Schizophrenia Bulletin vol. 42 no. 2 pp. 279–287, 2016
doi:10.1093/schbul/sbv119
Advance Access publication August 27, 2015

Meta-analysis of Positive and Negative Symptoms Reveals Schizophrenia Modifier Genes

Alexis C. Edwards*[,1], Tim B. Bigdeli[1], Anna R. Docherty[1], Silviu Bacanu[1], Donghyung Lee[1], Teresa R. de Candia[2,3], Arden Moscati[1], Dawn L. Thiselton[1,9], Brion S. Maher[4], Brandon K. Wormley[1], Molecular Genetics of Schizophrenia Collaboration (MGS)[†], Dermot Walsh[5], Francis A. O'Neill[6], Kenneth S. Kendler[1], Brien P. Riley[1], and Ayman H. Fanous[1,7,8]

[1]Virginia Institute for Psychiatric and Behavioral Genetics, Department of Psychiatry, Virginia Commonwealth University School of Medicine, Richmond, VA; [2]Department of Psychology and Neuroscience, University of Colorado, Boulder, CO; [3]Institute for Behavioral Genetics, University of Colorado, Boulder, CO; [4]Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD; [5]Health Research Board, Dublin, Ireland; [6]Centre for Public Health, Institute of Clinical Sciences, Queen's University Belfast, Belfast, UK; [7]Mental Health Service Line, Washington VA Medical Center, Washington, DC; [8]Department of Psychiatry, Georgetown University School of Medicine, Washington, DC

[9]Present address: Health Diagnostic Laboratory, Inc., Richmond, VA.

*To whom correspondence should be addressed; Virginia Institute for Psychiatric and Behavioral Genetics, Department of Psychiatry, Virginia Commonwealth University School of Medicine, PO Box 980126, Richmond, VA 23298-0126, US; tel: 1-804-828-8591, fax: 1-804-828-1471, e-mail: alexis.edwards@vcuhealth.org

[†]Members listed at the end of the manuscript.

Background: **Evidence suggests that genetic factors may influence both schizophrenia (Scz) and its clinical presentation. In recent years, genome-wide association studies (GWAS) have demonstrated considerable success in identifying risk loci. Detection of "modifier loci" has the potential to further elucidate underlying disease processes.** ***Methods:*** **We performed GWAS of empirically derived positive and negative symptom scales in Irish cases from multiply affected pedigrees and a larger, independent case–control sample, subsequently combining these into a large Irish meta-analysis. In addition to single-SNP associations, we considered gene-based and pathway analyses to better capture convergent genetic effects, and to facilitate biological interpretation of these findings. Replication and testing of aggregate genetic effects was conducted using an independent European-American sample.** ***Results:*** **Though no single marker met the genome-wide significance threshold, genes and ontologies/pathways were significantly associated with negative and positive symptoms; notably,** ***NKAIN2*** **and** ***NRG1*****, respectively. We observed limited overlap in ontologies/pathways associated with different symptom profiles, with immune-related categories over-represented for negative symptoms, and addiction-related categories for positive symptoms. Replication analyses suggested that genes associated with clinical presentation are generalizable to non-Irish samples.** ***Conclusions:*** **These findings strongly support the hypothesis that modifier loci contribute to the etiology of distinct Scz symptom profiles. The finding that previously implicated "risk loci" actually influence particular symptom dimensions has the potential to better delineate the roles of these genes in Scz etiology. Furthermore, the over-representation of distinct gene ontologies/pathways across symptom profiles suggests that the clinical heterogeneity of Scz is due in part to complex and diverse genetic factors.**

Key words: schizophrenia/meta-analysis/clinical heterogeneity/modifier genes/negative symptoms/positive symptoms

Introduction

Schizophrenia (Scz) is a complex psychiatric disorder with a prevalence of approximately 1% that incurs enormous economic, personal, and social costs.[1] Genetic factors contribute substantially to liability to Scz, with twin and family studies typically yielding heritability estimates of approximately 0.6–0.8.[2,3] The Scz spectrum is composed of multiple symptom dimensions,[4] and previous analyses have yielded factor structures consisting of negative, positive, and disorganization symptoms.[5] These dimensions have been shown to be familial and even heritable, though small sample sizes and differences in methods of data collection across studies have limited efforts to identify specific genetic factors.[6] Evidence of phenotypic dimensionality has led to the hypothesis that diverse genetic modifiers underlie Scz.[7]

Genetic factors that influence symptoms in a dimensional fashion, without necessarily conferring risk

for disorder, have been described as "modifier genes." Previous studies have employed genetic linkage,[8,9] candidate gene,[10–12] and genome-wide association approaches[13] in an effort to identify these loci. To maximize statistical power, studies have employed a variety of approaches, from rigorously assessing for clinical homogeneity among cases to using proxies for case/control status (eg, factor scores derived from symptom scales) that are continuous rather than dichotomous variables. These approaches have yielded some promising findings.[14–16]

The current study capitalizes on the availability of comprehensive item-level symptom data in 2 independently ascertained samples of individuals from Ireland and Northern Ireland who were diagnosed with Scz. Both samples have been the subject of analyses aimed at empirically describing clinical heterogeneity.[17–19] The first sample, a collection of 270 multiply affected Irish families, was previously subjected to a latent class analysis (LCA), and subsequent studies have yielded evidence of genetic loci differentially associated with specific clinical features.[8,10,19–21] Genetic studies of clinical heterogeneity among cases in the second, population-based sample have been less extensive.[10]

Here, we conducted analyses of empirically derived symptom dimensions in these 2 ethnically homogenous but differentially ascertained cohorts, combining results for each dimension in a large, case-only meta-analysis. We considered positive symptoms (Pos) and negative symptoms (Neg) as the focal phenotypes, given that comparable classes/factors representing these features have been empirically defined for both samples in previous studies.[17,19] The current study benefits from an improved sample size and continuous phenotypes. Furthermore, these phenotypes are consistent with the Research Domain Criteria (RDoC) project's conceptualization of psychiatric illness[22] and so may better reflect the dimensional nature of Scz psychopathology. In the tradition of previous research,[23] we applied a variety of analytic approaches to capture meaningful genetic influences at the single nucleotide polymorphism (SNP), gene, and ontology/network levels. We hypothesized that different genetic factors—be they SNPs, genes, or gene ontologies/pathways—might detectably influence the diverse clinical presentation of Scz.

