Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Lauren Canary Date

Genome-Wide Burden of Copy Number Variations in Schizophrenia and Bipolar I Disorder: A case-control study

By

Lauren Canary
Degree to be awarded: MPH

Epidemiology

_____

Jennifer Mullé
Committee Chair

Genome-Wide Burden of Copy Number Variations in Schizophrenia and Bipolar I Disorder: A case-control study

By

Lauren Canary

B.S.
The George Washington University
2011

Thesis Committee Chair: Jennifer Mullé, MHS, Ph.D.

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology
2014

Abstract

Genome-Wide Burden of Copy Number Variations in Schizophrenia and Bipolar I Disorder: A case-control study
By Lauren Canary

We used genome-wide data to evaluate the burden of copy number variations (CNV) in 479 patients with bipolar disorder type I (BPI), 554 patients with schizophrenia (SZ) and 1014 population-based controls. All study participants were of Ashkenazi Jewish descent. We used PLINK and R Studio software to analyze the statistical associations between presence of CNV with SZ and BPI. The burden of rare (<1%) CNV deletions was statistically significant ($P<0.05$) at various sizes (minimum sizes of 1kb, 100kb, 500kb, and 1mb) in SZ patients when compared to controls. The burden of rare CNV in BPI cases was not statistically significant, even when filtered by early onset, defined as mania occurring at age 21 or younger as well as onset defined as mania occurring at age 18 or younger. Rates of large, rare CNV were different between SZ and BPI patients. Our findings support a genetic distinction between schizophrenia and bipolar I disorder.

Genome-Wide Burden of Copy Number Variations in Schizophrenia and Bipolar I Disorder: A case-control study

By

Lauren Canary

B.S.
The George Washington University
2011

Thesis Committee Chair: Jennifer Mullé, MHS, Ph.D.

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology
2014

Introduction

Mental illness poses a significant burden to public health. According to the 2010 study on the Global Burden of Disease, mental and substance use disorders were the leading global cause of all non-fatal burden of disease [1]. It is imperative that we improve our capabilities to identify susceptible individuals and to identify targets for more efficacious treatments. Genetic studies of mental illness provide one source for achieving these goals. Though it has long been posited that one's risk for developing certain mental illnesses was a contribution of environmental and genetic factors, we are closer now than ever before to fully understanding the genetic architecture of various diseases. This exciting area of research has the potential to forge improvements in identifying susceptible individuals, diagnosing illness, and crafting effective, individualized treatments.

Identification of copy number variations in the human genome

Until 2004, it was common when investigating potential genetic contributions to disease to look at single-nucleotide polymorphisms (SNP) in DNA. However, years later researchers began looking at dosage changes of DNA known as copy number variation (CNV). A CNV is a type of genetic structural change wherein a large section (typically over 1000 nucleotides in length) of DNA is duplicated or deleted resulting in certain chromosomes with extra pieces of DNA and others missing segments[2]. In 2004, Sebat et al. identified CNV in healthy individuals spanning 70 different genes known to affect neurological function, cell growth, metabolism, and disease[3]. In the same year, Iafrate et al. identified hundreds of loci in the human genome which contain genomic imbalances that overlapped with genes showing distinct heterogeneity in 10% of subjects examined, suggesting large CNV play a role in genetic diversity in the population and a potentially important role in susceptibility to disease[4]. McCarroll et al. expanded on this knowledge in 2006 with the development of a systematic method to use publicly accessible genetic data on *HapMap* (http://hapmap.ncbi.nlm.nih.gov/) to identify deletion variants in linkage disequilibrium with nearby SNP, meaning that their association to disease could be evaluated in whole-genome association studies[5]. These seminal studies pushed the field of genetic epidemiology to a new level that allows us to systematically study CNV within populations and to appreciate their role in both health and disease.

Formation of CNV

CNV formation is achieved via four means of structural variation[2]. The first, non-allelic homologous recombination, involves alignment and crossing over between two sites in the genome usually during meiosis (less often in mitosis). Another method, non-homologous end-joining, occurs when a break in DNA is repaired with an error. Fork-stalling and template switching, on the other hand, involve errors in DNA replication[2]. Finally, retrotranspositions known as long interspersed nuclear elements are pieces of DNA that can amplify themselves and relocate to create changes in the genome. These sizeable changes in the genome (usually greater than 1,000 base pairs per generation) can have a significant functional impact in relation to disease.

The contribution of genetics to disease can be complex. There are two current models for genetic contribution to disease: common variant common disease (CVCD) and rare variant common disease (RVCD). The CVCD model states that genetic risk is due to many, high frequency variants that cause a moderate level of risk [6]. Alternatively, the RVCD model states that genetic risk is due to few, rare variants, which cause a high level of risk [7]. The CVCD and RVCD models can be employed for us to understand the importance of both small and large genetic changes association with disease, though most evidence specifically concerning CNV-associated risk follows the RVCD model. Notably, these large, rare genetic changes are not always inherited; some CNV are created *de novo*. *De novo* CNV formation is seemingly random, yet there are similarities in *de novo* CNV distribution across unrelated individuals because some areas of the genome, known as low copy repeats, are more prone to CNV formation than others.

Recent studies of the genetics of mental illness have employed the use of genome-wide CNV studies, which evaluate the statistical associations between large genetic variants and disease. These associations between genetic variants and disease can become confounded when genetically heterogenous samples are analyzed, due to hidden population stratification which can result in differences in CNV between cases and controls [8]. Ideally, genetic association studies should use homogenous ethnic populations so as to lessen the chance of finding spurious associations. The Ashkenazi Jewish population, which is living mostly in the U.S.

and central and eastern Europe, is often used in genetic association studies as it is particularly genetically homogenous, having descended from a small founder population approximately 500 years ago[9]. This thesis aims to carry-out a genome-wide CNV study in Ashkenazi Jewish subjects in order to better understand the genetic underpinnings and potential interrelatedness of two mental illnesses: schizophrenia and bipolar disorder-type I.

Schizophrenia

Schizophrenia, which affects about 1% of the population, is comprised of 3 families of symptoms: positive, negative, and cognitive [10]. Positive symptoms are comprised of hallucinations, delusions, thought disorders, and movement disorders, all of which may range in severity and frequency[11]. Negative symptoms are very similar to depressive symptoms and include flat affect, lack of pleasure, lack of ability to initiate or sustain planned activities, and lack of verbal communication[11]. Cognitive symptoms include problems with executive functioning, focusing, and working memory[11]. Lifetime rates of schizophrenia are similar in men and women, though later onset is more common in females and there are differences in symptomology between the sexes[12]. For decades there has been evidence that schizophrenia is partly-genetic[13, 14 15], though the genetic architecture is still under debate [16 7].

Recently, evidence of an excess of rare, large CNV in schizophrenia has been observed in case-control studies. Researchers have found as much as a 3-fold enrichment of rare CNV in schizophrenia cases versus controls [17], though more conservative estimates place CNV at a 1.1-1.5-fold enrichment (International Schizophrenia Consortium, 2008). Regardless of enrichment, the presence of CNV in schizophrenia is supported by several case-control studies [18, 19]. Family-based studies have also implicated a high rate (5-10%) of *de novo* CNVs in schizophrenia, though it is unclear how big of a difference, if any, exists between familial versus sporadic cases of schizophrenia [20 21 22].

Several CNV have been identified as being associated with schizophrenia. A large deletion (3mb) at 22q11.21 has been demonstrated to be a significant risk factor (25% of those with the CNV deletions will have symptoms of psychosis compared to an average of 1% in the general population) [2, 23], though recent studies have implicated other structural variations: deletions at chr 7q11.23[24], chr1q21[25] , chr3q29[26], chr15q11 & chr 15q13[27], chr2p16[17], Neurexin-1 [28], and duplications of chr16p11[29], chr7q36[30]. Some studies have identified

affected genes with functions related to synaptic activity and neurodevelopment [17, 21], while another study showed involvement of the genes coding for parts of the N-methyl-d-aspartate receptor (NMDAR-neuronal activity-regulated cytoskeleton-associated protein) and postsynaptic signaling complexes [22].

Bipolar Disorder

Patients with bipolar disorder experience episodes of depression that alternate with abnormally elevated energy levels, cognition, and mood[31]. Bipolar type I is distinguished from bipolar type II based on the severity of manic and mixed episodes. Both bipolar types I and II present with depressive episodes. However in bipolar type I, manic or mixed episodes may last 7-10 days or may be severe enough to require hospitalization whereas in bipolar type II, manic episodes are less severe and are characterized as 'hypomanic'[31].

