018): Genomic Dissection of Bipolar Disorder and Schizophrenia, Including 28 Subphenotypes. Cell. 173: 1705–1715.e16. [PubMed: 29906448]
33. Stefansson H, Meyer-Lindenberg A, Steinberg S, Magnusdottir B, Morgen K, Arnarsdottir S, et al. (2013): CNVs conferring risk of autism or schizophrenia affect cognition in controls. Nature. 361–366.
34. Kraepelin E (1921): Manic-depressive insanity and paranoia. E. & S. Livingstone Retrieved February 13, 2016, from https://books.google.com/books/about/Manic_depressive_insanity_and_paranoia.html?id=DuVMBi9ycfEC&pgis=1.
35. Kraepelin E, Barclay R (translator) (1919): Dementia Praecox and the Paraphrenia. (Robertson G, editor). Edinburgh: E & S Livingstone.
36. Kendall KM, Rees E, Escott-Price V, Einon M, Thomas R, Hewitt J, et al. (2016): Cognitive Performance Among Carriers of Pathogenic Copy Number Variants: Analysis of 152,000 UK Biobank Subjects. Biol Psychiatry. 103–110. [PubMed: 27773354]
37. Pato MT, Sobell JL, Medeiros H, Abbott C, Sklar BM, Buckley PF, et al. (2013): The genomic psychiatry cohort: partners in discovery. Am J Med Genet B Neuropsychiatr Genet. 162: 306–12.
38. Korn JM, Kuruvilla FG, McCarroll SA, Wysoker A, Nemesh J, Cawley S, et al. (2008): Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. Nat Genet. 40: 1253. [PubMed: 18776909]
39. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. (2007): PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 81: 559. [PubMed: 17701901]
40. Benjamini Y, Hochberg Y (1995): Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B. 57. doi: 10.2307/2346101.
41. International Schizophrenia Consortium (2009): Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature. 460: 748. [PubMed: 19571811]
42. Schizophrenia Psychiatric Genome-Wide Association Study Consortium (2011): Genome-wide association study identifies five new schizophrenia loci. Nat Genet. 43: 969–976. [PubMed: 21926974]
43. Burdick KE, Russo M, Frangou S, Mahon K, Braga RJ, Shanahan M, Malhotra AK (2014): Empirical evidence for discrete neurocognitive subgroups in bipolar disorder: clinical implications. Psychol Med. 44: 3083–3096. [PubMed: 25065409]
44. Van Rheenen TE, Lewandowski KE, Tan EJ, Ospina LH, Ongur D, Neill E, et al. (2017): Characterizing cognitive heterogeneity on the schizophrenia-bipolar disorder spectrum. Psychol Med. 47: 1848–1864. [PubMed: 28241891]

45. Russo M, Van Rheenen TE, Shanahan M, Mahon K, Perez-Rodriguez MM, Cuesta-Diaz A, et al. (2017): Neurocognitive subtypes in patients with bipolar disorder and their unaffected siblings. Psychol Med. 47: 2892–2905. [PubMed: 28587689]