Materials and Methods

Samples

Ascertainment for the Irish Study of High Density Schizophrenia Families (ISHDSF) has been described previously.[19] Briefly, the original sample consisted of 270 families (N = 1425 individuals) selected on the basis of having more than one family member who met criteria for DSM-III-R Scz or poor-outcome schizoaffective disorder. Interviews were conducted by Irish psychiatrists and trained social scientists between 1987 and 1992, using modified items from the Structured Clinical Interview for DSM Disorders (SCID) for selected Axis I disorders. For individuals with a history of psychotic illness, the Operational Criteria Checklist for Psychotic Disorders (OPCRIT[24]) was completed (N = 755, of whom 506 had genome-wide association study [GWAS] data available) based on hospital records and personal interviews.

Cases from the Irish Case-Control Study of Schizophrenia (ICCSS) were collected using records from inpatient and outpatient psychiatric facilities, as described previously.[17] To maintain consistency across samples, only individuals meeting DSM-III-R diagnostic criteria for Scz or poor-outcome schizoaffective disorder were eligible for inclusion. GWAS data were available for 665 individuals with Pos scores and 784 with Neg scores.

Phenotypes

The Operational Criteria for Psychotic Illness (OPCRIT) is a comprehensive lifetime symptom scale based on ratings from trained clinicians. For the ISHDSF, 60 of the 75 items on the OPCRIT were selected and entered into an exploratory factor analysis using VARIMAX rotation.[19] A 5-factor solution for the OPCRIT items was selected and further examined in confirmatory factor analysis. Factor-derived scales were obtained for each factor by summing the scores of all items. The 5 factors were: negative symptoms, delusions, hallucinations, mania, and depressive symptoms. For the current study, sum scores for the original negative factor were retained and used in the Neg meta-analysis. To improve comparability with the ICCSS, which yielded a single positive symptom factor, we summed delusion and hallucination factor scores in the ISHDSF to obtain a composite Pos score.

For the ICCSS, symptom ratings were derived from the SCID, and medical records were reviewed and rated using the Casenote Rating Scale.[17] Fanous and colleagues[17] performed an exploratory factor analysis using the VARIMAX rotation and a 3-factor solution was selected, consisting of a positive symptoms factor, a negative symptoms factor, and a Schneiderian symptoms factor. Positive and negative factor scores were retained for the current study for the Pos and Neg phenotypes, respectively.

Genotyping and Imputation

Genotyping and quality control procedures have been described previously for the ISHDSF and ICCSS.[25,26] Haplotype phasing was performed using SHAPEIT.[27] Imputation of additional SNPs was carried out with IMPUTE2 v.2.0[28] using the April 2012 release of the 1000 Genomes Project data (www.1000genomes.org).[29] Imputation analysis was performed for genomic windows of 5 Mb with an overlap interval of 500 Kb between adjacent segments. Following the recommendations of the authors of IMPUTE2, we did not limit our imputation procedure to European reference samples.[30] SNPs were

filtered using an imputation information threshold of 0.5 or higher. Phasing, imputation, and post-processing were performed independently for the ISHDSF and ICCSS samples. Within each study, monomorphic sites were excluded; we also excluded any SNP that yielded at least one Mendelian inconsistency in the family-based sample. Using the full datasets of genotyped and imputed variants, we conducted N-weighted (ie, random effects) meta-analyses using METAL[31] for Neg and Pos. Variants were retained for analysis if they had a minor allele frequency of at least 1%.

Gene-Based Analyses

We used KGG (v2.5)[32,33] to combine single-SNP results into gene-wide tests of association. Gene-based approaches represent a convenient, biology-driven strategy for combining results across studies.[34] KGG assigned SNPs to known genes mapping to the 22 human autosomes; we defined gene boundaries as 5 kb from both 3′ and 5′ ends. We used publicly available linkage disequilibrium (LD) data for European individuals from the 1000 Genomes Project (Phase 1 Release 3) to account for differences in gene size and patterns of intercorrelation among linked SNPs. Specifically, pairs of SNPs demonstrating a high degree of nonindependence ($r^2 > .9$) were "clumped" together; SNPs in low LD ($r^2 < .02$) were considered effectively "independent." Otherwise, correlation among SNPs was handled as described by KGG's authors.[32,33] We applied the HYST method, which combines the extended Simes' and scaled chi-square tests, and adjusted for multiple-testing (of genes) by applying a Benjamini and Hochberg[35] false discovery rate correction of 0.05 (q-values are reported where appropriate). Furthermore, the algorithm employed by KGG corrects for gene size to account for the fact that larger genes span more markers and would thus have a higher probability of including a top SNP by chance alone.

