Published in final edited form as:

Schizophr Res. 2015 February ; 161(0): 490–495. doi:10.1016/j.schres.2014.11.019.

The Impact of Clinical Heterogeneity in Schizophrenia on Genomic Analyses

Sherri G. Liang and Tiffany A. Greenwood

Department of Psychiatry, University of California San Diego, La Jolla, CA

Abstract

Though clinically useful, the diagnostic systems currently employed are not well equipped to capture the substantial clinical heterogeneity observed for most psychiatric disorders, as exemplified by the complex psychotic disorder(s) that Bleuler aptly labeled the “Group of Schizophrenias”. The clinical heterogeneity associated with schizophrenia has likely frustrated decades of attempts to illuminate the underlying genetic architecture, although recent genome-wide association studies have begun to provide valuable insight into the role of common genetic risk variants. Here we demonstrate the importance of using diagnostic information to identify a core form of the disorder and to eliminate potential comorbidities in genetic studies. We also demonstrate why applying a diagnostic screening procedure to the control dataset to remove individuals with potentially related disorders is critical. Additionally, subjects may participate in multiple studies at different institutions or may have genotype data released by more than one research group. It is thus good practice to verify that no identical subjects exist within or between samples prior to conducting any type of genetic analysis to avoid potential confounding of results. While the availability of genomic data for large collections of subjects has facilitated many investigations that would otherwise not have been possible, we clearly show why one must use caution when acquiring data from publicly available sources. Although the broad vs. narrow debate in terms of phenotype definition in genetic analyses will remain, it is likely that both approaches will yield different results and that both will have utility in resolving the genetic architecture of schizophrenia.

Keywords

genome-wide association; schizophrenia; heterogeneity; diagnostics

© 2014 Elsevier B.V. All rights reserved.

Address for Correspondence: Tiffany A. Greenwood, Ph.D., Department of Psychiatry, University of California San Diego, 9500 Gilman Drive, MC 0689, La Jolla, CA 92093, Phone: 858-246-1897, tgreenwood@ucsd.edu.

Contributors: Dr. Greenwood designed the study and drafted and critically revised the manuscript. Ms. Liang performed the association analyses and provided valuable edits to the text.

Conflict of Interest: Dr. Greenwood and Ms. Liang report no financial relationships with commercial interests.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

Schizophrenia (SZ) is a severe psychiatric disorder characterized by abnormalities in a patient's thoughts, perceptions, and behaviors, manifesting as hallucinations, delusions, and/or disorganized speech with significant social or occupational dysfunction (Andreasen, 1995). The substantial clinical heterogeneity associated with SZ, which Bleuler perhaps more appropriately labeled the "Group of Schizophrenias" (Bleuler, 1911), has likely combined with the inherent genetic heterogeneity to plague many attempts at identifying casual genetic variants (Karayiorgou and Gogos, 1997; Owen et al., 2007; Sanders et al., 2008; Schork et al., 2007). Although genome-wide association studies (GWAS) of increasingly large samples have finally begun to overcome this heterogeneity to provide valuable insight into the role of common genetic variants in SZ risk (O'Donovan et al., 2008; Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium, 2011; Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014; Shi et al., 2009), investigations of smaller samples may suffer unnecessary power losses if the clinical heterogeneity is not appropriately accommodated.

Here we demonstrate the importance of using specific diagnostic criteria to identify the core features of psychiatric disorders in genetic studies. We further show why screening the control population for disorders genetically related to the disorder of interest may be critical to the success of the study. For this purpose, we use data from the subset of the Molecular Genetics of Schizophrenia (MGS) that was genotyped as part of the Genetic Association Information Network (GAIN). Finally, we emphasize that one must use caution when acquiring data from publicly available sources and always verify that no identical subjects exist within or between samples prior to conducting any type of genetic analysis to avoid the potential confounding of results.

2. Methods

All data for these analyses were obtained from the database of Genotypes and Phenotypes (dbGaP) (<http://www.ncbi.nlm.nih.gov/gap>). The Genome-Wide Association Study of Schizophrenia (SZ GAIN) was obtained via accession number **phs000021.v3.p2** and contained 1343 cases and 1368 controls of European ancestry. Cases in this sample met criteria for SZ or schizoaffective disorder (SA) per the DSM-IV (American Psychiatric Association, 2000). Controls were screened briefly and only excluded from the original sample if they endorsed a history of SZ, SA, or bipolar disorder (BP). The SZ GAIN sample was genotyped together with the Whole Genome Association Study of Bipolar Disorder (BP GAIN) sample with a common set of controls through collaboration with the Genetic Association Information Network (GAIN) for the Affymetrix Genomewide Human SNP Array 6.0 at the Broad Institute. An initial round of genotype processing and quality control was performed to remove subjects with genotype call rate <98.5% and SNPs with genotype call rates <95%, Hardy Weinberg Equilibrium p values <10⁻⁶ in the controls, and minor allele frequencies <0.01 in the combined sample, resulting in the release of the filtered subjects above and 729,454 SNPs to dbGaP. Detailed information regarding the extensive quality control process and the clinical characteristics of this sample is provided elsewhere (Sanders et al., 2008; Shi et al., 2009), although we note that the SZ GAIN was ultimately

published as a combined analysis with the complete MGS sample, rather than as an independent study (Shi et al., 2009).