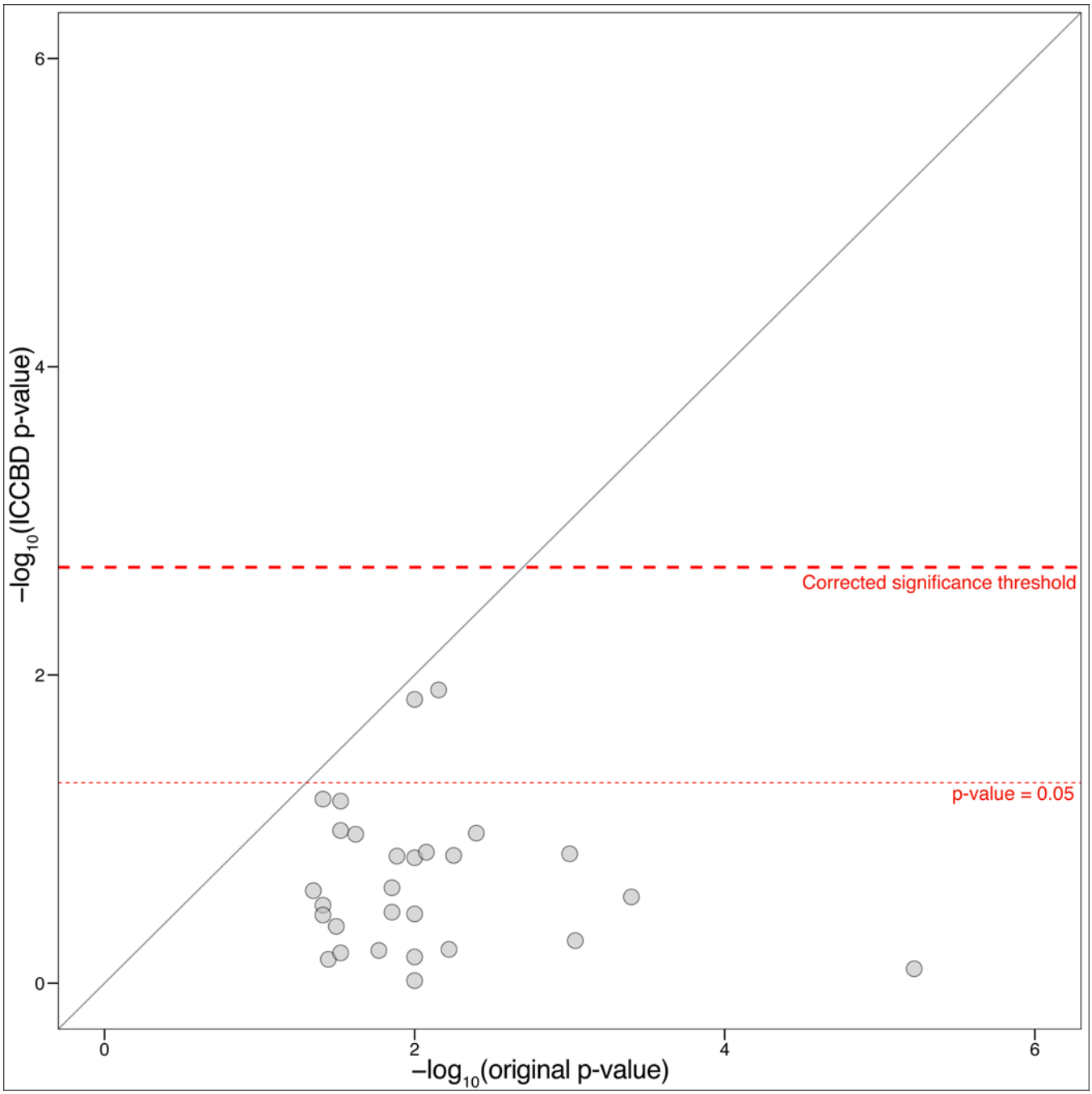

Figure 1.

Replication of previous reports of CNV burden in BD. Curation of literature on CNV burden in BD identified 36 instances where nominal association (p-value < 0.05) was reported. We were able to test 28 of these in the ICCBD. Plotted here are p-values in previous reports (x-axis) compared to the same test performed in ICCBD cohort (y-axis). There were 4 tests for which nominal significance was observed in the ICCBD data: (1) singleton deletions greater than 100KB in cases compared to controls, (2) proportion of individuals with a singleton deletion greater than 100KB in cases compared to controls, (3) singleton deletions greater

than 100KB in early onset cases compared to controls, and (4) proportion of individuals with a singleton deletion greater than 100KB in early onset cases compared to controls. None of these observations surpassed multiple test correction for the 27 tests we followed up in our data.