Rare CNV are implicated in bipolar disorder [21, 32-35], though the pattern is not as consistent as we see in schizophrenia, and there is limited evidence of their importance in relation to the disease. While some evidence exists, it appears that large deletions play very limited role in bipolar risk [21, 35-37]. Case-control studies on the subject are inconsistent; two studies reported enrichment of rare CNV [32 33] with the effect greater in those with early age of onset with a small effect size (OR~1.5), while two other studies [35, 38] did not support this finding. Family-based studies have demonstrated frequencies of *de novo* CNVs to be higher in bipolar compared to healthy individuals (4.3% vs. 0.09%) and even higher in early onset bipolar (5.6%), when defined as onset prior to 18 years of age [21]. Since inherited and *de novo* CNV are apparent in those with early-onset bipolar I disorder [21], some support the idea of subclasses or a continuum of bipolar disorder, where those with early onset represent one extreme or subclass [32, 33, 39, 40]. Further complicating how the field ought to classify bipolar disorder, there is limited evidence regarding which genes are involved in CNV formation. Some have posited enrichment of genes associated with psychological disorders and those associated with learning [33] and enrichment of genes affecting cell shape [21], but not those affecting neuronal function or development. To our knowledge the varying symptomology present in bipolar disorder, which make bipolar disorder types I and II distinct, has not been taken into account in prior large CNV studies.

Relationship between Schizophrenia and Bipolar Disorder

Though classically schizophrenia and bipolar have been categorized as psychotic and mood disorders, respectively, this dichotomy has been questioned in recent years raising questions about their potential existence along one spectrum[41]. Indeed, depressive and cognitive symptoms are defining features of both schizophrenia and bipolar, and the psychotic symptoms that underlie schizophrenia may appear quite similar to the disorganized and grandiose thoughts present in bipolar disorder, especially the more extreme manic symptoms associates with bipolar type I[42]. Further questioning the distinction between these two illnesses has been the pattern of inheritance observed in bipolar disorder and schizophrenia; family studies have shown co-aggregation of the disorders [43] [44] [45] [46]. It has been suggested that schizophrenia and bipolar are in fact the same disorder, but with schizophrenia on the more extreme end of the continuum[47, 48]. Supporting this, researchers have identified certain chromosomal regions[49, 50] and genes[51] [50, 52-54] [55] [56] shared by schizophrenia and bipolar.

The numerous reports indicating a potential role of CNV in schizophrenia, combined with the putative relationship between schizophrenia and bipolar disorder (especially bipolar type I), suggest that CNV may also increase risk for bipolar disorder. Working under this suggestion, the hypothesis of this study is that we will find a significant burden of large, rare CNV in schizophrenia and bipolar I cases when compared to unaffected controls, and that the rate of CNV in schizophrenia cases will be similar to that of bipolar I cases. To aid comparison of our data with other published studies, we also evaluate the CNV burden in early onset bipolar I disorder, defined as onset prior to age 18 (as in Grozeva et al.[37]) or onset prior to age 21 (as in Priebe et al.[32]). We hypothesize that if schizophrenia and bipolar I disorder are indeed related then the CNV frequency seen in bipolar cases will be similar to that seen in schizophrenia cases. This analysis is unique from previously published reports of CNV in bipolar disorder [32, 34, 37]; it aims to use a genetically homogeneous population; it aims to understand genetic differences in early-onset bipolar disorder by looking both at onset occurring at age 21 and under as well as age 18 and under; and finally it analyzes bipolar I cases in solidarity, rather than grouping them with bipolar II cases.

Methods

Subjects

Patients of Ashkenazi Jewish descent with schizophrenia (SZ) or bipolar disorder-type I (BPI) were recruited over a six-year period by the Johns Hopkins Epidemiology-Genetics Program in Psychiatry, via advertisements in newspapers and Jewish newsletters, talks to community organizations, meetings with Jewish leaders, and the study website. Cases were eligible for inclusion in these analyses if all four of the proband's grandparents were of Ashkenazi Jewish descent and if the proband met DSM-IV criteria for a SZ or BPI diagnosis. Interviewers collected the participant's psychiatric treatment history and the subjects were asked to sign release forms allowing study staff to receive copies of their psychiatric treatment records. Blood samples were obtained from probands, as well as parents when feasible. Probands were assessed for psychiatric illness according to an established consensus-based procedure[57, 58]. The study participant was interviewed in tape-recorded sessions by a clinical psychologist twice, first with the use of the Diagnostic Interview for Genetic Studies (DIGS, version 2.0; revised for DSM-IV), and secondly with the Structured Interview for DSM-IV Personality Disorders (SID-P)[59]. The interviewer also completed a collateral interview with at least one informant who knew the proband. No subject had a previous clinical genetic diagnosis. The SID-P interview was used in order to identify personality disorders in controls that might confound the analysis, including: antisocial, borderline, histrionic, compulsive, schizoid, schizotypal, and paranoid. Clinical materials and tape recordings of interview sessions were reviewed by a consensus committee, which consisted of two psychiatrists who went through a DSM-IV diagnostic checklist, which contained the necessary criterion for 26 Axis I DSM-IV disorders. Disorders were rated as absent, possibly present, probably present, definitely present, or unknown. When applicable, age of onset was recorded. Ratings assigned independently by the two members of the committee were compared, and if there existed any disagreement regarding ratings of the Axis I DSM-IV diagnoses, course of illness ratings for those presenting with psychosis, or age of onset (greater than a 4 year discrepancy), then the committee met to discuss discrepancies. Control subjects were recruited from three cohorts of Ashkenazi individuals: a Crohn's Disease cohort (n=258), the Ashkenazi Jewish Control Registry at Johns Hopkins University (n= 538), and a study of Parkinson's Disease and dystonia (n=266)[26]. Control

subjects from the Johns Hopkins University registry were administered a questionnaire regarding psychiatric conditions. Control subjects from the Crohn's Disease cohort and from the Parkinson's Disease and dystonia cohort were not screened for psychiatric illness. After all screening was completed, our analyses included 554 SZ cases, 479 BPI cases, and 1014 controls. All individuals completed an informed consent and recruitment methods and protocols for collecting clinical data and blood samples were approved by the Johns Hopkins University Internal Review Board. All data were anonymized before analysis at Emory University.

Genotyping and CNV analysis

DNA was extracted from blood samples using either the Gentra Puregene Kit or QIAGEN DNeasy Blood and Tissue Kit. Genotyping was performed at Emory University with Affymetrix Human Genome-wide SNP array 6.0 with use of the Birdseed Algorithm implemented in Affymetrix Power Tools software (version 1.10.0; Santa Clara, CA). Individual genotypes with confidence scores <0.9 were excluded. Affymetrix Power Tools was used to calculate normalization and log ratio data. Three algorithms (GLAD[60], GADA[61], and BEAST[62]) were used to extract and analyze log(2) ratio data. CNV called by only a single algorithm were removed from analysis. When algorithms disagreed on endpoint determination, the largest interval was considered. Samples were excluded (1) if they failed the manufacturer's recommended cutoffs for chip quality control; (2) patterns of SNP genotypes revealed inconsistent family structure (trios only) or unknown duplicates or if the heterozygosity was greater than three standard deviations (SD) above the mean and the missing data rate was greater than 3 SD above the mean (indicating potential sample mixing); (3) if patterns of SNP genotypes revealed cryptic relatedness determined by pairwise identity-by-state analysis (first- and second- degree relationships were removed, controls preferentially removed when possible); or (4) if the number of detected CNV were greater than 3 SD from the mean number of CNV in the overall sample[26]. CNV from the schizophrenia cohort were filtered against the Database of Genomic Variants (http://projects.tcag.ca/variation/), and only rare CNV were retained for analysis.

Statistical analysis

Permutation analysis of CNV burden in cases and controls was calculated using PLINK (http://pngu.mgh.harvard.edu/~purcell/plink/). Multiple analyses were performed in order to stratify CNV based on their size and frequency, and to stratify BPI cases based on age of onset and frequency. For all samples, CNV analyses were performed separately for CNV with sizes of 1kb, 100kb, 500kb, and 1mb. BPI CNV were also analyzed based on frequency with one analysis performed on all CNV and another on only rare CNV, which were defined here as those occurring in less than 1% of study subjects. An overlap of 50% genetic homogeneity was used to identify common CNV. For age-of-onset filtering of BPI CNV data, onset was defined as first episode of mania. Age of onset was analyzed in three ways: no filtering (n cases = 479); early onset defined as mania age of onset < 18 years (n cases = 174); and early onset defined as mania age at onset < 21 years (n cases = 260). These ages were chosen to aid comparison with prior published data[32, 37]. Fisher's exact test odds ratios and their associated 95% confidence intervals were calculated using R-Studio software (version 0.98.501; Boston, MA).

Results

Bipolar I disorder

We identified a total of 5147 putative CNV, comprised of 3078 deletion and 2069 duplication segments, in the initial sample of 479 bipolar I disorder cases and 1011 controls. In order to carry out a genome-wide burden analysis, we systematically analyzed the samples for significant differences in both rare (present in <1% of samples) and common (any frequency of occurrence) CNV occurrences of varying sizes (1kb, 100kb, 500kb, and 1mb) and at varying stages of disease onset for bipolar I disorder (18 years or younger, 21 years or younger).