Since we analyzed only a subset of the originally genotyped samples (i.e., the SZ cases and controls), we reapplied similar quality control thresholds, resulting in 724,067 SNPs for analysis. We note that initial quality control thresholds were applied to a subset that excluded 335 control subjects with recurrent major depression (MDD-R) as a primary hypothesis, consistent with the BP GAIN (Smith et al., 2009), and these subjects were later added for the evaluation of clinical heterogeneity in controls. Identity by descent (IBD) was calculated for all pairs of individuals to identify potential cryptic relatedness, and the genetic homogeneity of the sample was assured by multidimensional scaling (MDS). No related individuals or population outliers were detected. Association analyses were conducted using logistic regression, and the minimal genomic inflation identified was accommodated through the use of a genomic control correction of the p values. All analyses were performed in PLINK v.1.07 or v.109 (Purcell et al., 2007).

3. Results

3.1 Primary Analysis of the SZ GAIN Sample

The analysis of the complete SZ GAIN sample of 1343 cases and 1368 controls is provided for comparison in Figure 1A. The best result was found for rs11789399 on chromosome 9q33.1 with a p value of 3.0×10^{-6} . Another SNP, rs11789407, located just 260 bp from rs11789399 on 9q33.1 and in nearly complete linkage disequilibrium with it produced a p value of 5.3×10^{-6} . The best independent findings were for rs12745968 on chromosome 1p22 ($p=3.2 \times 10^{-6}$) and for rs271876 on chromosome 6q21 ($p=4.5 \times 10^{-6}$).

3.2 Evaluation of the Effect of Clinical Heterogeneity in the Controls

Although epidemiological and family studies indicate a genetic overlap between MDD-R and SZ (Maier et al., 1993; Somnath et al., 2002), the SZ GAIN sample includes a substantial number (N=335) of control subjects with a history of MDD-R (Sanders et al., 2008; Shi et al., 2009). The genetic relationship between MDD-R and SZ is further supported by recent data indicating a genetic correlation of 0.43, which is nearly equivalent to that seen between MDD-R and BP (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013). These data argue for the exclusion of controls with MDD-R in a GWAS of SZ. In an analysis comparing the complete set of SZ GAIN cases (N=1343) with controls screened for depression (N=1033), the best result remained at rs11789399 on 9q33.1 but with increased significance ($p=5.4 \times 10^{-7}$), as shown in Figure 1B.

3.3 Evaluation of the Effect of Clinical Heterogeneity in the Cases

Among the 1343 cases, 1191 had primary diagnoses of SZ, 57 had schizoaffective depressed type (SA-D), and 95 had schizoaffective bipolar type (SA-BP). Family studies have demonstrated the co-segregation of SA with both SZ and BP, with SA-D primarily co-segregating with SZ and SA-BP primarily co-segregating with BP (Gershon et al., 1988; Gershon et al., 1982; Kendler et al., 1993; Maier et al., 1993). Based on these prior studies, SA-D has commonly been grouped with SZ in genetic studies, and SA-BP has similarly

been considered a BP diagnosis and not generally included in genetic analyses of SZ. In fact, cases with SA-BP are included in both the BP GAIN sample and a replication sample genotyped by the same group (Smith et al., 2009; Smith et al., 2011). While the MGS study included both forms of SA, suggesting that the diagnostic criteria for differentiating SA-D and SA-BP from SZ have been shown to be somewhat subjective (Faraone et al., 1996), a recent study using dimensional rating scales demonstrated that SA-D and SA-BP are distinct diagnoses that are intermediate between BP and SZ (Keshavan et al., 2011). Thus, for an analysis of SZ, we removed all cases with either type of SA. Analysis of the 1191 cases with a primary diagnosis of SZ vs. the 1033 controls screened for depression produced a p value of 2.6×10^{-7} for rs11789399 (see Figure 1C).

Upon further inspection, we identified an additional 377 SZ cases with secondary diagnoses of mood disorders, including 149 cases with BP Not Otherwise Specified (BP-NOS), 10 cases with MDD-R, and 218 cases with Depressive Disorder Not Otherwise Specified (DD-NOS), all of which represent clinical heterogeneity in terms of co-morbid affective symptoms. Removal of these cases for an analysis of the 814 subjects with core SZ features and no history of prominent affective symptoms vs. the 1033 controls screened for depression resulted in genome-wide significant p values for both rs11789399 and rs11789407 of 1.2×10^{-8} and 2.1×10^{-8} , respectively, as shown in Figure 1D. The genomic region surrounding these SNPs is shown in Figure 2, and Table 2 presents a summary of the results for the SNPs in this region with p values $< 10^{-4}$. An additional seven SNPs with p values $< 10^{-4}$ spanning a total of 37.5 kb surrounding rs11789399 provide further evidence to support association to this region and also show a pattern of increasing significance with increasing stringency of diagnostic criteria. Given the high degree of linkage disequilibrium in the immediate region, we performed an analysis conditional on both rs11789399 and rs11789407. No other SNP in the 9q33.1 region remained significant, suggesting that the association signal is derived primarily through linkage disequilibrium with rs11789399.

4. Discussion

In this study, the most strongly associated SNP was rs11789399 on chromosome 9q33.1 with an initial p value of 3.0×10^{-6} that increased to 1.2×10^{-8} through the use of successively stringent diagnostic criteria for SZ. While we note that these are nested analyses and are nowhere near independent, the p value for rs11789399 remains genome-wide significant at 4.8×10^{-8} when corrected for multiple testing.