Gene Ontology Analyses

We used ConsensusPathDB[36] to examine potential enrichment of particular gene ontologies (GO) or pathways for disease-related variation, defined here as SNPs meeting modest significance criteria ($P < .001$). These SNPs were mapped to genes using the annotate function in Plink[37]; a curated list of non-redundant genes was uploaded to ConsensusPathDB. GO categories represent groups of genes with functional commonalities (eg, carbohydrate metabolism or drug binding). A single gene can belong to multiple such categories, and categories are nested based on a set of hierarchical definitions of functionality (eg, drug binding is a "child" category of the more general category of "binding"). GO categories with $P < .01$ were obtained. For pathway-based tests, we selected the KEGG, Biocarta, Reactome, and Pharmgkb options, again limiting our results to those with $P < .01$. Where possible, we used multiple other databases to corroborate the initial results, as described in the supplementary material.

Replication and Threshold Testing

We attempted to replicate our top findings using data from the Molecular Genetics of Schizophrenia study (MGS),[13] for which genome-wide SNP data (Affymetrix 6.0) is available in dbGAP.[38] A previous factor analysis identified 3 symptom factors using MGS cases: positive, negative, and mood-related or "affective" symptoms.[13] We conducted GWAS of positive ($N = 2198$) and negative ($N = 2231$) factor scores, as their item content was congruent with the Irish phenotypes. In the current study, we directly imputed summary statistics for additional, untyped SNPs from the 1000 Genomes Project (Phase 1 Release 3) using the DIST software.[39]

First, we conducted sign tests using varying P-value thresholds to determine whether effect directions were consistent across samples. Independent markers, identified using the clumping function in Plink,[37] were used for sign tests. We next selected genes with $P_{gene} < 1 \times 10^{-3}$ for Pos or Neg for follow-up in the MGS sample. Given its larger sample size (and thereby greater statistical power to detect associations), we constructed polygenic risk scores[40] based on estimated allelic effects in the MGS GWAS, testing these composite scores for association with Neg and Pos traits in each Irish cohort.

Results

Marker-Based Analyses

In the current meta-analyses of Neg (maximum $N = 1290$) and Pos (maximum $N = 1171$) symptoms among Irish Scz cases from Ireland and Northern Ireland, results were available for 9 481 181 and 9 485 578 variants, respectively. Results did not yield significant ($P < 5 \times 10^{-8}$) evidence of association with any one SNP for either phenotype. Manhattan plots are presented in supplementary figures 1 and 2. The observed genomic inflation factors did not indicate any overall inflation in the distribution of observed test-statistics (Neg $\lambda = 1.00$, SE $= 1.91 \times 10^{-6}$; Pos $\lambda = 1.02$, SE $= 1.08 \times 10^{-6}$).

For Neg, we observed the strongest evidence of association at chromosome 22q11 (rs1153415; $P = 2.42 \times 10^{-7}$), approximately 20 kb upstream of the uncharacterized locus *LOC388849*. The locus most significantly associated with Pos ($P = 3.90 \times 10^{-7}$) was an intergenic insertion/deletion (chr12:118441149:D) located >20 kb downstream from *SUDS3*. We conducted sign tests using independent markers and found only modest evidence that the directions of allelic effects for Pos were consistent between the Irish and MGS samples more frequently than expected by chance (at $P \leq .001$), and no evidence

of consistency for Neg (supplementary table 1). Similarly, polygenic scores derived using markers associated with Neg or Pos in the MGS sample at various *P*-value thresholds ($P < .0001$–$.5$) did not significantly predict Neg and Pos, respectively, in the Irish cohorts.

Gene-Based Analyses

For Neg, gene-based results were available for 23 000 known and predicted genes (supplementary table 2). Five loci—*NKAIN2*, *LSM6*, *GLRA1*, *G3BP1*, and *BCAT1*—had $q < 0.05$ (Table 1). Of particular note is the finding for *NKAIN2* ($P_{gene} = 7.35 \times 10^{-7}$), which has been previously associated with Scz case/control status in a study that included the ICCSS in a meta-analysis.[41] Of the 22 999 genes or predicted genes assessed (supplementary table 2) for their association with Pos, 3—*NRG1*, *KIAA1430*, and *PHACTR3*—had $q < 0.05$ (Table 1). *NRG1* ($P_{gene} = 1.12 \times 10^{-6}$) has been implicated in Scz case–control status and in bipolar disorder (see below). We also tested whether genes with $P_{gene} < .001$ (a threshold selected to obtain a balance between sample size and false positives) for Neg overlapped significantly with genes meeting a nominal significance threshold ($P_{gene} < .05$) for Pos, based on the null expectation under a binomial distribution, and found that there was significant overlap ($P < .00176$). Conversely, when we tested overlap between top Pos genes and those nominally implicated in Neg (as above), we found no evidence of enrichment ($P = .28$).

Next, we selected genes with $P_{gene} < .001$ ($N = 90$) in the Irish Neg results for follow-up in the MGS sample. Of these, results were available in MGS for 88 genes, 9 of which had $P_{gene} < .05$, which significantly exceeds the number expected by chance (exact binomial test, $P = .03$). Among genes with $q < 0.05$ in the Irish samples, only *BCAT1* was replicated in MGS ($P_{gene} = .0005$).

In the MGS sample, results were available for 83 of the 88 genes that had $P_{gene} < .001$ in the Irish Pos meta-analysis results. Of these, 11 had a $P_{gene} < .05$, which represents a significant enrichment (exact binomial test, $P = .003$). No genes with $q < 0.05$ in the Irish Pos results were replicated in MGS.

Gene Ontology and Pathway Analyses

Negative Symptoms. GO categories with $P < .01$ are provided in supplementary table 3. The most significantly over-represented category was *neuron projection* ($q = 3.48 \times 10^{-8}$), followed by a series of more general cellular process categories. Additional categories related more specifically to nervous system functions/processes were *postsynaptic membrane*, *dendritic development*, *olfactory bulb development*, and *telencephalon development* (among others). There was modest to substantial overlap among the genes driving the over-representation of these categories.