Our main outcome was the odds of having at least one CNV in the sample of cases compared to that of controls. We the used Fisher's exact test to evaluate the statistical significance associated with the odds ratios we calculated. These odds ratios are depicted in Figures 1 and 2. Notably, in the BPI analysis CNV were not associated with an odds ratio significantly greater than 1, regardless of age of onset, CNV size, and CNV frequency.

Numerous odds ratios were significantly less than 1, indicating a greater burden of CNV in controls compared to BPI cases (**Tables 1 & 2**).

Among all ages of BPI onset, the CNV rate was higher in cases (2.303 CNV per person) than controls (1.954 CNVs per person) for deletions of at least 1kb in size (*P*= 0.0039). This finding also held for duplications of 1mb in size (0.0626 vs 0.0336; *P*=0.0467), though the difference was not significantly different at minimum CNV sizes of 100kb or 500kb. When limited to cases with earlier ages of onset, there was no significant increase in CNV rate for cases compared to controls for deletions or duplications. When we limited our samples to rare CNV, the findings were similar. We found the rate of CNV to be higher in BPI cases than controls for CNV deletions of at least 1kb (1.5510 vs 1.1280; *P*=0.0008) and CNV duplications of at least 500kb (0.1169 vs. 0.0692; *P*=0.0245). Again, we found no statistically significant results when we limited to earlier ages of onset.

When comparing the CNV proportion, defined as the number of individuals with at least one CNV, among BPI cases and controls we found no statistically significant difference among all sizes of CNV and among all ages of disease onset. This finding was repeated when we limited the sample to rare CNV.

The CNV length, defined as the total length spanned by CNV per person, was statistically significantly larger in BPI cases than controls when evaluating CNV duplications of at least 1kb (518kb vs 398.8kb; *P*=0.0241) and 100kb (652.8kb vs 470.5kb; *P*=0.0035) in size. This finding was also observed in BPI cases limited to an age of onset less than or equal to 18 years with CNV duplications limited to at least 1kb (478.7kb vs 398.7kb; *P*=0.0387) and 100kb (565.5kb vs 470.5kb; *P*=0.0327). Notably, this finding was not significant in any subset of CNV deletions. When limited to rare CNV, the CNV length of duplications greater than or equal to 100kb in size was statistically significantly larger in BPI cases when compared to controls in all ages of onset (738.7kb vs 387.1kb; $P=9.99\text{x}10^{-5}$), in onset of age 21 years or earlier (519.3kb vs 417.2kb; *P*=0.0244), and in onset of age 18 or earlier (662.9kb vs 456.4kb; *P*=0.0018). There were no significant differences for duplications of other sizes, or for any rare CNV deletions. The average CNV length was greater in BPI cases than controls only for duplications of at least 100kb for all CNV (399.8kb vs 357.7kb; *P*=0.0444) and when limited to rare CNV (400.5kb vs 321.6kb; *P*=0.0096). When we limited analyses to cases with earlier BPI onset the difference was not statistically significant. There were no

statistically significant differences in CNV deletion length for any minimum size or age of onset.

Schizophrenia

We identified 2356 putative CNV, comprised of 1248 deletion CNV and 1108 duplication CNV, among the 554 schizophrenia cases and 1014 controls. We systematically analyzed the samples for significant differences in CNV occurrences of varying sizes (1kb, 100kb, 500kb, and 1mb), though we did not filter samples by age of onset nor frequency. All of the CNV in the schizophrenia dataset were considered to be rare CNV.

The Fisher's exact odds ratios reflecting the burden of CNV in schizophrenia cases compared to controls are depicted in figures 3 and 4. When analyzing CNV deletions, those of at least 1kb in size were nearly one and a half times (OR=1.4441; 95% CI=1.1557-1.8042) more likely to occur in schizophrenia cases than controls (P=9.6490x10^{-4}). This effect size was similar in CNV deletions of at least 100kb in size wherein the odds were 1.3572 (95% CI= 1.0226-1.7973) times greater in SZ cases than controls (P=0.03267). With increasing size of CNV deletions, we observed a greater effect size; those deletions of at least 500kb in size were associated with an odds ratio of 2.7822 (95% CI=1.0385, 7.8992; P=0.03134). The largest effect size was observed with CNV deletions of at least 1mb in size with cases being more than 4.5 times (OR=4.6367) more likely to appear in controls (P=0.008475), though the confidence interval for this effect size ranged greatly (95% CI=1.3296-20.3440).

In contrast to our observations of deletions, small CNV duplications were not associated with schizophrenia. However, duplications of at least 500kb in size were twice (OR=2.0937) as likely to appear in SZ cases than controls (95% CI= 0.9987-4.4276; P=0.0442). The odds ratio associated with even larger duplications of at least 1mb in size were not statistically significant (OR=2.46; 95% CI= 0.7442, 8.6482; P=0.0973), but this is likely due to the infrequent number of observations, which have decreased the precision of the estimate.

We found that the CNV rate was significantly larger in SZ cases than controls when CNV deletions were limited to at least 1kb in size (1.1100 vs 0.5759; P=9.99 x10^{-5}) and when limited to at least 1mb in size (0.0181 vs 0.0059; P=0.0436). The difference at 100kb approached statistical significance (0.3069 vs 0.2367; P=0.0546). This finding was also

apparent in duplications of 500kb in size (0.0325 vs 0.0168; *P*=0.0377), and approached significance in duplications of both 100kb (0.2292 vs 0.1775; *P*=0.0700) and 1mb (0.0144 vs 0.0059; *P*=0.0776) in size.

We found statistically significant differences between cases and controls for all sizes of CNV. The proportion of individuals with at least one CNV was highest for both SZ cases and controls when filtered by at least 1kb in size with 38.99% of cases having CNV and 30.67% of controls with CNV (*P*=0.0005). As we filtered CNV by increasing size, the proportions of individuals with an event decreased, but the statistical effect remained. At a minimum 100kb filter, 19.31% of cases and 14.99% of controls exhibited CNV (*P*=0.0192). At a minimum 500kb filter, 2.17% of cases and 0.79% of controls exhibited CNV (*P*=0.0205). With large CNV of at least 1mb in size, 1.81% of cases and 0.39% of controls exhibited CNV (*P*=0.0082). Notably, the statistically significant differences observed in all sizes of CNV deletions were not observed with CNV duplications. There were no statistically significant differences between smaller CNV duplications limited to at least 1kb or 100kb in size in relation to SZ. However, larger duplications were more apparent in SZ cases than controls; at a minimum size of 500kb, CNV duplications were significantly more common in cases than controls (2.35% vs 1.58%; *P*=0.0235). At a minimum CNV size of 1mb, the difference in the proportion of duplications was greater in cases (1.44%) when compared to controls (0.59%), though this difference only approached statistical significance (*P*=0.0766).

In SZ cases the total length spanned by CNV per person was statistically significantly larger in CNV duplications of at least 1kb (333.8kb vs 242kb; *P*=0.0159) and 100kb (620.8kb vs 348.9kb; *P*=0.0002) in size. The average segment size was larger in SZ cases than controls for a minimum size of 1kb deletions (219.1kb vs 117.5kb; *P*=0.0172) and duplications (188.4kb vs 142kb; *P*=0.0410) as well as minimum 100kb deletions (443kb vs 215.7kb; *P*=0.0037) and duplications (454.3kb vs 294.7kb; *P*=0.0040).

Discussion

There is increasing interest in the research of mental illness to better understand the interrelatedness of various disorders rather than thinking about them as discrete illnesses[56, 63].

Building on previous studies, we set out to use genetic analyses to shed light on the disputed relationship between schizophrenia and bipolar disorder. We have shown here that large rare deletions, and to a lesser extent duplications, of DNA are associated with schizophrenia consistent with other studies[17-30]. In bipolar I disorder, however, we find limited evidence to suggest a role played by CNV, even when limited to various ages of onset and sizes of CNV, and when we analyzed both common and rare CNV.

Prior studies have demonstrated a role for CNV in schizophrenia [17-30] and we build on that here by illustrating the role played by large CNV in comparison to smaller CNV. Particularly in CNV deletions, we saw increasingly large effect sizes with increasingly large CNV. The highest effect size we witnessed was the odds of a deletion of at least 1mb in size being four times greater in schizophrenia cases then controls (P=0.0085). It is logical to presume that larger CNV would span a greater number of genes and thus would be more likely to have a potential effect on an area of the genome necessary for proper psychological function. Future studies should aim to better understand which regions are highly affected in association with various types of symptoms of schizophrenia.