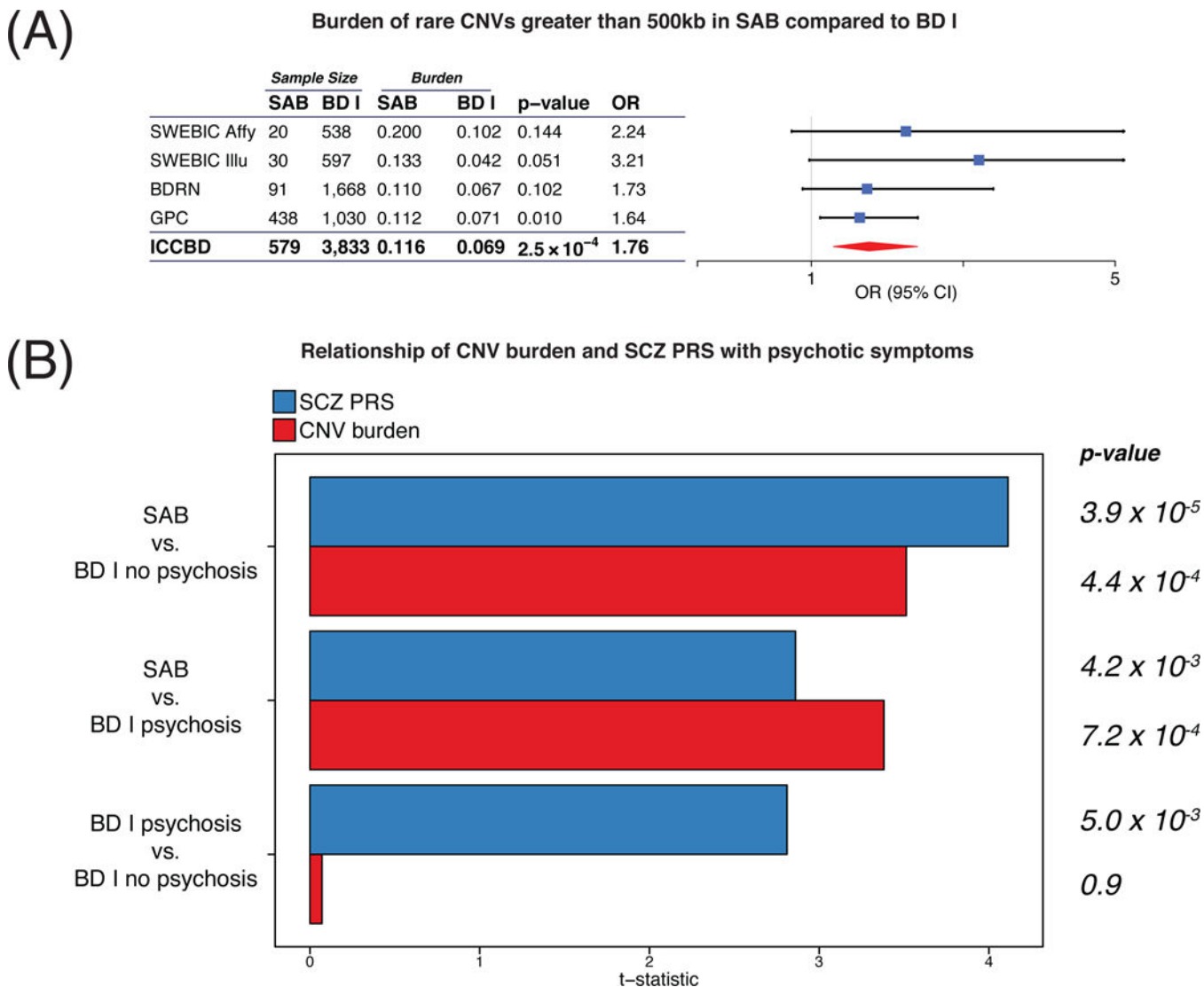

Figure 2.

Burden of rare CNVs (frequency < 1 %) greater than 500KB in SAB compared to BD I. (a) Forest plot of CNV burden partitioned by site of collection, with the full ICCBD sample at the bottom. CNV burden is calculated by combining CNV deletions and duplications. The p-values presented here for burden tests used a logistic regression model predicting SAB-BD I status by CNV burden along with covariates. The odds ratio (OR) is the exponential of the logistic regression coefficient, and OR > 1 predicts increased SAB risk. (b) Comparison of BD and SAB to one another with respect to polygenic risk scores and CNV burden. Regression analyses were performed of phenotype (stratified by history of psychosis) on polygenic scores derived from a previous GWAS for SCZ (blue) and burden of CNVs with frequency less than 1 % and size greater than 500KB (red). MDS components, study site and gender were used as covariates. The t-statistic plotted on the x-axis is the ratio of the coefficient of the polygenic score or CNV burden variable and its standard error from the generalized linear model regression equation. The direction of the plotted bars indicates higher CNV burden or PRS in the phenotype listed first in the y-axis label. The p-values for

whether polygenic risk scores or CNV burden differed significantly between phenotypes are shown at the far right. The Nagelkerke's R^2 for the corresponding polygenic risk score comparisons were as 0.004 for SAB vs. BD I with psychosis, 0.011 for SAB vs. BD I without psychosis and 0.003 for BD I with psychosis vs. BD I without psychosis.