Pathway analyses revealed that nervous system-related categories were consistently over-represented (supplementary table 3) with agreement across different source databases. The top pathway was *axon guidance* ($q = 0.0005$ from KEGG; $q = 0.04$ from Reactome); other nervous system-related categories include *signaling by Robo receptor*, *Netrin-1 signaling*, and *long-term depression*. In addition, several categories related to immune system functioning/processes were implicated, such as *interferon alphalbeta signaling*, *cytokine signaling in immune system*, and *regulation of IFNA signaling*.

Positive Symptoms. As observed among the Neg results, the most strongly implicated category was *neuron projection* ($q = 3.07 \times 10^{-8}$; supplementary table 3). Other enriched nervous system-related categories included *hindbrain development*, *olfactory bulb development*, *cerebellum development, telencephalon development, dendritic spine*, *synapse organization*, and *metencephalon development*.

In the pathway analyses, multiple categories related to addiction phenotypes were over-represented: the top category was *alcoholism* ($q = 3.66 \times 10^{-3}$), and *morphine addiction*, *nicotine addiction*, and *amphetamine addiction* were implicated as well, with substantial overlap among the genes driving the results. Perhaps unsurprisingly, given the relationship between addictive phenotypes and dopaminergic signaling, *dopaminergic synapse* was also over-represented. As with Neg, one or more of the other databases queried also supported most of the categories implicated by ConsensusPathDB.

Discussion

In the current study, we examined evidence of association between common genetic variants and empirically derived symptom factor scores for 2 well-characterized Irish samples, contextualizing the meta-analytic results in terms of specific genes and biological processes. Though no individual marker met the genome-wide significance threshold, gene-based analyses derived from single-SNP test-statistics yielded evidence of genes influencing either positive or negative symptoms but not both, some of which were previously implicated in Scz risk. We also confirmed a role for immune-related and neurobiological processes in Scz etiology. Taken together, these findings suggest that the specificity of genetic variation underlying these processes is nuanced[7] and that select biological pathways might underlie specific clinical dimensions. We discuss the evidence in support of particular biological pathways in the context of emergent Scz findings and extend the putative roles of these processes to clinical presentation of the disease.

Gene-Based Analyses

We observed gene-wise evidence for association between Neg and *NKAIN2*, which has been reported to show association with Scz and with other psychiatric phenotypes (Table 1). This raises the possibility that some

Table 1. Genes with Gene-Based $q < 0.05$ for Negative (Top Panel) and Positive (Bottom Panel) Symptoms

Gene	Gene-Based *P* -Value	Gene/Protein Function	Additional Notes
Negative symptoms			
NKAIN2	$P = 7.35 \times 10^{-7}$	Transmembrane protein that interacts with a Na+/K+ transporting ATPase	Previously associated with case–control status in a meta-analysis that included one of the current samples.[41] That result was based on rs6917824 ($P = .80$ in current Neg results). Previously associated with cognitive decline in Alzheimer's patients[42] based on rs117780815 ($P = .70$ in current Neg results). Previously associated with neuroticism[43] (all associated SNPs $P > .7$ in current Neg results). See text for additional details.
LSM6	$P = 1.34 \times 10^{-6}$	Involved in RNA processing	Perturbations of genes in the Lsm family have been associated with spinal muscular atrophy.
GLRA1	$P = 2.01 \times 10^{-6}$	Glycine receptor subunit; receptor mediates postsynaptic inhibition in CNS	Mutations can inhibit glycine signaling in the CNS. Mutations can cause hereditary hyperekplexia.[44,45]
G3BP1	$P = 2.34 \times 10^{-6}$	DNA-unwinding protein responsive to environmental stress[46,47]	No known previous associations with neurobiological or psychiatric phenotypes.
BCAT1	$P = 2.92 \times 10^{-6}$	Cytosolic form of transaminase; essential for cell growth	Cytosolic BCAT genes are expressed in GABA-ergic and glutamatergic neurons in the brain. These proteins potentially play a role in glutamate toxicity, which contributes to the development of some neurodegenerative disorders.[48]
Positive symptoms			
KIAA1430	$P = 3.23 \times 10^{-7}$	Coding protein of unknown function	Flanked by SLC25A4 ($P = 1.17 \times 10^{-5}$, $q = 0.06$) and SNX25 ($P = 6.15 \times 10^{-5}$, $q = 0.09$). SLC25A4 encodes a mitochondrial ADP transporter that is responsive to escitalopram treatment in a mouse model of depression.[49] SNX25 is expressed in neurons and astrocytes; its expression is increased in the temporal neocortex of patients with temporal lobe epilepsy.[50]
NRG1	$P = 1.12 \times 10^{-6}$	Encodes a glycoprotein that mediates cell–cell signaling; plays a role in growth and development of multiple organ systems; expressed in glutamatergic synaptic vesicles	Dysregulation has been previously associated with schizophrenia and bipolar disorder. See "Discussion" section for additional details.
PHACTR3	$P = 6.39 \times 10^{-6}$	Member of the phosphatase and actin regulator protein family	Encoded protein scapinin is preferentially expressed in the brain; has been shown to enhance cell spreading, potentially through the modulation of actin cytoskeleton structures.[51] Some evidence suggests that scapinin has a regulatory role in neuroplasticity.[52]

NKAIN2 variants impact general liability to Scz, while others affect clinical presentation. Importantly, *NKAIN2* is not associated with case–control status in the combined samples ($P_{gene} = .93$), suggesting that it is acting as a modifier of symptom presentation rather than as a general risk locus, though potentially as a mixed susceptibility/modifier gene[7] in other samples. The remaining genes demonstrating association with Neg at our stringent significance threshold have not previously been shown to confer Scz risk.