It has been hypothesized that there is also some role played by CNV in bipolar disorder [21, 32, 33]. Here, we found the occasional statistically significant test without any appreciable trend. Our probability of a type I error was 5%, and the number of significant results in our one-sided analysis of rare CNV in bipolar disorder was only 6% (6 significant results out of 96 tests), so it is possibly that the significant results we did find in the bipolar analyses were due to improper rejection of our null hypothesis. In contrast, our one-sided analysis of rare CNV Schizophrenia resulted in 36% statistically significant findings. It is noteworthy, though, that *controls* had significantly higher odds of CNV than BPI cases at various sizes and ages of onset in this study, a finding which has been reported previously[37]. Moreover, this effect size actually decreased with lower ages of onset which is supported by prior findings form Grozeva et al.[37], but counters others who found increased effect sizes in earlier onset cases of bipolar disorder[32]. Of note, we defined 'onset' in our study as the time at which manic symptoms began, so as to maintain consistency with other studies. It may be helpful in future studies to explore various definitions of 'onset'. Our analyses here support numerous studies showing limited role for CNV in bipolar disorder[37, 38], though we brought a more

critical assessment than previously performed by looking at many different sizes of CNV, various ages of onset, rare and common CNV, and by using a genetically homogenous population. Though we limited our inclusion criteria to patients with bipolar I disorder, future studies which include both BPI and BPII, but distinguish the two, may be helpful in better understanding the genetic underpinnings of bipolar disorder as they may better elucidate the connection between genetic and symptomatic variation.

We have found here that CNV, especially large rare deletions, appear to play a role in schizophrenia, but not in bipolar I disorder. While these mental illnesses may show some overlap in symptomology, there is not sufficient evidence to suggest that this overlap is due to similarities in structural variations in DNA. We hypothesized that if bipolar disorder type I and schizophrenia were similar in their CNV burden, then we would expect to see similar rates of CNV in both disorders. Here, when we analyzed large, rare CNV, consistent with the RVCD model, the rate of CNV in SZ cases was more than twice the rate of BP cases (1.81 vs 0.59 CNV per 100 cases, respectively) even when the age of BPI onset was limited to 21 years (0.77 CNV per 100 cases) or 18 years (0.57 CNV per 100 cases). Moreover, in our schizophrenia analysis we found statistically significant differences between cases and controls for all sizes of CNV, whereas in our bipolar disorder analysis we found no statistically significant difference in CNV proportion regardless of size of CNV. These findings, however, do not mean that we should necessarily approach schizophrenia and bipolar disorder as distinct entities.

Just because we may be closer to saying that large, rare CNV are involved in schizophrenia but not in BPI does not necessarily mean that there are no genetic similarities between the two disorders. The current state of genetic research of bipolar disorder promotes the effects of a number of small, common variants that individually might indicate a small increase in genetic risk [64]. These smaller genetic events show less penetrance and genotypic risk than large, rare CNV, though in large numbers they could have a significant biological effect. Thus, it is plausible to presume, as has been promoted by others, that some of the SNP implicated in bipolar disorder and the CNV implicated in schizophrenia may have overlapping genetic regions[65], resulting in similar changes in gene expression and protein synthesis. For example, variation in calcium-channel activity genes may have pleiotropic

effects on both schizophrenia and bipolar disorder, and this variation is associated with changes in amygdala activity during emotional processing[66]. This is just one model that might explain why symptoms are sometimes similar in patients with schizophrenia and those with bipolar type I disorder even though these disorders have different genetic underpinnings.

Here we have provided further support for large, rare CNV in schizophrenia but not for bipolar disorder type I. This analysis would suggest that schizophrenia and bipolar disorder type I and genetically distinct. However, further exploration to identify the source of functional impairment in the brain, and to examine the nature of the overlap in symptomology in these two disorders is necessitated.

Sources

1. Whiteford HA, Baxter AJ. The Global Burden of Disease 2010 Study: what does it tell us about mental disorders in Latin America? Rev Bras Psiquiatr. 2013; 35:111-2.
2. Malhotra D, Sebat J. CNVs: harbingers of a rare variant revolution in psychiatric genetics. Cell. 2012; 148:1223-41.
3. Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, et al. Large-scale copy number polymorphism in the human genome. Science. 2004; 305:525-8.
4. Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, et al. Detection of large-scale variation in the human genome. Nature genetics. 2004; 36:949-51.
5. McCarroll SA, Hadnott TN, Perry GH, Sabeti PC, Zody MC, Barrett JC, et al. Common deletion polymorphisms in the human genome. Nature genetics. 2006; 38:86-92.
6. Risch N, Merikangas K. The future of genetic studies of complex human diseases. Science. 1996; 273:1516-7.
7. Bodmer W, Bonilla C. Common and rare variants in multifactorial susceptibility to common diseases. Nature genetics. 2008; 40:695-701.
8. Liu J, Lewinger JP, Gilliland FD, Gauderman WJ, Conti DV. Confounding and heterogeneity in genetic association studies with admixed populations. American journal of epidemiology. 2013; 177:351-60.
9. Ostrer H. A genetic profile of contemporary Jewish populations. Nature reviews Genetics. 2001; 2:891-8.
10. Liddle PF. The symptoms of chronic schizophrenia. A re-examination of the positive-negative dichotomy. The British journal of psychiatry : the journal of mental science. 1987; 151:145-51.
11. Health NIoM. Schizophrenia. In: Health NIo, editor. Department of Health and Human Services2009.
12. Mendrek A, Stip E. Sexual dimorphism in schizophrenia: is there a need for gender-based protocols? Expert review of neurotherapeutics. 2011; 11:951-9.
13. Kety SS. Schizophrenic illness in the families of schizophrenic adoptees: findings frmo the Danish national sample. Schizophrenia bulletin. 1988; 14:217-22.
14. Risch N. Genetic linkage and complex diseases, with special reference to psychiatric disorders. Genetic epidemiology. 1990; 7:3-16; discussion 7-45.
15. Gottesman I. Schizophrenia Genesis: The Origins of Madness. Freeman, New York1991.
16. Pickard B. Progress in defining the biological causes of schizophrenia. Expert reviews in molecular medicine. 2011; 13:e25.
17. Walsh T, McClellan JM, McCarthy SE, Addington AM, Pierce SB, Cooper GM, et al. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. Science. 2008; 320:539-43.
18. Buizer-Voskamp JEea. Genome-wide analysis shows increased frequency of copy number variation deletions in Dutch
schizophrenia patients. . Genetic Risk and Outcome in Psychosis (GROUP) Consortium Members Biol Psychiatry. 2011; 70:655-62.
19. Kirov G, Zaharieva I, Georgieva L, Moskvina V, Nikolov I, Cichon S, et al. A genome-wide association study in 574 schizophrenia trios using DNA pooling. Molecular psychiatry. 2009; 14:796-803.

20. Xu B, Roos JL, Levy S, van Rensburg EJ, Gogos JA, Karayiorgou M. Strong association of de novo copy number mutations with sporadic schizophrenia. Nature genetics. 2008; 40:880-5.
21. Malhotra D, McCarthy S, Michaelson JJ, Vacic V, Burdick KE, Yoon S, et al. High frequencies of de novo CNVs in bipolar disorder and schizophrenia. Neuron. 2011; 72:951-63.
22. Kirov G, Pocklington AJ, Holmans P, Ivanov D, Ikeda M, Ruderfer D, et al. De novo CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia. Molecular psychiatry. 2012; 17:142-53.
23. Karayiorgou M, Morris MA, Morrow B, Shprintzen RJ, Goldberg R, Borrow J, et al. Schizophrenia susceptibility associated with interstitial deletions of chromosome 22q11. Proceedings of the National Academy of Sciences of the United States of America. 1995; 92:7612-6.
24. Mulle JG, Pulver AE, McGrath JA, Wolyniec PS, Dodd AF, Cutler DJ, et al. Reciprocal duplication of the Williams-Beuren syndrome deletion on chromosome 7q11.23 is associated with schizophrenia. Biological psychiatry. 2014; 75:371-7.
25. Consortium IS. Rare chromosomal deletions and duplications increase risk of schizophrenia. Nature. 2008; 455:237-41.
26. Mulle JG, Dodd AF, McGrath JA, Wolyniec PS, Mitchell AA, Shetty AC, et al. Microdeletions of 3q29 confer high risk for schizophrenia. American journal of human genetics. 2010; 87:229-36.
27. Stefansson H, Rujescu D, Cichon S, Pietilainen OP, Ingason A, Steinberg S, et al. Large recurrent microdeletions associated with schizophrenia. Nature. 2008; 455:232-6.
28. Kirov G, Rujescu D, Ingason A, Collier DA, O'Donovan MC, Owen MJ. Neurexin 1 (NRXN1) deletions in schizophrenia. Schizophrenia bulletin. 2009; 35:851-4.
29. McCarthy SE, Makarov V, Kirov G, Addington AM, McClellan J, Yoon S, et al. Microduplications of 16p11.2 are associated with schizophrenia. Nature genetics. 2009; 41:1223-7.
30. Vacic V, McCarthy S, Malhotra D, Murray F, Chou HH, Peoples A, et al. Duplications of the neuropeptide receptor gene VIPR2 confer significant risk for schizophrenia. Nature. 2011; 471:499-503.
31. Health NIoM. Bipolar Disorder In: Health NIo, editor. Department of Health and Human Services2009.
32. Priebe L, Degenhardt FA, Herms S, Haenisch B, Mattheisen M, Nieratschker V, et al. Genome-wide survey implicates the influence of copy number variants (CNVs) in the development of early-onset bipolar disorder. Molecular psychiatry. 2012; 17:421-32.
33. Zhang D, Cheng L, Qian Y, Alliey-Rodriguez N, Kelsoe JR, Greenwood T, et al. Singleton deletions throughout the genome increase risk of bipolar disorder. Molecular psychiatry. 2009; 14:376-80.
34. Bergen SE, O'Dushlaine CT, Ripke S, Lee PH, Ruderfer DM, Akterin S, et al. Genome-wide association study in a Swedish population yields support for greater CNV and MHC involvement in schizophrenia compared with bipolar disorder. Molecular psychiatry. 2012; 17:880-6.
35. Grozeva D, Kirov G, Ivanov D, Jones IR, Jones L, Green EK, et al. Rare copy number variants: a point of rarity in genetic risk for bipolar disorder and schizophrenia. Archives of general psychiatry. 2010; 67:318-27.