Table 1:

Sample characteristics for ICCBD CNV analyses.

Cohort	Ethnicity	Chip	Batch[a]	Cases	Controls	BD I			BD II	SAB	NOS
						All	Psychosis	No psychosis			
SWEBIC	European	Affymetrix 6.0	Wave3	0	499	0	0	0	0	0	0
		Affymetrix 6.0	Wave4	917	1170	538	355	162	114	20	245
		Omni Express	Wave6	1,356	1140	597	355	199	500	30	229
				0							
BDRN	European	Combo Chip	-	1,084	0	677	389	185	361	34	12
		Omni Express	-	1,493	0	991	694	161	426	57	19
		Illumina 1.2M	58BC	0	2428	0	0	0	0	0	0
		Illumina 1.2M	NBS	0	2149	0	0	0	0	0	0
				0							
GPC	European-American	Omni Express	-	1,503	1,270	1,030	883	143	35	438	0
ICCBD				**6,353**	**8,656**	**3,833**	**2,676**	**850**	**1,436**	**579**	**505**

[a]CNV analyses including Swedish samples from waves 2–4 have been previously reported in Bergen et al. CNV results for SWEBIC controls from all waves have been reported in Marshall et al. CNV analyses for BD using subsets of the BDRN controls have been reported in Grozeva et al 2010 and Green et al 2016. BDRN cases were included in analyses by Green et al 2016.

Biol Psychiatry. Author manuscript; available in PMC 2020 July 15.

Table 2:

CNV burden in bipolar disorder. Burden metrics are presented for the following groups (N): BD treated as a single diagnostic entity (6353), controls (8656), BD I (3833), BD II (1436) and SAB (579). Three classes of CNV burden were assessed, indicated by the description above the dotted lines. Burden was compared between cases and controls, as well as between case subtypes. P-values are two-sided, uncorrected for multiple testing, and based on 10,000 permutations testing for relative burden between the two groups. CNV type, size, and frequency refer to the filters applied for the test being reported. Events refers to the number of CNVs total observed in the groups being compared for the specified parameters. P-values surpassing study-wide significance are shown in bold. Singletons are those CNVs that occur once within the full ICCBD case cohort when filtered for those greater than 100KB in size. KB=kilobases

				Burden					p-value						
CNV type	**Size (KB)**	**Frequency**	**Events**	**BD**	**CON**	**BD I**	**BD II**	**SAB**	**BD CON**	**BD I CON**	**BD II CON**	**SAB CON**	**BD I SAB**	**BD II SAB**	**BD I BD II**
Number of CNVs per individual															
All	100	<1%	10515	0.698	0.702	0.682	0.683	0.800	0.855	0.345	0.753	0.036	0.016	0.025	1.000
		Singleton	1849	0.125	0.122	0.129	0.114	0.131	0.532	0.274	0.588	0.318	0.562	0.274	0.276
	500	<1%	1014	0.070	0.066	0.069	0.054	0.111	0.353	0.545	0.423	**0.001**	**3×10^{-4}**	**7×10^{-4}**	0.145
		Singleton	270	0.020	0.017	0.019	0.015	0.035	0.243	0.556	0.791	0.009	0.006	0.037	0.481
Deletions	100	<1%	3970	0.266	0.264	0.251	0.280	0.323	0.778	0.193	0.418	0.099	0.010	0.146	0.055
		Singleton	918	0.067	0.057	0.067	0.065	0.070	0.021	0.057	0.348	0.141	0.282	0.686	0.884
	500	<1%	231	0.017	0.014	0.016	0.016	0.022	0.112	0.308	0.553	0.024	0.098	0.117	1.000
		Singleton	107	0.009	0.006	0.008	0.008	0.012	0.064	0.265	0.458	0.071	0.078	0.495	1.000
Duplications	100	<1%	6545	0.433	0.439	0.432	0.403	0.477	0.722	0.714	0.241	0.441	0.258	0.068	0.328
		Singleton	1442	0.094	0.098	0.098	0.083	0.103	0.570	0.933	0.201	0.677	0.719	0.319	0.259
	500	<1%	783	0.053	0.052	0.052	0.038	0.089	0.851	0.896	0.313	0.007	0.002	**0.002**	0.118
		Singleton	226	0.015	0.015	0.015	0.009	0.034	1.000	1.000	0.221	0.004	**0.001**	**0.002**	0.112
Number of genes within CNVs per individual															
All	100	<1%	10515	1.470	1.517	1.476	1.239	1.579	0.664	0.728	0.056	0.869	0.600	0.016	0.068
		Singleton	1849	0.333	0.303	0.346	0.272	0.285	0.366	0.209	0.519	0.773	0.440	0.751	0.172
	500	<1%	1014	0.402	0.329	0.416	0.227	0.580	0.091	0.058	0.332	0.023	0.102	0.004	0.030
		Singleton	270	0.124	0.082	0.120	0.088	0.169	0.053	0.088	0.877	0.059	0.369	0.221	0.401
Deletions	100	<1%	3970	0.365	0.318	0.359	0.329	0.437	0.056	0.179	0.783	0.026	0.393	0.062	0.592