For Pos, we observed the strongest gene-wide association with *KIAA1430,* which encodes a protein of unknown function. This locus lies between *SNX25* and *SLC25A4* on chromosome 4 (Table 1). Given low recombination in the region, it is possible that signals in one or both of these flanking genes are relevant, particularly

given the role of *SLC25A4* in other psychiatric phenotypes.[49,53] *Neuregulin-1* (*NRG1*) was also strongly implicated; this protein is present in glutamatergic synaptic vesicles and impacts the regulation of NMDA-receptor expression and has roles in neurodevelopment and synaptic plasticity.[54] It has been implicated as a risk locus for Scz case–control status in multiple populations.[55–58] However, linkage and single-marker analyses in the ISHDSF showed no association between *NRG1* and case–control status (P_{gene} = .54), nor was it associated with Scz in an earlier linkage/association study.[59] Different clinical features across samples of cases could drive these inconsistencies: if cases in one sample were enriched for positive symptoms, it is possible that an association with variation in *NRG1* would be detected. Conversely, case–control analyses (rather than analyses of symptom profiles) of samples whose cases represent a balance of Pos and Neg symptoms would not be expected to detect these loci if they are truly modifiers. Previous evidence suggests that *NRG1* variants have different effects across samples.[60] As with *NKAIN2*, the current results suggest that *NRG1* acts as a modifier gene, though it is potentially a mixed susceptibility/modifier gene. See Table 1 for details on *KIAA140* and *PHACTR3*.

Genes that are strongly associated (P_{gene} < .001) with Neg overlap significantly with genes implicated at a modest significance threshold (P_{gene} < .05) with Pos; however, the reverse does not hold true. Previous studies have found that individuals who experience pronounced negative symptoms are more likely to have a family history of Scz,[61] suggesting that this clinical presentation might be more heritable than others.[18] Negative symptoms might also be more temporally stable, and are easier for clinicians to assess.[17] Our results are consistent with a model in which genes that strongly impact Neg are also relevant to Pos, while those that strongly impact Pos are more specific: essentially, the observed overlap supports modest genetic correlation across the clinical dimensions. Critically, among the genes that were the most strongly associated (q_{gene} < .05) with either Neg or Pos, we observed no overlap. Furthermore, those genes were not associated with case–control status, consistent with the expectation that symptom profiles are affected by modifier genes that are largely phenotype-specific, with other genetic factors influencing overall severity.

Gene Ontology and Pathway Analyses

Enrichment analyses of canonical pathways and GO definitions yielded results consistent with emergent Scz findings[62] while also providing some insight into potential etiological differences underlying clinical dimensions. For both symptom factors, *neuron projection* represented the most strongly over-represented category, meaning that suggestive SNPs (P < .001) mapped to genes in this category more frequently than expected by chance. Furthermore, various nervous system-related ontologies were over-represented in the results for both Neg and Pos, which is consistent with our observation that these dimensions appear genetically related to some degree. However, while genes involved in olfactory bulb and telencephalon development were enriched for both symptom dimensions, genes involved in the development of a broader range of brain regions were enriched only among variants associated with Pos. This raises the possibility that perturbation of functioning in a wide range of brain regions can influence the manifestation of positive symptoms, while negative symptoms are influenced by a more limited set of brain regions.

Categories yielding evidence of enrichment for either Neg or Pos (but not both) included immune function and putative involvement in addiction, respectively. Immune dysregulation has long been thought to play a role in Scz, with support from epidemiological studies[63] and as suggested by the repeated and consistent implication of the major histocompatibility complex.[26,40] Recently, the largest GWAS of Scz conducted to date detected an enrichment of associations among genes expressed in tissues with important roles in immunity, independent of brain-expressed genes.[62]

Over-representation of addiction-related processes with respect to Pos is intriguing, given previous research that suggests individuals with a dual diagnosis of Scz and substance use disorder (SUD) experience more positive and fewer negative symptoms than cases without SUD.[64] Many drugs of abuse enhance dopaminergic activity in the mesolimbic system,[64] and hyperdopaminergic activity/tone in the mesolimbic system has been associated with positive symptoms.[65,66] Furthermore, most dopamine-blocking agents ameliorate positive rather than negative symptoms.[67] The current results suggest that variation in genes that contribute to liability to SUD also impact liability to positive symptoms; this possibility should be explored further as a potential mechanism underlying comorbidity of Scz and SUD. We also note that the dopaminergic system has been implicated in Scz case–control status in a recent mega-analysis.[62]

Limitations

The results presented herein should be interpreted in light of a few key limitations. Given emergent realizations regarding requisite sample sizes for GWAS of complex traits, both Irish samples are relatively limited in size. We were also unable to control for the potential effects of medication. Furthermore, the use of different diagnostic instruments in each sample could conceivably limit comparability, highlighting an important challenge of such studies to date. However, both samples were drawn from the same, ethnically homogenous population, and previous work by our group found concordance between instruments with respect to the assessment of negative

symptoms in the ICCSS,[17] mitigating this concern. In addition, we recently demonstrated the validity of using different instruments to assess the same underlying symptom dimensions in Scz.[68] It is possible that the genetic architecture of Scz differs between the case–control sample and the densely affected family sample.

We observed cross-sample replication at the gene level, including the promising *BCAT1* locus, but polygenic and sign tests were less promising. This could be due to population-level differences, or it could suggest that the current analyses are underpowered. Finally, the wide variety of correlated tests undertaken precludes any straightforward correction for the number of tests conducted. In spite of these limitations, the biologically plausible results of our ontology/pathway analyses support their validity and merit further consideration via follow-up/replication studies.