36. Xu W, Cohen-Woods S, Chen Q, Noor A, Knight J, Hosang G, et al. Genome-wide association study of bipolar disorder in Canadian and UK populations corroborates disease loci including SYNE1 and CSMD1. BMC medical genetics. 2014; 15:2.
37. Grozeva D, Kirov G, Conrad DF, Barnes CP, Hurles M, Owen MJ, et al. Reduced burden of very large and rare CNVs in bipolar affective disorder. Bipolar disorders. 2013; 15:893-8.
38. McQuillin A, Bass N, Anjorin A, Lawrence J, Kandaswamy R, Lydall G, et al. Analysis of genetic deletions and duplications in the University College London bipolar disorder case control sample. European journal of human genetics : EJHG. 2011; 19:588-92.
39. Faraone SV, Glatt SJ, Tsuang MT. The genetics of pediatric-onset bipolar disorder. Biological psychiatry. 2003; 53:970-7.
40. Potash JB, Toolan J, Steele J, Miller EB, Pearl J, Zandi PP, et al. The bipolar disorder phenome database: a resource for genetic studies. The American journal of psychiatry. 2007; 164:1229-37.
41. Ruderfer DM, Fanous AH, Ripke S, McQuillin A, Amdur RL, Gejman PV, et al. Polygenic dissection of diagnosis and clinical dimensions of bipolar disorder and schizophrenia. Molecular psychiatry. 2013.
42. Fernandes CP, Christoforou A, Giddaluru S, Ersland KM, Djurovic S, Mattheisen M, et al. A genetic deconstruction of neurocognitive traits in schizophrenia and bipolar disorder. PloS one. 2013; 8:e81052.
43. Pope HG, Jr., Yurgelun-Todd D. Schizophrenic individuals with bipolar first-degree relatives: analysis of two pedigrees. The Journal of clinical psychiatry. 1990; 51:97-101.
44. Tsuang MT, Winokur G, Crowe RR. Morbidity risks of schizophrenia and affective disorders among first degree relatives of patients with schizophrenia, mania, depression and surgical conditions. The British journal of psychiatry : the journal of mental science. 1980; 137:497-504.
45. Valles V, Van Os J, Guillamat R, Gutierrez B, Campillo M, Gento P, et al. Increased morbid risk for schizophrenia in families of in-patients with bipolar illness. Schizophrenia research. 2000; 42:83-90.
46. Cardno AG, Owen MJ. Genetic Relationships Between Schizophrenia, Bipolar Disorder, and Schizoaffective Disorder. Schizophrenia bulletin. 2014.
47. Moller HJ. Bipolar disorder and schizophrenia: distinct illnesses or a continuum? The Journal of clinical psychiatry. 2003; 64 Suppl 6:23-7; discussion 8.
48. Lapierre YD. Schizophrenia and manic-depression: separate illnesses or a continuum? Canadian journal of psychiatry Revue canadienne de psychiatrie. 1994; 39:S59-64.
49. Berrettini W. Evidence for shared susceptibility in bipolar disorder and schizophrenia. American journal of medical genetics Part C, Seminars in medical genetics. 2003; 123C:59-64.
50. Badner JA, Gershon ES. Meta-analysis of whole-genome linkage scans of bipolar disorder and schizophrenia. Molecular psychiatry. 2002; 7:405-11.
51. Lencz T, Guha S, Liu C, Rosenfeld J, Mukherjee S, DeRosse P, et al. Genome-wide association study implicates NDST3 in schizophrenia and bipolar disorder. Nature communications. 2013; 4:2739.
52. Green EK, Raybould R, Macgregor S, Gordon-Smith K, Heron J, Hyde S, et al. Operation of the schizophrenia susceptibility gene, neuregulin 1, across traditional diagnostic boundaries to increase risk for bipolar disorder. Archives of general psychiatry. 2005; 62:642-8.

53. Schumacher J, Jamra RA, Freudenberg J, Becker T, Ohlraun S, Otte AC, et al. Examination of G72 and D-amino-acid oxidase as genetic risk factors for schizophrenia and bipolar affective disorder. Molecular psychiatry. 2004; 9:203-7.
54. Craddock N, O'Donovan MC, Owen MJ. The genetics of schizophrenia and bipolar disorder: dissecting psychosis. Journal of medical genetics. 2005; 42:193-204.
55. Williams HJ, Craddock N, Russo G, Hamshere ML, Moskvina V, Dwyer S, et al. Most genome-wide significant susceptibility loci for schizophrenia and bipolar disorder reported to date cross-traditional diagnostic boundaries. Human molecular genetics. 2011; 20:387-91.
56. Craddock N, Owen MJ. The beginning of the end for the Kraepelinian dichotomy. The British journal of psychiatry : the journal of mental science. 2005; 186:364-6.
57. Fallin MD, Lasseter VK, Wolyniec PS, McGrath JA, Nestadt G, Valle D, et al. Genomewide linkage scan for schizophrenia susceptibility loci among Ashkenazi Jewish families shows evidence of linkage on chromosome 10q22. American journal of human genetics. 2003; 73:601-11.
58. Fallin MD, Lasseter VK, Wolyniec PS, McGrath JA, Nestadt G, Valle D, et al. Genomewide linkage scan for bipolar-disorder susceptibility loci among Ashkenazi Jewish families. American journal of human genetics. 2004; 75:204-19.
59. Stangl D, Pfohl B, Zimmerman M, Bowers W, Corenthal C. A structured interview for the DSM-III personality disorders. A preliminary report. Archives of general psychiatry. 1985; 42:591-6.
60. Hupe P, Stransky N, Thiery JP, Radvanyi F, Barillot E. Analysis of array CGH data: from signal ratio to gain and loss of DNA regions. Bioinformatics. 2004; 20:3413-22.
61. Pique-Regi R, Monso-Varona J, Ortega A, Seeger RC, Triche TJ, Asgharzadeh S. Sparse representation and Bayesian detection of genome copy number alterations from microarray data. Bioinformatics. 2008; 24:309-18.
62. Satten GA, Allen AS, Ikeda M, Mulle JG, Warren ST. Robust regression analysis of copy number variation data based on a univariate score. PloS one. 2014; 9:e86272.
63. Cuthbert BN, Insel TR. Toward new approaches to psychotic disorders: the NIMH Research Domain Criteria project. Schizophrenia bulletin. 2010; 36:1061-2.
64. Kerner B. Genetics of bipolar disorder. The application of clinical genetics. 2014; 7:33-42.
65. Disorder PGCB, Group W. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. Lancet. 2013; 381:1371-9.
66. Tesli M, Skatun KC, Ousdal OT, Brown AA, Thoresen C, Agartz I, et al. CACNA1C risk variant and amygdala activity in bipolar disorder, schizophrenia and healthy controls. PloS one. 2013; 8:e56970.