				Burden					p-value						
CNV type	**Size (KB)**	**Frequency**	**Events**	**BD**	**CON**	**BD I**	**BD II**	**SAB**	**BD CON**	**BD I CON**	**BD II CON**	**SAB CON**	**BD I SAB**	**BD II SAB**	**BD I BD II**
		Singleton	918	0.132	0.088	0.142	0.107	0.173	0.010	0.003	0.447	0.008	0.702	0.151	0.387
	500	<1%	231	0.091	0.048	0.097	0.062	0.161	0.010	0.007	0.560	**0.001**	0.241	0.054	0.449
		Singleton	107	0.052	0.021	0.052	0.033	0.131	0.015	0.015	0.542	$\mathbf{7 \times 10^{-4}}$	0.018	0.106	0.597
Duplications	100	<1%	6545	1.105	1.199	1.117	0.910	1.142	0.285	0.373	0.018	0.843	0.864	0.106	0.054
		Singleton	1442	0.271	0.277	0.279	0.208	0.259	0.868	0.953	0.130	0.798	0.817	0.256	0.118
	500	<1%	783	0.312	0.281	0.319	0.165	0.420	0.449	0.365	0.194	0.185	0.254	0.009	0.021
		Singleton	226	0.106	0.080	0.103	0.056	0.167	0.245	0.276	0.470	0.052	0.257	0.033	0.155
Total distance (KB) covered by CNVs per individual															
All	100	<1%	10515	388	372	391	349	467	0.100	0.100	0.531	**0.001**	0.003	**0.001**	0.095
		Singleton	1849	372	327	381	314	424	0.068	0.064	0.718	0.124	0.425	0.083	0.150
	500	<1%	1014	935	911	942	839	984	0.602	0.570	0.549	0.403	0.627	0.127	0.313
		Singleton	270	1075	955	1171	869	930	0.345	0.157	0.556	0.918	0.352	0.602	0.175
Deletions	100	<1%	3970	274	256	275	258	324	0.055	0.094	0.896	0.005	0.083	0.029	0.444
		Singleton	918	322	289	327	286	378	0.213	0.236	0.944	0.110	0.453	0.185	0.463
	500	<1%	231	954	938	973	111	1196	0.837	0.715	0.292	0.151	0.235	0.041	0.274
		Singleton	107	1069	1052	1162	813	1237	0.899	0.539	0.374	0.520	0.821	0.085	0.213
Duplications	100	<1%	6545	379	368	380	337	469	0.352	0.397	0.414	0.012	0.010	0.003	0.133
		Singleton	1442	389	344	399	303	500	0.149	0.138	0.458	0.063	0.137	**0.002**	0.083
	500	<1%	783	923	893	928	850	922	0.588	0.601	0.796	0.781	0.957	0.543	0.491
		Singleton	226	1096	946	1179	881	910	0.317	0.191	0.799	0.899	0.442	0.666	0.262

Biol Psychiatry. Author manuscript; available in PMC 2020 July 15.

Table 3.