The associated region in this study is intergenic but directly adjacent to a cluster of eight human spliced ESTs in a very intriguing genomic region (see Figure 2). One of the flanking genes, toll-like receptor 4 precursor gene (*TLR4*), plays a fundamental role in pathogen recognition and the activation of innate immunity (Piccinini and Midwood, 2010; Takeuchi and Akira, 2010). *TLR4* is expressed in the uterus, cervix, and placenta (Gonzalez et al., 2007; Patni et al., 2009), as well as the brain, where it has been shown to play a significant role in both neurodevelopment and neuroplasticity (Akira et al., 2001; Bsibsi et al., 2002; Kaul et al., 2012; Larsen et al., 2007; Okun et al., 2011). It has been proposed that *TLR4* mediates increased SZ risk in response to maternal infection, given the above observations and evidence suggesting that immuno-inflammatory reactions during neurodevelopment

contribute to SZ (Ganguli et al., 1994; Muller et al., 1999; Venkatasubramanian and Debnath, 2013). The other flanking gene, deleted in bladder cancer 1 (*DBC1*, formerly *BRINP1*), suppresses cell cycle progression in neural stem cells and is highly expressed in the brain (Kawano et al., 2004). The absence of *BRINP1* in mice causes deregulation of neurogenesis and impairments of neuronal differentiation in adult hippocampal circuitry, resulting in hyperactivity and poor social behavior (Kobayashi et al., 2014). Furthermore, linkage to a nearby marker in the 9q33.1 region was also observed for genetic modifiers of the clinical course of SZ (Fanous et al., 2007). Finally, a 262 kb deletion of 9q33.1 was detected by whole genome bacterial artificial chromosome (BAC) array in a patient with developmental delay (Neill et al., 2010). While the boundaries of this deletion technically include only *TLR4*, gaps in BAC coverage on both sides of the alteration span 4.5 Mb proximally and 4.0 Mb distally. As such, this alteration may be as large as 8.7 Mb and include the associated region in this study, as well as astrotactin 2 (*ASTN2*), which has been found to be deleted/duplicated in SZ and to predict metabolic side effects to antipsychotic medications (Adkins et al., 2011; Vrijenhoek et al., 2008). Thus, several independent lines of evidence offer support for this genomic region in SZ susceptibility.

The two other regions of potential interest identified in the initial analysis of the complete sample did not show this same pattern of increasing significance with successive stringency of diagnostic criteria. The region on chromosome 1p22 gave an initial p value of 3.1×10^{-6} for a SNP in the *FAM69A* gene, which belongs to a family of cysteine-rich type II transmembrane proteins of unknown function. Although support remained for this large gene, the peak significance decreased to $<10^{-4}$ and was observed for a different SNP when MDD-R was eliminated from the controls, with a further decrease to $<10^{-3}$ observed with successive stringency of criteria. The peak on this chromosome shifted to 1p31 in the final analysis of the core features of SZ with a p value of 1.1×10^{-6} observed for an intergenic region. The region on chromosome 6q21 gave an initial p value of 3.1×10^{-6} , which decreased to p values in the range of approximately 10^{-3} – 10^{-4} in the subsequent analyses. This region is intergenic but near *GRIK2*, which has been implicated in treatment-emergent suicide associated with antidepressants and appeared as a prominent finding in a recent large study of SZ, which included the MGS sample described here without removal of controls with MDD-R or SZ cases with SA or co-morbid affective diagnoses (Laje et al., 2007; Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014). It is thus likely that both regions implicated initially, particularly 6q21, relate more to depression, or possibly affective symptoms in SZ, than to the core features of SZ.

While SZ and BP are generally considered separate disorders by current diagnostic criteria, emerging data has challenged this view with increasingly strong evidence for a partially shared etiology and genetic risk, with some investigators suggesting that these disorders exist on opposite ends of a functional psychosis continuum, or Schizo-Bipolar spectrum, in which SA-BP lies intermediate between the disorders, while SA-D lies somewhat closer to SZ (Craddock and Owen, 2005; Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013; Crow, 1990; Ivleva et al., 2008; Kendler et al., 1998; Keshavan et al., 2011; Lichtenstein et al., 2009). Since the SZ GAIN sample includes cases with SZ and both subtypes of SA, it may be well-suited to investigations of the Schizo-Bipolar spectrum,

although one would also likely want to include BP in that case to capture the full spectrum, not just some portion of it. In fact, strategies to combine these disorders in genetic studies to identify genes that influence risk across the spectrum have proven successful in large samples (Ruderfer et al., 2013). However, for an investigation of SZ, subjects with SA-BP should be removed, and those with SA-D likely just increase the clinical heterogeneity and further dilute the genetic signal. Additionally, SZ with vs. without co-morbid affective symptoms may represent very different underlying genetic and biological causes, as was recently suggested by a large linkage study (Vieland et al., 2014). Therefore, using more specific phenotype definitions of the core features of SZ to create more genetically homogenous subgroups of patients may increase power to detect a signal, as our data clearly demonstrate. This may appear to conflict with the recent institution of Research Domain Criteria (RDoC), which challenges investigators to develop alternative classifications of psychopathology based on behavioral phenotypes that transcend traditional diagnostic boundaries (Cuthbert and Insel, 2010; Insel et al., 2010; Insel and Cuthbert, 2009). However, the use of biological phenotypes, such as endophenotypes, as objective measurements related to specific neurological functions also aim to reduce heterogeneity, thereby facilitating the detection of SZ risk variants (Braff and Freedman, 2002; Gottesman and Gould, 2003). These are simply two different strategies that accomplish the same goal.