Tables and Figures

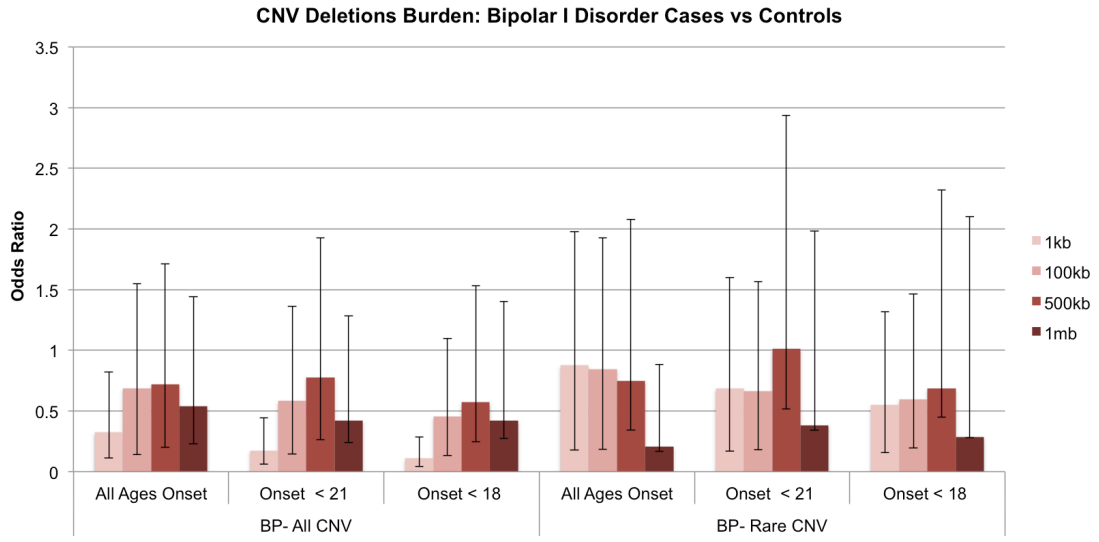

Figure 1. *Fisher's exact odds ratios for CNV deletions in Bipolar Disorder Type I cases and controls.* Fisher's exact odds ratios measure the odds of the individuals having at least one CNV in the group of cases divided by that of controls. Odds ratios are stratified by size of CNV duplication. BP-All CNV: analyses of all CNV deletions in bipolar I sample. BP- Rare CNV: analyses of CNV deletions occurring in less than 1% of bipolar I sample. Onset < 21: cases limited to onset of mania at or younger than 21 years of age. Onset < 18: cases limited to onset of mania at or younger than 18 years of age. Bars represent 95% confidence intervals.

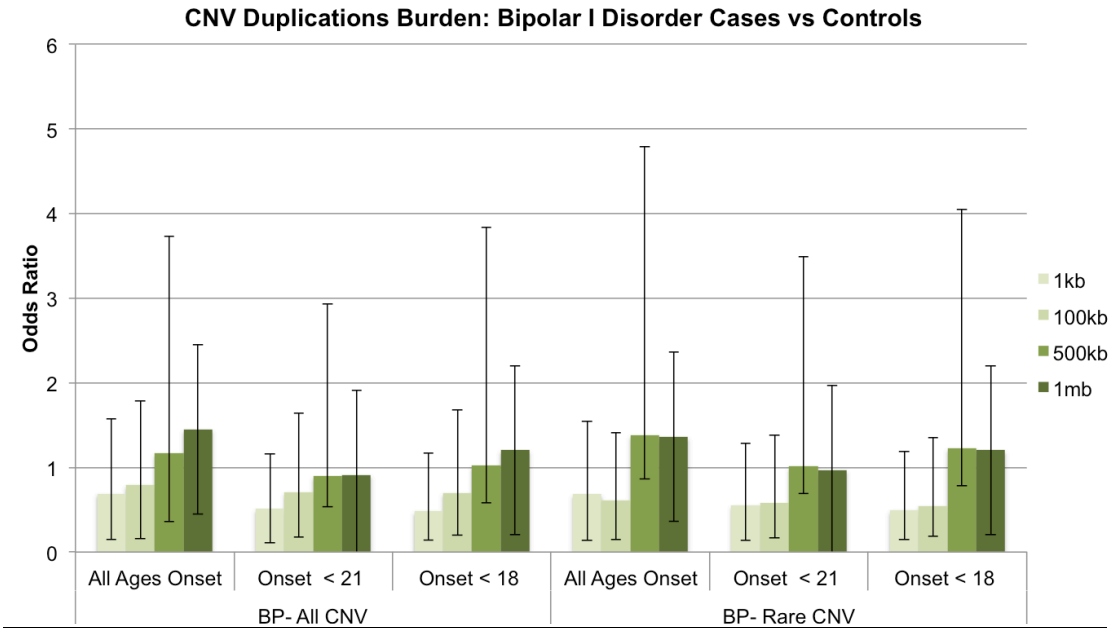

Figure 2. *Fisher's exact odds ratios for CNV duplications in Bipolar Disorder Type I cases and controls.* Fisher's exact odds ratios measure the odds of the individuals having at least one CNV in the group of cases divided by that of controls. Odds ratios are stratified by size of CNV duplication. BP-All CNV: analyses of all CNV duplications in bipolar I sample. BP- Rare CNV: analyses of CNV duplications occurring in less than 1% of bipolar I sample. Onset < 21: cases limited to onset of mania at or younger than 21 years of age. Onset < 18: cases limited to onset of mania at or younger than 18 years of age. Bars represent 95% confidence intervals.

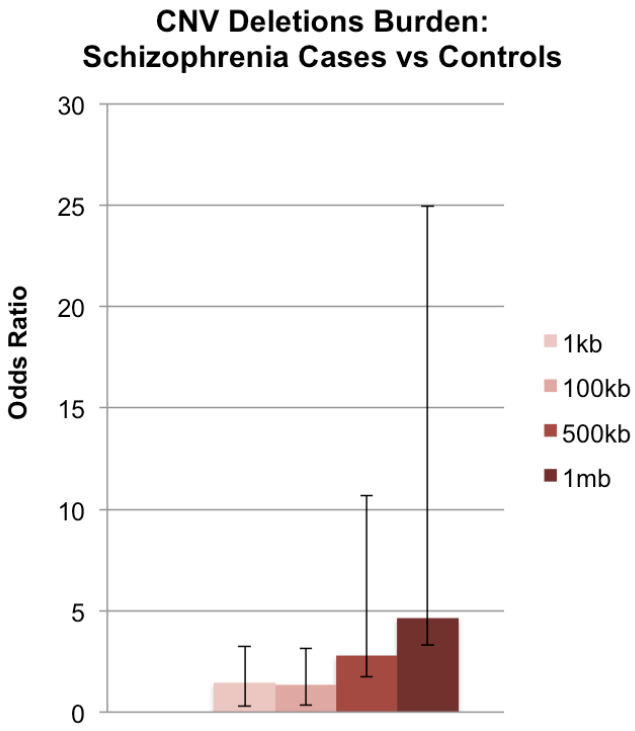

Figure 3. *Fisher's exact odds ratios for CNV deletions in Schizophrenia cases and controls.* Fisher's exact odds ratios measure the odds of the individuals having at least one CNV in the group of cases divided by that of controls. Odds ratios are stratified by size of CNV duplication. Bars represent 95% confidence intervals.

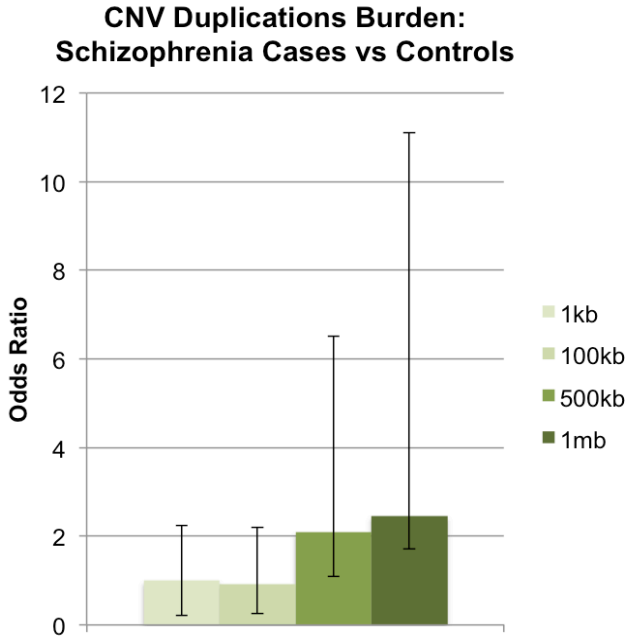

Figure 4. *Fisher's exact odds ratios for CNV duplications in Schizophrenia cases and controls.* Fisher's exact odds ratios measure the odds of the individuals having at least one CNV in the group of cases divided by that of controls. Odds ratios are stratified by size of CNV duplication. Bars represent 95% confidence intervals.