Follow-up of previously reports of increased copy number variant burden in BD. Listed are all findings identified from manual curation of literature on rare copy number variant (CNV) burden in BD with a p-value below 0.05 in the original report. Details of the test performed in the original report that was reproduced in ICCBD are described in the test parameter fields and include the phenotypes compared, the definition of burden, and the filters applied for CNV frequency, type and size. Ratios were calculated as the burden in the first phenotype in the comparison field relative to the second phenotype. When the original report did not specify the CNV size studied all CNVs greater than 100KB were included in the ICCBD. Original reports where the test included either SCZ cases or BD trios could not be followed-up in the ICCBD, but are included in this table so as to consolidate all of the previously significant findings in rare CNV studies of BD. BD cases and controls in Bergen (2011) are part of the ICCBD, so for these follow-up tests only the ICCBD samples not in the original report were utilized. The controls in Grozeva (2010) comprise half of the BDRN controls in ICCBD; since there was no case overlap between these studies, all of these controls were included in the follow-up test in the ICCBD. The cases in Green (2015) are the BDRN cases in ICCBD, though these findings were not followed up in ICCBD as they involved SCZ cases. Early-onset was defined as less than 21 in Priebe (2012) and less than 18 in Zhang (2009) and ICCBD. Family history in Malhotra (2011) was defined as having a relative with bipolar disorder (I, II, SAB), schizophrenia, autism, MDD or intellectual disability; in ICCBD, it was defined as having a family member with any psychiatric history. Reported p-values for ICCBD are 1-sided from using 10,000 permutations to test for enrichment in the direction observed in the original report. Asterisks denote nominal significance observed in ICCBD. BD – bipolar disorder; CON – control; EO – early-onset bipolar disorder; FAM – bipolar disorder with a family history of psychiatric illness; SPOR – sporadic bipolar disorder (i.e., no family history of psychiatric illness); SCZ – schizophrenia; SING – singleton; CNV – copy number variant; SCZ – schizophrenia; NR – not reported; NA – not applicable; N – number of individuals for the groups listed in the comparison field

Test parameters					**Original report**				**ICCBD**		
Comparison	**Freq**	**Type**	**Burden definition**	**Size (KB)**	**Author (year; reference)**	**N**	**Ratio**	**p-value**	**N**	**Ratio**	**p-value**
BD/CON	< 1%	CNV	Gene set enrichment: “psychological disorders”	>=100	Zhang (2009; 15)	998/1001	NR	6×10^{-6}	6353/8656	0.90	0.806
BD/CON	< 1%	CNV	Gene set enrichment: “behavior (learning)”	>=100	Zhang (2009; 15)	998/1001	NR	6×10^{-3}	6353/8656	1.04	0.603
BD/CON	< 1%	CNV	Number of CNV per individual	>=100	McQuillin (2011; 21)[a]	500/500	0.82	0.0360	6353/8656	1.01	0.700
BD/CON	< 1%	CNV	Number of CNV per individual	>=1000	Grozeva (2013; 23)	1650/10259	0.58	0.0100	6353/8656	0.84	0.962
BD/CON	< 1%	CNV	Number of *de novo* CNV per individual	>=10	Malhotra (2011; 16)	185/426	4.78	0.0090	NA	NA	NA
BD/CON	< 1%	CNV	Number of *de novo* CNV per individual	>=10	Georgieva (2014; 17)	768/3208	2.15	0.0007	NA	NA	NA
BD/CON	< 1%	DEL	Number genes hit by CNV per individual	>=500	Bergen (2012; 20)[b]	834/2087	2.12	0.0390	5436/6987	1.73	0.064
BD/CON	< 1%	DEL	Number of CNV per individual	>=100	Grozeva (2010; 22)	1697/2806	0.83	0.0100	6353/8656	1.01	0.355
BD/CON	< 1%	DEL	Number of CNV per individual	>=1000	Grozeva (2013; 23)	1650/10259	0.44	0.0300	6353/8656	1.45	0.102
BD/CON	< 1%	DEL	Number of CNV per individual	100–200	Grozeva (2010; 22)	1697/2806	0.83	0.0300	6353/8656	0.98	0.636