We have also shown why screening the control dataset for disorders related to the disorder of interest may be critical to the success of the study. While one could argue that the inclusion of controls with MDD-R in a SZ study provides a control for depressive symptoms in SZ to give a better estimation of the genetic contribution to psychotic symptoms, depression in otherwise healthy individuals may be the result of entirely different genetic mechanisms than depressive symptoms in SZ. In this case, the inclusion of controls with MDD-R would be expected to dilute the genetic signal, as we have found. Additionally, many would argue that controls with MDD-R are simply not appropriate for inclusion in a genetic study of SZ, given the observed genetic correlation between SZ and MDD-R of 0.43, which is nearly equivalent to the 0.47 genetic correlation observed between BP and MDD-R (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013). In fact, the BP GAIN excluded these controls from their analyses, given the epidemiological data indicating a genetic overlap of MDD-R with BP (Gershon et al., 1982; Maier et al., 1993; Smith et al., 2009; Weissman et al., 1984; Winokur et al., 1982). Thus, in light of epidemiological and family studies, as well as more recent studies of unrelated individuals indicating a substantial genetic correlation between MDD-R and SZ (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013; Maier et al., 1993; Somnath et al., 2002), it would seem prudent to exclude controls with MDD-R from genetic analyses of SZ as well.

It should be noted that in a study by Wang and colleagues, rs11789399 reached genome-wide significance in a meta-analysis of the SZ and BP GAIN samples with p values of 2.4×10^{-6} , 5.7×10^{-4} , and 5.6×10^{-9} observed for SZ, BP, and the meta-analysis, respectively (Wang et al., 2010). However, the BP GAIN and SZ GAIN samples share a common set of controls, and this study made no apparent correction for the complete overlap of 1033 subjects between these two datasets in the meta-analysis. Therefore, the p value is likely inflated. This clearly demonstrates why it is a good general practice to verify the absence of overlapping subjects within and between all datasets obtained from public sources prior to

conducting any type of genetic analysis to avoid potential confounding of results. Subjects have been known to participate in multiple studies across time or even in studies at different institutions, and relatives may also participate in the same study without disclosing this fact to the investigators. While the duplication of subjects or unintentional recruitment of relatives within and among large consortia may not be preventable, such subjects are easily identified through measures of genomic similarity (Shi et al., 2009; Smith et al., 2011).

This study thus demonstrates the importance of more specific phenotype definitions for SZ, which have broader implications for genetic studies of psychiatric illness in general. While such strategies may not produce identical results in every dataset, we illustrate how an association to a very interesting genomic region was present in the SZ GAIN sample initially but may have gone unnoticed because the p value did not reach genome-wide significance in the original sample. Through the use of successively stringent diagnostic criteria, we were able to bring this finding to a level of significance worthy of note. This may therefore be a useful and alternative strategy to combining subjects across portions of the Schizo-Bipolar spectrum to identify genes specific to the core features of SZ, which will be critical for understanding the complex etiology of the disorder. A recent study by Manchia and colleagues further demonstrated the impact of phenotypic heterogeneity on genetic studies of other complex diseases and strikingly concluded that the use of accurate phenotype definitions may be even more important for detecting true genetic associations than increased sample sizes (Manchia et al., 2013). While extremely large-scale efforts may be able to overcome this clinical heterogeneity to identify variants that increase risk for SZ specifically and psychiatric illness in general (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014), strategies focused on specific phenotype definitions may provide important mechanistic insight into the illness and make better use of smaller, extensively characterized samples (Kohane, 2014). The availability of genomic data for large collections of subjects has facilitated many investigations, such as these, that would otherwise never have been possible; yet, one must use caution when acquiring data from publicly available sources to ensure appropriate hypothesis testing.

Acknowledgments

Role of Funding Source: Other than proving support, the NIH had no further role in this manuscript.

This study was supported by grant K01-MH087889 from the National Institute of Mental Health. The authors thank Drs. David Braff and John Kelsoe for their expert advice and guidance in the preparation of this manuscript.

Funding support for the Genome-Wide Association of Schizophrenia Study was provided by the National Institute of Mental Health (R01 MH67257, R01 MH59588, R01 MH59571, R01 MH59565, R01 MH59587, R01 MH60870, R01 MH59566, R01 MH59586, R01 MH61675, R01 MH60879, R01 MH81800, U01 MH46276, U01 MH46289, U01 MH46318, U01 MH79469, and U01 MH79470) and the genotyping of samples was provided through the Genetic Association Information Network (GAIN). The datasets used for the analyses described in this manuscript were obtained from the database of Genotypes and Phenotypes (dbGaP) found at <http://www.ncbi.nlm.nih.gov/gap> through dbGaP accession number phs000021.v3.p2. Samples and associated phenotype data for the Genome-Wide Association of Schizophrenia Study were provided by the Molecular Genetics of Schizophrenia Collaboration (PI: Pablo V. Gejman, Evanston Northwestern Healthcare (ENH) and Northwestern University, Evanston, IL, USA).