Conclusions

In spite of these limitations, the biological plausibility of these results supports their validity and merits further consideration via follow-up/replication studies. Broadly speaking, we have demonstrated that negative and positive symptoms of Scz are influenced in part by distinct genetic factors. The results indicate that some genes previously associated with Scz risk might also be modifier or mixed susceptibility-modifier loci. If confirmed, the latter are likely to contribute to "genetic heterogeneity"—long understood to underlie complex phenotypes—as they impact not just clinical presentation but also Scz risk more generally.[7] The gene ontology/pathway analyses suggest that nervous system-related genes influence both negative and positive symptoms, perhaps via influencing overall illness severity or drug response. However, distinctions also exist, specifically with respect to immune- and addiction-related pathways. These findings have potential implications for our understanding of the etiology of Scz from both genetic and neurobiological perspectives, and could be incorporated into future studies on risk assessment and treatment. In summary, these analyses emphasize the complex nature of overall genetic liability to Scz and to its clinical dimensions, and provide insight as to the relevance of common and distinct pathways influencing both.

Supplementary Material

Supplementary material is available at http://schizophreniabulletin.oxfordjournals.org.

Funding

National Institutes of Health (MH083094, MH068881, MH041953, AA021399, MH020030, MH016880, and DA026119), Department of Veterans Affairs (I01 CX000278), and the Wellcome Trust (WTCCC-084710). For the MGS, recruitment, clinical evaluation and clinical data analyses of the MGS cohort were supported by NIH R01 grants (MH67257 to N.G.B., MH59588 to B.J.M., MH59571 to P.V.G., MH59565 to R.F., MH59587 to F.A., MH60870 to W.F.B., MH59566 to D.W.B., MH59586 to J.M.S., MH61675 to D.F.L., MH60879 to C.R.C., and MH81800 to P.V.G.), NIH U01 grants (MH79469 to P.V.G., and MH79470 to D.F.L.), Veterans' Administration grant 1I01CX000278 to A.F., and National Alliance for Research on Schizophrenia and Depression Young Investigator Awards (to A.R.S.).

Acknowledgments

We would like to thank the Molecular Genetics of Schizophrenia Consortium for generously providing data access. The authors have no financial or other potential conflicts of interest to declare.

Members of the Molecular Genetics of Schizophrenia Consortium include the following: Farooq Amin, Department of Psychiatry and Behavioral Sciences, Emory University, Atlanta, GA, and Department of Psychiatry and Behavioral Sciences, Atlanta Veterans Affairs Medical Center, Atlanta, GA; Donald W. Black, Department of Psychiatry, University of Iowa Carver College of Medicine, Iowa City, IA; Nancy G. Buccola, School of Nursing, Louisiana State University Health Sciences Center, New Orleans, LA; William Byerley, Department of Psychiatry, University of California at San Francisco, San Francisco, CA, and Northern California Institute for Research And Education, San Francisco, CA; C. Robert Cloninger, Department of Psychiatry, Washington University, St. Louis, MO; Robert Freedman, Department of Psychiatry, University of Colorado Denver, Aurora, CO; Pablo V. Gejman, Department of Psychiatry and Behavioral Sciences, NorthShore University HealthSystem, Evanston, IL, and Department of Psychiatry and Behavioral Neuroscience, University of Chicago, Chicago, IL; Douglas F. Levinson, Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA; Bryan J. Mowry, Queensland Brain Institute, The University of Queensland, Brisbane, Queensland, Australia, and Queensland Centre for Mental Health Research, University of Queensland, Brisbane, Queensland, Australia; Ann Olincy, Department of Psychiatry, University of Colorado Denver, Aurora, CO; Alan R. Sanders, Department of Psychiatry and Behavioral Sciences, NorthShore University HealthSystem, Evanston, IL, and Department of Psychiatry and Behavioral Neuroscience, University of Chicago, Chicago, IL; Jeremy M. Silverman, Department of Psychiatry, Mount Sinai School of Medicine, New York, NY, and Department of Psychiatry, Bronx Veterans Affairs Medical Center, New York, NY; Dragan M. Svrakic, Department of Psychiatry, Washington University, St. Louis, MO.