Table 1. Odds Ratios in Bipolar I Disorder, All CNV

CNV Deletions						
Age at mania onset	CNV minimum size	Cases with Event (n=479)	Controls with Event (n=1011)	Fisher's Exact Odds Ratio	95% CI	P-Value
All ages	1kb	419	966	0.3256	[0.2123, 0.4962]	P= 5.443e-08
	100kb	282	684	0.6846	[0.5434, 0.8627]	P= 0.0011
	500kb	64	178	0.7219	[0.5211, 0.9908]	P= 0.0422
	1mb	21	79	0.5411	[0.3136, 0.8981]	P= 0.0144
≤ 21 years	1kb	205	966	0.174	[0.1112, 0.2708]	P=1.33E-15
	100kb	143	684	0.5845	[0.4387, 0.7796]	P=1.91E-04
	500kb	37	178	0.7766	[0.5138, 1.1503]	P=0.2277
	1mb	9	79	0.4232	[0.1841, 0.8606]	P=0.01309
≤ 18 years	1kb	122	966	0.1097	[0.0686, 0.1744]	P< 2.2e-16
	100kb	85	684	0.4569	[0.3255, 0.6409]	P=2.89E-06
	500kb	19	178	0.5739	[0.3274, 0.9577]	P=0.02757
	1mb	6	79	0.4216	[0.1478, 0.9811]	P=0.0384
CNV Duplications						
Age at mania onset	CNV minimum size	Cases with Event (n=479)	Controls with Event (n=1011)	Fisher's Exact Odds Ratio	95% CI	P-Value
All ages	1kb	318	749	0.6911	[0.5422, 0.8819]	P= 0.0026
	100kb	235	555	0.7914	[0.6326, 0.9899]	P= 0.0397
	500kb	74	137	1.1655	[0.8453, 1.5981]	P= 0.3401
	1mb	23	34	1.449	[0.8050, 2.5658]	P= 0.1935
≤ 21 years	1kb	284	749	0.5097	[0.4022, 0.6460]	P= 1.371E-08
	100kb	120	555	0.7045	[0.5305, 0.9344]	P= 0.0122
	500kb	32	137	0.8955	[0.5735, 1.3652]	P= 0.6821
	1mb	8	34	0.9123	[0.3603, 2.0400]	P= 1.0000
≤ 18 years	1kb	101	749	0.4843	[0.3430, 0.6859]	P= 2.54E-05
	100kb	80	555	0.6995	[0.4993, 0.9781]	P= 0.0324
	500kb	24	137	1.0207	[0.6112, 1.6472]	P= 0.9050
	1mb	7	34	1.2042	[0.4431, 2.8200]	P= 0.6529

Table 2. Odds Ratios in Bipolar I Disorder- Rare CNV (present in <1% of individuals)

Rare CNV Deletions						
Age at mania onset	CNV minimum size	*Cases with Event (n=479)*	*Controls with Event (n=1011)*	*Odds Ratio*	*95% CI*	*P-Value*
All ages	1kb	267	596	0.8770	[0.7000, 1.0994]	P= 0.2612
	100kb	138	327	0.8466	[0.6619, 1.0800]	P= 0.1878
	500kb	18	50	0.7506	[0.4073, 1.3267]	P= 0.3531
	1mb	3	20	0.2083	[0.0405, 0.6751]	P= 0.0038
≤ 21 years	1kb	134	614	0.6879	[0.5180, 0.9135]	P= 0.0088
	100kb	69	357	0.6620	[0.4807, 0.9043]	P= 0.007947
	500kb	13	50	1.0116	[0.4959, 1.9272]	P= 1
	1mb	2	20	0.3843	[0.0433, 1.5988]	P= 0.2841
≤ 18 years	1kb	80	615	0.5483	[0.3909, 0.7675]	P= 3.21E-04
	100kb	43	359	0.5964	[0.4025, 0.8701]	P= 5.47E-03
	500kb	6	50	0.6866	[0.2368, 1.6374]	P= 0.5603
	1mb	1	20	0.2866	[0.0069, 1.8156]	P= 0.3457
Rare CNV Duplications						
Age at mania onset	CNV minimum size	Cases with Event (n=479)	Controls with Event (n=1011)	Odds Ratio	95% CI	P-Value
All ages	1kb	203	524	0.6838	[0.5455, 0.8562]	P= 0.0007087
	100kb	103	312	0.6139	[0.4703, 0.7974]	P= 0.0001573
	500kb	44	69	1.3806	[0.9076, 2.0822]	P= 0.1163
	1mb	9	14	1.3634	[0.5166, 3.4099]	P= 0.5022
≤ 21 years	1kb	104	554	0.5502	[0.4123, 0.7321]	P= 2.13E-05
	100kb	58	336	0.5771	[0.4112, 0.8008]	P= 6.92E-04
	500kb	19	73	1.0130	[0.5658, 1.7365]	P= 1
	1mb	6	24	0.9715	[0.3213, 2.4733]	P= 1
≤ 18 years	1kb	66	560	0.4925	[0.3480, 0.6930]	P= 2.49E-05
	100kb	38	342	0.5468	[0.3625, 0.8093]	P= 1.52E-03
	500kb	17	82	1.2265	[0.6634, 2.1545]	P= 0.4590
	1mb	7	34	1.2042	[0.4431, 2.8200]	P= 0.6529

Table 3. Odds Ratios in Schizophrenia- Rare CNV (present in <1% of individuals)

Rare CNV Deletions					
CNV minimum size	Cases with Event (n=554)	Controls with Event (n=1014)	Odds Ratio	95% CI	P-Value
1kb	216	311	1.4441	[1.1557, 1.8042]	P=0.0009649
100kb	107	152	1.3572	[1.0226, 1.7973]	P= 0.03267
500kb	12	8	2.7822	[1.0385, 7.8992]	P= 0.03134
1mb	10	4	4.6367	[1.3296, 20.3440]	P= 0.008475
Rare CNV Duplications					
CNV minimum size	Cases with Event (n=554)	Controls with Event (n=1014)	Odds Ratio	95% CI	P-Value
1kb	173	317	0.9984	[0.7928, 1.2553]	P= 1
100kb	72	141	0.9249	[0.6715, 1.2658]	P= 0.6442
500kb	18	16	2.0937	[0.9987, 4.4276]	P= 0.0442
1mb	8	6	2.46	[0.7442, 8.6482]	P= 0.0973

Table 4a. *Bipolar I Disorder- All CNV burden analysis*			**CNV Rate** No. of segments per person		**CNV Proportion** Proportion of samples with ≥1 segment		**CNV length** Total length spanned per person (kb)		**Average CNV size** Avg segment size (kb)	
			Deletions	*Duplications*	*Deletions*	*Duplications*	*Deletions*	*Duplications*	*Deletions*	*Duplications*
All ages of onset	1kb	Significance	**P=0.0039**	P=0.9860	P=1	P=0.9992	P=0.8699	**P= 0.0241**	P=0.9984	P=0.0524
		Cases	**2.303**	1.19	0.8747	0.6639	400.2	**518**	182.7	286.9
		Controls	**1.954**	1.483	0.9555	0.7409	445.7	**398.8**	229.3	255.1
	100kb	Significance	P=0.9938	P=0.7098	P=0.9996	P=0.985	P=0.9538	**P= 0.0035**	P=0.9593	**P= 0.0444**
		Cases	0.881	0.7286	0.5887	0.4906	468.3	**652.8**	311	**399.8**
		Controls	1.039	0.7587	0.6766	0.549	554.3	**470.5**	353.8	**357.7**
	500kb	Significance	P= 0.9877	P=0.0670	P=0.9858	P=0.1827	P=0.8796	P= 0.1629	P=0.9202	P= 0.2972
		Cases	0.1441	0.1879	0.1336	0.1545	1036	1295	927.9	915.7
		Controls	0.1958	0.1434	0.1761	0.1355	1197	934.1	1043	884.1
	1mb	Significance	P=0.9972	**P= 0.0467**	P=0.9962	P=0.1145	P=0.5039	P=0.2723	P=0.2441	P=0.2433
		Cases	0.04384	**0.0626**	0.04384	0.04802	1764	2431	1764	1553
		Controls	0.0821	**0.0336**	0.07814	0.03363	1776	1476	1650	1476
Onset ≤ 21 years of age	1kb	Significance	P=0.6981	P=0.9989	P=1	P=1	P=0.9630	P= 0.2247	P= 0.9949	P=0.2466
		Cases	1.881	1.015	0.7885	0.5923	362.8	425.9	177.1	270.3
		Controls	1.954	1.483	0.9555	0.7409	445.7	398.8	229.3	255.1
	100kb	Significance	P=1	P=0.9457	P= 0.9999	P=0.9958	P=0.9935	P=0.1881	P=0.9125	P=0.4003
		Cases	0.7462	0.6577	0.55	0.4615	413.3	506.2	312.9	364.1
		Controls	1.039	0.7587	0.6766	0.549	554.3	470.5	353.8	357.7
	500kb	Significance	P=0.9607	P=0.7143	P=0.9211	P=0.7320	P=0.9964	P=0.4295	P=0.9955	P=0.3918
		Cases	0.1462	0.1308	0.1423	0.1231	840.4	943.8	810.9	900.1
		Controls	0.1958	0.1434	0.1761	0.1355	1197	934.1	1043	884.1
	1mb	Significance	P=0.9984	P=0.6600	P=0.9979	P=0.6600	P=0.7643	P=0.2962	P=0.7132	P=0.2962
		Cases	0.03462	0.03077	0.03462	0.03077	1523	1540	1523	1540
		Controls	0.0821	0.03363	0.07814	0.03363	1776	1476	1650	1476
Onset ≤ 18 years of age	1kb	Significance	P=0.9308	P=0.9955	P=1	P=1	P= 0.9409	**P= 0.0387**	P=0.9922	P=0.0779
		Cases	1.724	1.023	0.7011	0.5805	358	**478.7**	171.1	295.2
		Controls	1.954	1.483	0.9555	0.7409	445.7	**398.8**	229.3	255.1
	100kb	Significance	P=1	P=0.8528	P=1	P= 0.9885	P=0.9889	**P= 0.0327**	P= 0.9366	P= 0.1696
		Cases	0.6552	0.6782	0.4885	0.4598	399.8	**565.5**	298.9	390.6
		Controls	1.039	0.7587	0.6766	0.549	554.3	**470.5**	353.8	357.7
	500kb	Significance	P= 0.9952	P=0.4553	P=0.9906	P=0.5046	P=0.9616	P=0.2737	P=0.9007	P=0.2838
		Cases	0.1092	0.1494	0.1092	0.1379	893.7	990.8	893.7	932.5
		Controls	0.1958	0.1434	0.1761	0.1355	1197	934.1	1043	884.1
	1mb	Significance	P=0.9931	P=0.3959	P=0.9922	P=0.3959	P=0.7061	P=0.3399	P=0.6864	P=0.3399
		Cases	0.03448	0.04023	0.03448	0.04023	1517	1533	1517	1533
		Controls	0.0821	0.03363	0.07814	0.03363	1776	1476	1650	1476