References

- Adkins DE, Aberg K, McClay JL, Bukszar J, Zhao Z, Jia P, Stroup TS, Perkins D, McEvoy JP, Lieberman JA, Sullivan PF, van den Oord EJ. Genomewide pharmacogenomic study of metabolic side effects to antipsychotic drugs. *Mol Psychiatry*. 2011; 16(3):321–332. [PubMed: 20195266]
- Akira S, Takeda K, Kaisho T. Toll-like receptors: critical proteins linking innate and acquired immunity. *Nat Immunol*. 2001; 2(8):675–680. [PubMed: 11477402]
- American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders 4-Text Revision edition. Washington DC: 2000.
- Andreasen NC. Symptoms, signs, and diagnosis of schizophrenia. *Lancet*. 1995; 346(8973):477–481. [PubMed: 7637483]
- Bleuler, E. *Dementia Praecox oder Gruppe der Schizophrenien*. Deuticke; Leipzig: 1911.
- Braff, DL.; Freedman, R. Endophenotypes in studies of the genetics of schizophrenia. In: Davis, KL.; Charney, D.; Coyle, JT.; Nemeroff, C., editors. *Neuropsychopharmacology: The Fifth Generation of Progress: An Official Publication of the American College of Neuropsychopharmacology*. Lippincott Williams & Wilkins; Philadelphia: 2002.
- Bsibsi M, Ravid R, Gveric D, van Noort JM. Broad expression of Toll-like receptors in the human central nervous system. 2002; 61(11):1013–1021.
- Craddock N, Owen MJ. The beginning of the end for the Kraepelinian dichotomy. *Br J Psychiatry*. 2005; 186:364–366. [PubMed: 15863738]
- Cross-Disorder Group of the Psychiatric Genomics Consortium. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet*. 2013; 45 (9):984–994. [PubMed: 23933821]
- Crow TJ. The continuum of psychosis and its genetic origins. The sixty-fifth Maudsley lecture. *Br J Psychiatry*. 1990; 156:788–797. [PubMed: 2207509]
- Cuthbert BN, Insel TR. Toward new approaches to psychotic disorders: the NIMH Research Domain Criteria project. *Schizophr Bull*. 2010; 36 (6):1061–1062. [PubMed: 20929969]
- Fanous AH, Neale MC, Webb BT, Straub RE, Amdur RL, O'Neill FA, Walsh D, Riley BP, Kendler KS. A genome-wide scan for modifier loci in schizophrenia. *Am J Med Genet B Neuropsychiatr Genet*. 2007; 144B (5):589–595. [PubMed: 17262803]
- Faraone SV, Blehar M, Pepple J, Moldin SO, Norton J, Nurnberger JI, Malaspina D, Kaufmann CA, Reich T, Cloninger CR, DePaulo JR, Berg K, Gershon ES, Kirch DG, Tsuang MT. Diagnostic accuracy and confusability analyses: an application to the Diagnostic Interview for Genetic Studies. *Psychol Med*. 1996; 26 (2):401–410. [PubMed: 8685296]
- Ganguli R, Brar JS, Rabin BS. Immune abnormalities in schizophrenia: evidence for the autoimmune hypothesis. *Harv Rev Psychiatry*. 1994; 2 (2):70–83. [PubMed: 9384885]
- Gershon ES, DeLisi LE, Hamovit J, Nurnberger JI Jr, Maxwell ME, Schreiber J, Dauphinais D, Dingman CW 2nd, Guroff JJ. A controlled family study of chronic psychoses. Schizophrenia and schizoaffective disorder. *Arch Gen Psychiatry*. 1988; 45 (4):328–336. [PubMed: 3355320]
- Gershon ES, Hamovit J, Guroff JJ, Dibble E, Leckman JF, Sceery W, Targum SD, Nurnberger JI Jr, Goldin LR, Bunney WE Jr. A family study of schizoaffective, bipolar I, bipolar II, unipolar, and normal control probands. *Arch Gen Psychiatry*. 1982; 39 (10):1157–1167. [PubMed: 7125846]
- Gonzalez JM, Xu H, Ofori E, Elovitz MA. Toll-like receptors in the uterus, cervix, and placenta: is pregnancy an immunosuppressed state? *Am J Obstet Gynecol*. 2007; 197 (3):296.e1–6. [PubMed: 17826427]
- Gottesman II, Gould TD. The endophenotype concept in psychiatry: etymology and strategic intentions. *Am J Psychiatry*. 2003; 160 (4):636–645. [PubMed: 12668349]
- Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, Sanislow C, Wang P. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *Am J Psychiatry*. 2010; 167 (7):748–751. [PubMed: 20595427]
- Insel TR, Cuthbert BN. Endophenotypes: bridging genomic complexity and disorder heterogeneity. *Biol Psychiatry*. 2009; 66 (11):988–989. [PubMed: 19900610]