Test parameters					Original report				ICCBD		
Comparison	**Freq**	**Type**	**Burden definition**	**Size (KB)**	**Author (year; reference)**	**N**	**Ratio**	**p-value**	**N**	**Ratio**	**p-value**
BD/CON	< 1%	DEL	Number of CNV per individual	200–500	McQuillin (2011; 21)[a]	500/500	0.64	0.0390	6353/8656	1.04	0.312
BD/CON	SING	CNV	Number genes hit by CNV per individual	>=100	Bergen (2012; 20)[b]	834/2087	1.37	0.0320	5436/6987	1.00	0.428
BD/CON	SING	DEL	Average size of genic CNVs	NR	Priebe (2012; 26)	882/872	1.90	0.0140	6353/8656	1.04	0.346
BD/CON	SING	DEL	Proportion of individuals with a CNV	>=100	Zhang (2009; 15)	998/1001	1.33	0.0070	6353/8656	1.17	0.012
BD/CON	SING	DEL	Number genes hit by CNV per individual	>=100	Bergen (2012; 20)[b]	834/2087	1.93	0.0100	5436/6987	1.24	0.153
BD/CON	SING	DEL	Number of CNV per individual	>=100	Zhang (2009; 15)	998/1001	1.38	0.0100	6353/8656	1.17	0.014
BD/CON	SING	DEL	Number of CNV per individual	>=100	Bergen (2012; 20)[b]	834/2087	1.35	0.0130	5436/6987	1.12	0.150
BD/CON	SING	DEL	Total size of genic CNVs	NR	Priebe (2012; 26)	882/872	1.84	0.0140	6353/8656	1.07	0.240
BD/SCZ	< 1%	CNV	Number of *de novo* CNV per individual	>=100	Georgieva (2014; 17)	768/1115	0.51	0.0150	NA	NA	NA
BD/SCZ	< 1%	DEL	Number of exonic CNV per individual	>=1000	Green (2016; 13)	2591/6882	0.40	0.0009	NA	NA	NA
BD/SCZ	< 1%	DUP	Number of exonic CNV per individual	500–1000	Green (2016; 13)	2591/6882	0.80	0.0450	NA	NA	NA
EO/CON	< 1%	CNV	Proportion of individuals with a genic CNV	NR	Priebe (2012; 26)	291/872	1.26	0.0009	2235/8656	0.96	0.529
EO/CON	< 1%	CNV	Number of *de novo* CNV per individual	>=10	Malhotra (2011; 16)	107/426	6.22	0.0060	NA	NA	NA
EO/CON	< 1%	DUP	Proportion of individuals with a genic CNV	NR	Priebe (2012; 26)	291/872	1.34	0.0004	2235/8656	0.97	0.275
EO/CON	< 1%	DUP	Number of genic CNV per individual	NR	Priebe (2012; 26)	291/872	NR	0.0240	2235/8656	1.00	0.108
EO/CON	SING	CNV	Proportion of individuals with a CNV	>=100	Zhang (2009; 15)	NR/1001	1.17	0.0390	2235/8656	1.05	0.361
EO/CON	SING	CNV	Number of genic CNV per individual	NR	Priebe (2012; 26)	291/872	NR	0.0450	2235/8656	1.14	0.251
EO/CON	SING	DEL	Average size of genic CNVs	NR	Priebe (2012; 26)	291/872	2.65	0.0056	2235/8656	1.20	0.148
EO/CON	SING	DEL	Proportion of individuals with a CNV	>=100	Zhang (2009; 15)	NR/1001	1.58	0.0010	2235/8656	1.22	0.145
EO/CON	SING	DEL	Number of CNV per individual	>=100	Zhang (2009; 15)	NR/1001	1.54	0.0040	2235/8656	1.25	0.106
EO/CON	SING	DEL	Total size of genic CNVs	NR	Priebe (2012; 26)	291/872	2.60	0.0084	2235/8656	1.19	0.141
EO/CON	SING	DUP	Proportion of individuals with a genic CNV	NR	Priebe (2012; 26)	291/872	1.45	0.0170	2235/8656	0.98	0.613
EO/CON	SING	DUP	Number of genic CNV per individual	NR	Priebe (2012; 26)	291/872	NR	0.0100	2235/8656	0.97	0.675
FAM/CON	< 1%	DUP	Number of genic CNV per individual	>=500	Malhotra (2011; 16)	107/426	1.83	0.0300	3072/8656	1.09	0.066
SPOR/CON	< 1%	CNV	Number of *de novo* CNV per individual	>=10	Malhotra (2011; 16)	78/426	4.22	0.0190	NA	NA	NA
SPOR/FAM	< 1%	CNV	Number of *de novo* CNV per individual	>=10	Georgieva (2014; 17)	318/50	0.31	0.0390	NA	NA	NA

[a] Sample sizes approximated from earlier studies of the same cohort

[b] This study included the ICCBD SWEBIC Affymetrix cohort, thus our replication test excluded this cohort

Biol Psychiatry. Author manuscript; available in PMC 2020 July 15.