Table 4b. Bipolar I Disorder- Rare CNV burden analysis			CNV Rate: No. of segments per person		CNV Proportion: Proportion of samples with ≥1 segment		CNV length: Total length spanned per person (kb)		Average CNV size: Avg segment size (kb)	
			Deletions	*Duplications*	*Deletions*	*Duplications*	*Deletions*	*Duplications*	*Deletions*	*Duplications*
All ages of onset	1kb	Significance	**P= 0.0008**	P=0.9949	P=0.8927	P=0.9997	P=0.7891	P=0.0913	P=0.9932	P=0.1382
		Cases	**1.5510**	0.7641	0.5574	0.4238	287.8	426.7	120.8	209.2
		Controls	**1.1280**	1.1050	0.5895	0.5183	331.4	298.5	160.5	187.3
	100kb	Significance	P=0.9700	P=0.9539	P= 0.9233	P=1	P= 0.9441	**P= 9.99x10-5**	P= 0.9716	**P= 0.0096**
		Cases	0.3946	156.0000	0.2881	0.2150	357.5	**738.7**	234.6	**400.5**
		Controls	0.4965	410.0000	0.3234	0.3086	497.5	**387.1**	288.8	**321.6**
	500kb	Significance	P=0.9384	**P= 0.0245**	P=0.8761	P=0.0677	P=0.8620	P=0.3046	P=0.8786	P=0.6602
		Cases	0.0397	**0.1169**	0.0376	0.0919	1065	1324	881.9	771.4
		Controls	0.0623	**0.0692**	0.0495	0.0683	1542	819.5	1128	801.4
	1mb	Significance	P=0.9932	P=0.1321	P=0.9904	P=0.3023	P=0.3872	P=0.2034	P=0.1463	P=0.1326
		Cases	0.0063	0.0334	0.0063	0.0188	2695	3928	2695	1683
		Controls	0.0237	0.0139	0.0198	0.0139	2467	1444	1971	1444
Onset ≤ 21 years of age	1kb	Significance	P=0.3152	P=0.9995	P=0.9969	P=1	P=0.8907	P=0.3506	P=0.9287	P=0.2820
		Cases	1.2190	0.6731	0.5154	0.4000	254.7	334.4	133.5	217.9
		Controls	1.1530	1.1520	0.6073	0.5480	333	320	164.4	203.6
	100kb	Significance	P=0.9995	P=0.9825	P=0.9975	P=1	P=0.9274	**P= 0.0244**	P=0.6270	P=0.2235
		Cases	0.3308	0.3269	0.2654	0.2231	332.5	**519.3**	269.2	373.7
		Controls	0.5312	0.4372	0.3531	0.3323	477.6	**417.2**	283.1	344.5
	500kb	Significance	P=0.7496	P=0.4623	P=0.5365	P=0.5259	P=0.9954	P=0.5626	P=0.9876	P=0.6484
		Cases	0.0500	0.0769	0.0500	0.0731	738.5	858	738.5	821.8
		Controls	0.0623	0.0732	0.0495	0.0722	1542	880.8	1128	863.8
	1mb	Significance	P=0.9639	P=0.5918	P=0.9591	P=0.5918	P=0.879512	P=0.2477	P= 0.9282	P=0.2477
		Cases	0.0077	0.0231	0.0077	0.0231	1227	1532	1227	1532
		Controls	0.0237	0.0237	0.0198	0.0237	2467	1428	1971	1428
Onset < 18 years of age	1kb	Significance	P=0.5605	P=0.9974	P=0.9999	P=1	P=0.8811	P=0.0613	P= 0.9760	P=0.1106
		Cases	1.1320	0.6839	0.4598	0.3793	245.4	427.4	119.9	265
		Controls	1.1570	1.1620	0.6083	0.5539	333.8	345	164.6	221.8
	100kb	Significance	P=0.9997	P=0.9284	P=0.9982	P= 0.9994	P=0.9637	**P= 0.0018**	P= 0.8894	P=0.0693
		Cases	0.2931	0.3506	0.2471	0.2184	291.8	**662.9**	232.7	451.5
		Controls	0.5341	0.4471	0.3551	0.3383	476.8	**456.4**	282.6	372.1
	500kb	Significance	P=0.9130	P=0.2382	P=0.8595	P= 0.2763	P= 0.9715	P=0.5770	P=0.9575	P=0.6236
		Cases	0.0345	0.1034	0.0345	0.0977	731.4	943.2	731.4	902.7
		Controls	0.0623	0.0831	0.0495	0.0811	1542	978.4	1128	948.5
	1mb	Significance	P=0.9681	P=0.3981	P=0.9681	P=0.3981	P= 0.8541	P= 0.3394	P= 0.8541	P= 0.3394
		Cases	0.0057	0.0402	0.0057	0.0402	1197	1533	1197	1533
		Controls	0.0237	0.0336	0.0198	0.0336	2467	1476	1971	1476

Table 5. Schizophrenia Rare CNV burden analysis		CNV Rate: No. of segments per person		CNV Proportion: Proportion of samples with ≥1 segment		CNV length: Total length spanned per person (kb)		Average CNV size: Avg segment size (kb)	
		Deletions	*Duplications*	*Deletions*	*Duplications*	*Deletions*	*Duplications*	*Deletions*	*Duplications*
1kb	Significance	***P*= 9.99x10-05**	p= 0.5761	***P*= 0.0005**	p=0.5259	p= 0.0537	***P*= 0.0159**	***P*= 0.0172**	***P*= 0.0410**
	Cases	**1.1100**	0.6913	**0.3899**	0.3123	389.8	**333.8**	**219.1**	**188.4**
	Controls	**0.5759**	0.7150	**0.3067**	0.3126	268.4	**242**	**117.5**	**142**
100kb	Significance	p=0.0546	p=0.0770	***P*= 0.0192**	p= 0.7073	p= 0.1598	***P*= 0.0002**	***P*= 0.0073**	***P*= 0.0040**
	Cases	0.3069	0.2292	**0.1931**	0.1300	576.3	**620.8**	**443**	**454.3**
	Controls	0.2367	0.1775	**0.1499**	0.1391	432.9	**348.9**	**215.7**	**294.7**
500kb	Significance	p=0.1990	***P*= 0.0377**	***P*= 0.0205**	***P*= 0.0235**	p=0.474253	p=0.2060	p=0.1230	p=0.1204
	Cases	0.0217	**0.0325**	**0.0217**	**0.0325**	2758	1220	2758	1220
	Controls	0.0128	**0.0168**	**0.0079**	**0.0158**	2726	1008	1527	930.5
1mb	Significance	***P*= 0.0436**	p=0.0776	***P*= 0.0082**	p=0.0776	p=0.7177	p=0.2076	p=0.3160	p=0.2076
	Cases	**0.0181**	0.0144	**0.0181**	0.0144	3147	1807	3147	1807
	Controls	**0.0059**	0.0059	**0.0039**	0.0059	4050	1476	2560	1476