- Ivleva E, Thaker G, Tamminga CA. Comparing genes and phenomenology in the major psychoses: schizophrenia and bipolar 1 disorder. *Schizophr Bull.* 2008; 34(4):734–742. [PubMed: 18515820]
- Karayiorougou M, Gogos JA. A turning point in schizophrenia genetics. *Neuron.* 1997; 19(5):967–979. [PubMed: 9390512]
- Kaul D, Habbel P, Derkow K, Kruger C, Franzoni E, Wulczyn FG, Bereswill S, Nitsch R, Schott E, Veh R, Naumann T, Lehnardt S. Expression of Toll-like receptors in the developing brain. *PLoS One.* 2012; 7(5):e37767. [PubMed: 22666391]
- Kawano H, Nakatani T, Mori T, Ueno S, Fukaya M, Abe A, Kobayashi M, Toda F, Watanabe M, Matsuoka I. Identification and characterization of novel developmentally regulated neural-specific proteins, BRINP family. *Brain Res Mol Brain Res.* 2004; 125(1–2):60–75. [PubMed: 15193423]
- Kendler KS, Karkowski LM, Walsh D. The structure of psychosis: latent class analysis of probands from the Roscommon Family Study. *Arch Gen Psychiatry.* 1998; 55(6):492–499. [PubMed: 9633666]
- Kendler KS, McGuire M, Gruenberg AM, O'Hare A, Spellman M, Walsh D. The Roscommon Family Study. IV. Affective illness, anxiety disorders, and alcoholism in relatives. *Arch Gen Psychiatry.* 1993; 50(12):952–960. [PubMed: 8250681]
- Keshavan MS, Morris DW, Sweeney JA, Pearlson G, Thaker G, Seidman LJ, Eack SM, Tamminga C. A dimensional approach to the psychosis spectrum between bipolar disorder and schizophrenia: the Schizo-Bipolar Scale. *Schizophr Res.* 2011; 133(1–3):250–254. [PubMed: 21996268]
- Kobayashi M, Nakatani T, Koda T, Matsumoto K, Ozaki R, Mochida N, Takao K, Miyakawa T, Matsuoka I. Absence of BRINP1 in mice causes increase of hippocampal neurogenesis and behavioral alterations relevant to human psychiatric disorders. *Mol Brain.* 2014; 7:12. [PubMed: 24528488]
- Kohane IS. Deeper, longer phenotyping to accelerate the discovery of the genetic architectures of diseases. *Genome Biol.* 2014; 15:115. [PubMed: 25165795]
- Laje G, Paddock S, Manji H, Rush AJ, Wilson AF, Charney D, McMahon FJ. Genetic markers of suicidal ideation emerging during citalopram treatment of major depression. *Am J Psychiatry.* 2007; 164 (10):1530–1538. [PubMed: 17898344]
- Larsen PH, Holm TH, Owens T. Toll-like receptors in brain development and homeostasis. *Sci STKE.* 2007; 2007(402):pe47. [PubMed: 17785714]
- Lichtenstein P, Yip BH, Bjork C, Pawitan Y, Cannon TD, Sullivan PF, Hultman CM. Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet.* 2009; 373(9659):234–239. [PubMed: 19150704]
- Maier W, Lichtermann D, Minges J, Hallmayer J, Heun R, Benkert O, Levinson DF. Continuity and discontinuity of affective disorders and schizophrenia. Results of a controlled family study. *Arch Gen Psychiatry.* 1993; 50(11):871–883. [PubMed: 8215813]
- Manchia M, Cullis J, Turecki G, Rouleau GA, Uher R, Alda M. The impact of phenotypic and genetic heterogeneity on results of genome wide association studies of complex diseases. *PLoS One.* 2013; 8 (10):e76295. [PubMed: 24146854]
- Muller N, Riedel M, Ackenheil M, Schwarz MJ. The role of immune function in schizophrenia: an overview. *Eur Arch Psychiatry Clin Neurosci.* 1999; 249(Suppl 4):62–68. [PubMed: 10654111]
- Neill NJ, Torchia BS, Bejjani BA, Shaffer LG, Ballif BC. Comparative analysis of copy number detection by whole-genome BAC and oligonucleotide array CGH. *Mol Cytogenet.* 2010; 3:11. [PubMed: 20587050]
- O'Donovan MC, Craddock N, Norton N, Williams H, Peirce T, Moskva V, Nikolov I, Hamshere M, Carroll L, Georgieva L, Dwyer S, Holmans P, Marchini JL, Spencer CC, Howie B, Leung HT, Hartmann AM, Moller HJ, Morris DW, Shi Y, Feng G, Hoffmann P, Propping P, Vasilescu C, Maier W, Rietschel M, Zammit S, Schumacher J, Quinn EM, Schulze TG, Williams NM, Giegling I, Iwata N, Ikeda M, Darvasi A, Shifman S, He L, Duan J, Sanders AR, Levinson DF, Gejman PV, Cichon S, Nothen MM, Gill M, Corvin A, Rujescu D, Kirov G, Owen MJ, Buccola NG, Mowry BJ, Freedman R, Amin F, Black DW, Silverman JM, Byerley WF, Cloninger CR. Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nat Genet.* 2008; 40(9):1053–1055. [PubMed: 18677311]