References

1. Whiteford HA, Degenhardt L, Rehm J, et al. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *Lancet*. 2013;382:1575–1586.
2. Sullivan PF, Kendler KS, Neale MC. Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. *Arch Gen Psychiatry*. 2003;60:1187–1192.
3. Lichtenstein P, Yip BH, Bjork C, et al. Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet*. 2009;373:234–239.
4. MacDonald AW, Schulz SC. What we know: findings that every theory of schizophrenia should explain. *Schizophr Bull*. 2009;35:493–508.
5. Peralta V, Cuesta MJ. How many and which are the psychopathological dimensions in schizophrenia? Issues influencing their ascertainment. *Schizophr Res*. 2001;49:269–285.
6. McGrath JA, Avramopoulos D, Lasseter VK, et al. Familiality of novel factorial dimensions of schizophrenia. *Arch Gen Psychiatry*. 2009;66:591–600.
7. Fanous AH, Kendler KS. Genetic heterogeneity, modifier genes, and quantitative phenotypes in psychiatric illness: searching for a framework. *Mol Psychiatry*. 2005;10:6–13.
8. Fanous AH, Neale MC, Webb BT, et al. A genome-wide scan for modifier loci in schizophrenia. *Am J Med Genet B Neuropsychiatr Genet*. 2007;144B:589–595.
9. Wilcox MA, Faraone SV, Su J, Van Eerdewegh P, Tsuang MT. Genome scan of three quantitative traits in schizophrenia pedigrees. *Biol Psychiatry*. 2002;52:847–854.
10. Bergen SE, Fanous AH, Kuo PH, et al. No association of dysbindin with symptom factors of schizophrenia in an Irish case-control sample. *Am J Med Genet B Neuropsychiatr Genet*. 2009;153B:700–705.
11. Suchanek R, Owczarek A, Paul-Samojedny M, et al. BDNF val66met polymorphism is associated with age at onset and intensity of symptoms of paranoid schizophrenia in a Polish population. *J Neuropsychiatry Clin Neurosci*. 2013;25:88–94.
12. Tovilla-Zárate C, Medellín BC, Fresán A, et al. No association between catechol-o-methyltransferase Val108/158Met polymorphism and schizophrenia or its clinical symptomatology in a Mexican population. *Mol Biol Rep*. 2013;40:2053–2058.
13. Fanous AH, Zhou B, Aggen SH, et al. Genome-wide association study of clinical dimensions of schizophrenia: polygenic effect on disorganized symptoms. *Am J Psychiatry*. 2012;169:1309–1317.
14. Cummings E, Donohoe G, McDonald C, et al. Clinical symptomatology and the psychosis risk gene ZNF804A. *Schizophr Res*. 2010;122:273–275.
15. DeRosse P, Funke B, Burdick KE, et al. Dysbindin genotype and negative symptoms in schizophrenia. *Am J Psychiatry*. 2006;163:532–534.
16. DeRosse P, Funke B, Burdick KE, et al. COMT genotype and manic symptoms in schizophrenia. *Schizophr Res*. 2006;87:28–31.
17. Fanous AH, Amdur RL, O'Neill FA, Walsh D, Kendler KS. Concordance between chart review and structured interview assessments of schizophrenic symptoms. *Compr Psychiatry*. 2012;53:275–279.
18. Fanous AH, Neale MC, Webb BT, et al. Novel linkage to chromosome 20p using latent classes of psychotic illness in 270 Irish high-density families. *Biol Psychiatry*. 2008;64:121–127.
19. Fanous AH, van den Oord EJ, Riley BP, et al. Relationship between a high-risk haplotype in the *DTNBP1* (Dysbindin) gene and clinical features of schizophrenia. *Am J Psychiatry* 2005;162:1824–1832.
20. Bigdeli TB, Maher BS, Zhao Z, et al. Comprehensive gene-based association study of a chromosome 20 linked region implicates novel risk loci for depressive symptoms in psychotic illness. *PLoS One*. 2011;6:e21440.
21. Fanous AH, Zhao Z, van den Oord EJ, et al. Association study of SNAP25 and schizophrenia in Irish family and case-control samples. *Am J Med Genet B Neuropsychiatr Genet*. 2010;153B:663–674.
22. Cuthbert BN, Insel TR. Toward new approaches to psychotic disorders: the NIMH Research Domain Criteria project. *Schizophr Bull*. 2010;36:1061–1062.
23. Ayalew M, Le-Niculescu H, Levey DF, et al. Convergent functional genomics of schizophrenia: from comprehensive understanding to genetic risk prediction. *Mol Psychiatry*. 2012;17:887–905.
24. Farmer AE, Jones I, McGuffin P. Defining schizophrenia: operational criteria. *J Mental Health*. 1993;2:209–222.
25. Bigdeli TB, Bacanu SA, Webb BT, et al. Molecular validation of the schizophrenia spectrum. *Schizophr Bull*. 2014;40:60–65.
26. Irish Schizophrenia Genomics Consortium, Wellcome Trust Case Control Consortium. Genome-wide association study implicates HLA-C*01:02 as a risk factor at the major histocompatibility complex locus in schizophrenia. *Biol Psychiatry*. 2012;72:620–628.
27. Delaneau O, Marchini J, Zagury JF. A linear complexity phasing method for thousands of genomes. *Nat Methods*. 2012;9:179–181.
28. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet*. 2009;5:e1000529.
29. The 1000 Genomes Project Consortium, Abecasis GR, Auton A, et al. An integrated map of genetic variation from 1,092 human genomes. *Nature*. 2012;491:56–65.
30. Howie B, Marchini J, Stephens M. Genotype imputation with thousands of genomes. *G3 (Bethesda)*. 2011;1:457–470.
31. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. 2010;26:2190–2191.
32. Li MX, Gui HS, Kwan JS, Sham PC. GATES: a rapid and powerful gene-based association test using extended Simes procedure. *Am J Hum Genet*. 2011;88:283–293.
33. Li MX, Kwan JS, Sham PC. HYST: a hybrid set-based test for genome-wide association studies, with application to protein-protein interaction-based association analysis. *Am J Hum Genet*. 2012;91:478–488.
34. Neale BM, Sham PC. The future of association studies: gene-based analysis and replication. *Am J Hum Genet*. 2004;75:353–362.
35. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Stat Soc Ser B (Stat Method)*. 1995;57:289–300.
36. Kamburov A, Pentchev K, Galicka H, Wierling C, Lehrach H, Herwig R. ConsensusPathDB: toward a more complete picture of cell biology. *Nucleic Acids Res*. 2011;39:D712–D717.
37. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81:559–575.

38. Mailman MD, Feolo M, Jin Y, et al. The NCBI dbGaP database of genotypes and phenotypes. *Nat Genet*. 2007;39:1181–1186.
39. Lee D, Bigdeli TB, Riley BP, Fanous AH, Bacanu SA. DIST: direct imputation of summary statistics for unmeasured SNPs. *Bioinformatics*. 2013;29:2925–2927.
40. International Schizophrenia Consortium, Purcell SM, Wray NR, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*. 2009;460:748–752.
41. Aberg KA, Liu Y, Bukszár J, et al. A comprehensive family-based replication study of schizophrenia genes. *JAMA Psychiatry*. 2013;70:573–581.
42. Sherva R, Tripodis Y, Bennett DA, et al. Genome-wide association study of the rate of cognitive decline in Alzheimer's disease. *Alzheimers Dement*. 2014;10:45–52.
43. Calboli FC, Tozzi F, Galwey NW, et al. A genome-wide association study of neuroticism in a population-based sample. *PLoS One*. 2010;5:e11504.
44. Ryan SG, Dixon MJ, Nigro MA, et al. Genetic and radiation hybrid mapping of the hyperekplexia region on chromosome 5q. *Am J Hum Genet*. 1992;51:1334–1343.
45. Shiang R, Ryan SG, Zhu YZ, Hahn AF, O'Connell P, Wasmuth JJ. Mutations in the alpha 1 subunit of the inhibitory glycine receptor cause the dominant neurologic disorder, hyperekplexia. *Nat Genet*. 1993;5:351–358.
46. Gao X, Ge L, Shao J, et al. Tudor-SN interacts with and co-localizes with G3BP in stress granules under stress conditions. *FEBS Lett*. 2010;584:3525–3532.
47. Kwon S, Zhang Y, Matthias P. The deacetylase HDAC6 is a novel critical component of stress granules involved in the stress response. *Genes Dev*. 2007;21:3381–3394.
48. Hull J, Hindy ME, Kehoe PG, Chalmers K, Love S, Conway ME. Distribution of the branched chain aminotransferase proteins in the human brain and their role in glutamate regulation. *J Neurochem*. 2012;123:997–1009.
49. Malki K, Campbell J, Davies M, et al. Pharmacoproteomic investigation into antidepressant response in two mouse inbred strains. *Proteomics*. 2012;12:2355–2365.
50. Du Y, Zou Y, Yu W, et al. Expression pattern of sorting Nexin 25 in temporal lobe epilepsy: a study on patients and pilocarpine-induced rats. *Brain Res*. 2013;1509:79–85.
51. Sagara J, Arata T, Taniguchi S. Scapinin, the protein phosphatase 1 binding protein, enhances cell spreading and motility by interacting with the actin cytoskeleton. *PLoS One*. 2009;4:e4247.
52. Farghaian H, Chen Y, Fu AW, et al. Scapinin-induced inhibition of axon elongation is attenuated by phosphorylation and translocation to the cytoplasm. *J Biol Chem*. 2011;286:19724–19734.
53. Kato M, Nakamura M, Ichiba M, et al. Mitochondrial DNA deletion mutations in patients with neuropsychiatric symptoms. *Neurosci Res*. 2011;69:331–336.
54. Buonanno A, Fischbach GD. Neuregulin and ErbB receptor signaling pathways in the nervous system. *Curr Opin Neurobiol*. 2001;11:287–296.
55. Stefansson H, Sigurdsson E, Steinthorsdottir V, et al. Neuregulin 1 and susceptibility to schizophrenia. *Am J Hum Genet*. 2002;71:877–892.
56. Williams NM, Preece A, Spurlock G, et al. Support for genetic variation in neuregulin 1 and susceptibility to schizophrenia. *Mol Psychiatry*. 2003;8:485–487.
57. Munafò MR, Thiselton DL, Clark TG, Flint J. Association of the NRG1 gene and schizophrenia: a meta-analysis. *Mol Psychiatry*. 2006;11:539–546.
58. Kukshal P, Bhatia T, Bhagwat AM, et al. Association study of neuregulin-1 gene polymorphisms in a North Indian schizophrenia sample. *Schizophr Res*. 2013;144:24–30.
59. Thiselton DL, Webb BT, Neale BM, et al. No evidence for linkage or association of neuregulin-1 (NRG1) with disease in the Irish study of high-density schizophrenia families (ISHDSF). *Mol Psychiatry*. 2004;9:777–783; image 729.
60. Munafo MR, Attwood AS, Flint J. Neuregulin 1 genotype and schizophrenia. *Schizophr Bull*. 2008;34:9–12.
61. Kirkpatrick B, Buchanan RW, Ross DE, Carpenter WT Jr. A separate disease within the syndrome of schizophrenia. *Arch Gen Psychiatry*. 2001;58:165–171.
62. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature*. 2014;511:421–427.
63. Benros ME, Mortensen PB, Eaton WW. Autoimmune diseases and infections as risk factors for schizophrenia. *Ann N Y Acad Sci*. 2012;1262:56–66.
64. Thoma P, Daum I. Comorbid substance use disorder in schizophrenia: a selective overview of neurobiological and cognitive underpinnings. *Psychiatry Clin Neurosci*. 2013;67:367–383.
65. Davis KL, Kahn RS, Ko G, Davidson M. Dopamine in schizophrenia: a review and reconceptualization. *Am J Psychiatry*. 1991;148:1474–1486.
66. Howes OD, Kapur S. The dopamine hypothesis of schizophrenia: version III–the final common pathway. *Schizophr Bull*. 2009;35:549–562.
67. Miyamoto S, Miyake N, Jarskog LF, Fleischhacker WW, Lieberman JA. Pharmacological treatment of schizophrenia: a critical review of the pharmacology and clinical effects of current and future therapeutic agents. *Mol Psychiatry*. 2012;17:1206–1227.
68. Ruderfer DM, Fanous AH, Ripke S, et al. Polygenic dissection of diagnosis and clinical dimensions of bipolar disorder and schizophrenia. *Mol Psychiatry*. 2014;19:1017–1024.