- Okun E, Griffioen KJ, Mattson MP. Toll-like receptor signaling in neural plasticity and disease. *Trends Neurosci.* 2011; 34(5):269–281. [PubMed: 21419501]
- Owen MJ, Craddock N, Jablensky A. The genetic deconstruction of psychosis. *Schizophr Bull.* 2007; 33(4):905–911. [PubMed: 17551090]
- Patni S, Wynen LP, Seager AL, Morgan G, White JO, Thornton CA. Expression and activity of Toll-like receptors 1–9 in the human term placenta and changes associated with labor at term. *Biol Reprod.* 2009; 80(2):243–248. [PubMed: 18815357]
- Piccinini AM, Midwood KS. DAMPening inflammation by modulating TLR signalling. *Mediators Inflamm.* 2010; 2010
- Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, Boehnke M, Abecasis GR, Willer CJ. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics.* 2010; 26(18):2336–2337. [PubMed: 20634204]
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am J Hum Genet.* 2007; 81(3):559–575. [PubMed: 17701901]
- Ruderfer DM, Fanous AH, Ripke S, McQuillin A, Amdur RL, Gejman PV, O'Donovan MC, Andreassen OA, Djurovic S, Hultman CM, Kelsoe JR, Jamain S, Landen M, Leboyer M, Nimgaonkar V, Nurnberger J, Smoller JW, Craddock N, Corvin A, Sullivan PF, Holmans P, Sklar P, Kendler KS. Polygenic dissection of diagnosis and clinical dimensions of bipolar disorder and schizophrenia. *Mol Psychiatry.* 2013
- Sanders AR, Duan J, Levinson DF, Shi J, He D, Hou C, Burrell GJ, Rice JP, Nertney DA, Olincy A, Rozić P, Vinogradov S, Buccola NG, Mowry BJ, Freedman R, Amin F, Black DW, Silverman JM, Byerley WF, Crowe RR, Cloninger CR, Martinez M, Gejman PV. No significant association of 14 candidate genes with schizophrenia in a large European ancestry sample: implications for psychiatric genetics. *Am J Psychiatry.* 2008; 165(4):497–506. [PubMed: 18198266]
- Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium. Genome-wide association study identifies five new schizophrenia loci. *Nat Genet.* 2011; 43(10):969–976. [PubMed: 21926974]
- Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature.* 2014; 511:421–427. [PubMed: 25056061]
- Schork NJ, Greenwood TA, Braff DL. Statistical genetics concepts and approaches in schizophrenia and related neuropsychiatric research. *Schizophr Bull.* 2007; 33(1):95–104. [PubMed: 17035359]
- Shi J, Levinson DF, Duan J, Sanders AR, Zheng Y, Pe'er I, Dudbridge F, Holmans PA, Whittemore AS, Mowry BJ, Olincy A, Amin F, Cloninger CR, Silverman JM, Buccola NG, Byerley WF, Black DW, Crowe RR, Oksenberg JR, Mirel DB, Kendler KS, Freedman R, Gejman PV. Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature.* 2009; 460(7256):753–757. [PubMed: 19571809]
- Smith EN, Bloss CS, Badner JA, Barrett T, Belmonte PL, Berrettini W, Byerley W, Coryell W, Craig D, Edenberg HJ, Eskin E, Foroud T, Gershon E, Greenwood TA, Hipolito M, Koller DL, Lawson WB, Liu C, Lohoff F, McInnis MG, McMahon FJ, Mirel DB, Murray SS, Nievergelt C, Nurnberger J, Nwulia EA, Paschall J, Potash JB, Rice J, Schulze TG, Scheftner W, Panganiban C, Zaitlen N, Zandi PP, Zollner S, Schork NJ, Kelsoe JR. Genome-wide association study of bipolar disorder in European American and African American individuals. *Mol Psychiatry.* 2009; 14(8):755–763. [PubMed: 19488044]
- Smith EN, Koller DL, Panganiban C, Szelinger S, Zhang P, Badner JA, Barrett TB, Berrettini WH, Bloss CS, Byerley W, Coryell W, Edenberg HJ, Foroud T, Gershon ES, Greenwood TA, Guo Y, Hipolito M, Keating BJ, Lawson WB, Liu C, Mahon PB, McInnis MG, McMahon FJ, McKinney R, Murray SS, Nievergelt CM, Nurnberger JI Jr, Nwulia EA, Potash JB, Rice J, Schulze TG, Scheftner WA, Shilling PD, Zandi PP, Zollner S, Craig DW, Schork NJ, Kelsoe JR. Genome-Wide Association of Bipolar Disorder Suggests an Enrichment of Replicable Associations in Regions near Genes. *PLoS Genet.* 2011; 7(6):e1002134. [PubMed: 21738484]
- Somnath CP, Janardhan Reddy YC, Jain S. Is there a familial overlap between schizophrenia and bipolar disorder? *J Affect Disord.* 2002; 72(3):243–247. [PubMed: 12450641]
- Takeuchi O, Akira S. Schizophrenia Working Group of the Psychiatric Genomics C, Bipolar Disorder Working Group of the Psychiatric Genomics C, Cross-Disorder Working Group of the Psychiatric

- Genomics C. Pattern recognition receptors and inflammation. *Cell*. 2010; 140(6):805–820. [PubMed: 20303872]
- Venkatasubramanian G, Debnath M. The TRIPS (Toll-like receptors in immuno-inflammatory pathogenesis) Hypothesis: a novel postulate to understand schizophrenia. *Prog Neuropsychopharmacol Biol Psychiatry*. 2013; 44:301–311. [PubMed: 23587629]
- Vieland VJ, Walters KA, Lehner T, Azaro M, Tobin K, Huang Y, Brzustowicz LM. Revisiting schizophrenia linkage data in the NIMH Repository: reanalysis of regularized data across multiple studies. *Am J Psychiatry*. 2014; 171 (3):350–359. [PubMed: 24170318]
- Vrijenhoek T, Buizer-Voskamp JE, van der Stelt I, Strengman E, Sabatti C, Geurts van Kessel A, Brunner HG, Ophoff RA, Veltman JA. Recurrent CNVs disrupt three candidate genes in schizophrenia patients. *Am J Hum Genet*. 2008; 83(4):504–510. [PubMed: 18940311]
- Wang KS, Liu XF, Aragam N. A genome-wide meta-analysis identifies novel loci associated with schizophrenia and bipolar disorder. *Schizophr Res*. 2010; 124(1–3):192–199. [PubMed: 20889312]
- Weissman MM, Gershon ES, Kidd KK, Prusoff BA, Leckman JF, Dibble E, Hamovit J, Thompson WD, Pauls DL, Guroff JJ. Psychiatric disorders in the relatives of probands with affective disorders. The Yale University--National Institute of Mental Health Collaborative Study. *Arch Gen Psychiatry*. 1984; 41(1):13–21. [PubMed: 6691780]
- Winokur G, Tsuang MT, Crowe RR. The Iowa 500: affective disorder in relatives of manic and depressed patients. *Am J Psychiatry*. 1982; 139(2):209–212. [PubMed: 7055292]

Figure 1.

Genome-wide association results for the SZ GAIN sample demonstrating the effects of clinical heterogeneity in both cases and controls. (A) All cases (N=1343) vs. all controls (N=1368) following genotyping quality control procedures only. (B) All cases (N=1343) vs. controls screened for depression (N=1033). (C) Cases screened for SA (N=1191) vs. controls screened for depression (N=1033). (D) Cases screened for SA and secondary mood diagnoses (see Table 1, N=814) vs. controls screened for depression (N=1033). In each panel, a blue line indicates $p < 10^{-4}$, a red line indicates genome-wide significance at $p = 5 \times 10^{-8}$, and an arrow indicates the peak SNP, rs11789399.

Figure 2.

Details of the 9q33 region from LocusZoom (Pruim et al., 2010), showing the chromosomal context, linkage disequilibrium structure, and patterns of recombination surrounding the peak SNP, rs11789399, which is indicated as a purple diamond. All locations are based on the hg 18 assembly, and linkage disequilibrium patterns across the region are shown according to the CEPH reference population from the HapMap release 22 with red indicating complete disequilibrium ($D'=1$). The association results shown correspond to the analysis of SZ cases with SA and secondary mood diagnoses removed vs. controls screened for depression (see Figure 1D), with a peak p value of 6.7×10^{-9} observed for rs11789399. Nearby genes and two spliced ESTs (representative of a cluster of eight total ESTs) are also shown, as are the locations of a linkage microsatellite marker (D9S934; Fanous et al., 2007) and a BAC deletion at 119,452,279–119,714,054 bp (Neill et al., 2010). Although this deletion is estimated to be 262 kb, due to gaps in coverage on both sides it may actually be as large as 8.7 Mb and include the entire region shown.

Table 1

Clinical characteristics of the SZ GAIN sample

	DSM-IV Code(s)	N
Primary Diagnoses		
SZ	295	814
SZ + secondary mood diagnosis ^a	295 + (see below)	377
SA-D	295.70D	57
SA-BP	295.70M	95
Total		1343
Secondary Mood Diagnoses^a		
BP-NOS	296.8	149
MDD-R	296.3	10
DD-NOS	311	218
Total		377

Key: SZ = Schizophrenia; SA-D = Schizoaffective, Depressed type; SA-BP = Schizoaffective, Bipolar-type; BP-NOS = Bipolar Disorder Not Otherwise Specified; MDD-R = Major Depressive Disorder, Recurrent; DD-NOS = Depressive Disorder Not Otherwise Specified.

^aMood diagnoses secondary to the primary SZ diagnosis representing clinical heterogeneity in terms of co-morbid affective symptoms.

Table 2

Comparison of association results for the chromosome 9q33.1 region across analyses using successively stringent diagnostic criteria.

SNP	Location (bp)	A1	A2	MAF SZ	MAF CTL	(A) All cases (N=1343) vs. All controls (N=1368)			(B) All cases (N=1343) vs. Controls screened for MDD-R (N=1033)			(C) Cases screened for SA (N=1191) vs. Controls screened for MDD-R (N=1033)			(D) Cases screened for SA and secondary mood diagnoses (N=814) vs. Controls screened for MDD-R (N=1033)		
						OR	P		OR	P		OR	P		OR	P	
rs1335258	120367144	C	A	0.39	0.32	1.27	4.5×10^{-5}		1.33	7.8×10^{-6}		1.33	1.2×10^{-5}		1.43	5.4×10^{-7}	
rs7874294	120372743	T	G	0.38	0.32	1.25	1.5×10^{-4}		1.31	2.4×10^{-5}		1.31	4.2×10^{-5}		1.41	1.5×10^{-6}	
rs11789399	120399107	G	A	0.52	0.44	1.30	3.0×10^{-6}		1.36	5.4×10^{-7}		1.38	2.6×10^{-7}		1.47	1.2×10^{-8}	
rs11789407	120399367	C	A	0.52	0.44	1.29	5.3×10^{-6}		1.35	7.2×10^{-7}		1.37	3.5×10^{-7}		1.47	2.1×10^{-8}	
rs10759986	120403249	A	C	0.42	0.35	1.26	6.4×10^{-5}		1.31	1.7×10^{-5}		1.31	2.9×10^{-5}		1.42	5.2×10^{-7}	
rs10733631	120403878	A	C	0.42	0.35	1.27	4.7×10^{-5}		1.32	1.1×10^{-5}		1.31	2.4×10^{-5}		1.42	4.3×10^{-7}	
rs10759987	120403955	A	C	0.42	0.35	1.26	7.8×10^{-5}		1.31	1.9×10^{-5}		1.31	3.3×10^{-5}		1.42	5.5×10^{-7}	
rs7866602	120404624	A	C	0.42	0.35	1.25	1.1×10^{-4}		1.30	2.9×10^{-5}		1.30	5.8×10^{-5}		1.41	1.1×10^{-6}	

Key: A1 = minor allele; A2 = major allele; MAF = minor allele frequency; OR = odds ratio; P = association p value. Panels of association results correspond to those shown in Figure 1. Note that the MAFs are presented for the analyses reflected in A, B, and C, which were identical. The MAFs in SZ subjects were only slightly higher (by 0.01–0.02) for the analysis shown in D.