

Inherited L1 Retrotransposon Insertions Associated With Risk for Schizophrenia and Bipolar Disorder

Benjamin C. Reiner^{*.1.○}, Glenn A. Doyle¹, Andrew E. Weller¹, Rachel N. Levinson¹, Aditya M. Rao¹, Emilie Davila Perea¹, Esin Namoglu¹, Alicia Pigeon¹, Gabriella Arauco-Shapiro¹, Cyndi Shannon Weickert^{2,3}, Gustavo Turecki⁴, Richard C. Crist¹, and Wade H. Berrettini¹

¹Molecular and Neural Basis of Psychiatric Disease Section, Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA; ²Schizophrenia Research Laboratory, Neuroscience Research Australia & School of Psychiatry, Faculty of Medicine, University of New South Wales, Sydney, New South Wales, Australia; ³Department of Neuroscience & Physiology, Upstate Medical University, Syracuse, NY, USA; ⁴McGill Group for Suicide Studies, Douglas Mental Health University Institute, McGill University, Montreal, Canada

*To whom correspondence should be addressed; 125 S. 31st St., Philadelphia, PA 19104, United States; tel: 215-573-0278, e-mail: bcreiner@pennmedicine.upenn.edu

Studies of the genetic heritability of schizophrenia and bipolar disorder examining single nucleotide polymorphisms (SNPs) and copy number variations have failed to explain a large portion of the genetic liability, resulting in substantial missing heritability. Long interspersed element 1 (L1) retrotransposons are a type of inherited polymorphic variant that may be associated with risk for schizophrenia and bipolar disorder. We performed REBELseq, a genome wide assay for L1 sequences, on DNA from male and female persons with schizophrenia and controls ($n = 63$ each) to identify inherited L1 insertions and validated priority insertions. L1 insertions of interest were genotyped in DNA from a replication cohort of persons with schizophrenia, bipolar disorder, and controls ($n = 2268$ each) to examine differences in carrier frequencies. We identified an inherited L1 insertion in *ARHGAP24* and a quadallelic SNP (rs74169643) inside an L1 insertion in *SNTG2* that are associated with risk for developing schizophrenia and bipolar disorder (all odds ratios ~ 1.2). Pathway analysis identified 15 gene ontologies that were differentially affected by L1 burden, including multiple ontologies related to glutamatergic signaling and immune function, which have been previously associated with schizophrenia. These findings provide further evidence supporting the role of inherited repetitive genetic elements in the heritability of psychiatric disorders.

Key words: L1Hs/REBELseq

Introduction

Schizophrenia is a chronic psychotic illness, with typical onset in late adolescence or early adulthood. Schizophrenia has substantial heritability, with estimates as high as 80%.^{1,2} Rare copy number variations (CNVs) have been shown to be associated with a substantial increase in schizophrenia disease risk but are only found in a small fraction of individuals with schizophrenia.³⁻⁵ A recent genome wide association study (GWAS) meta-analysis identified 145 genomic loci associated with schizophrenia,⁶ but estimates suggest that common single nucleotide polymorphisms (SNPs) captured by GWAS explain less than half of the genetic liability,⁷ leaving a substantial portion of missing heritability that represents an important gap in our understanding of the genetic etiology of schizophrenia. Similarly, bipolar disorder has high heritability, with estimates up to 80%⁸ and substantial missing heritability.⁹ GWAS of schizophrenia and bipolar disorder demonstrated extensive shared genetic risk.¹⁰ Taken together, this evidence suggests that examination of alternative types of inherited genetic variants may be necessary to develop a more complete understanding of the heritable variants associated with risk for developing schizophrenia and bipolar disorder.

One alternative type of inherited genetic variant is long interspersed element 1 (L1) retrotransposons. L1s are a class of mobile DNA elements that comprise $\sim 17\%$ of the human genome¹¹ and >100 inherited L1 insertions

are known to cause disease.¹² Recent literature supports a role for L1s in the etiology of schizophrenia. Misiak et al.¹³ demonstrated that individuals with schizophrenia have lower L1 methylation, which is an epigenetic change known to increase the transcription and insertion of L1s.^{14,15} Similarly, people with schizophrenia and bipolar disorder show a reduction in neuronal proteins that suppress L1 retrotransposition.^{16,17} Neuronal DNA of individuals with schizophrenia was shown to have an increase in de novo intragenic L1 insertions¹⁸ and a recent report identified a pair of neuronal DNA-specific intronic somatic L1 insertions capable of affecting gene expression.¹⁹ A study of postmortem brain samples from three persons with schizophrenia also showed increased L1 copy number and that L1 insertions occurred in synaptically expressed genes that have been previously associated with schizophrenia.²⁰ Examination of a family where schizophrenia segregated allowed Guffanti et al. to identify the first inherited L1 insertions that were associated with risk for schizophrenia, which occurred in open reading frames for genes that belong to biological pathways already associated with schizophrenia.²¹ However, the family specific nature of these L1 insertions meant they were not generalizable to the broader schizophrenia population. Taken together, these data support a role for L1 retrotransposons in psychiatric disease and suggest that it may be possible to identify additional individually relevant and heritable L1 retrotransposon insertions that are associated with risk for the development of schizophrenia and bipolar disorder. To this end, we identified L1 retrotransposon insertions in the neuronal genomes of a cohort of individuals with schizophrenia and controls and we validated the association of a subset of heritable L1 insertions in independent cohorts of schizophrenia and bipolar disorder.

Methods and Materials

Human Tissue Samples and Identification of L1s

Human postmortem dorsolateral prefrontal cortex samples ($n = 126$) were provided by the Douglas-Bell Canada Brain Bank at McGill University, the Human Brain and Spinal Fluid Resource Center at UCLA and the Sydney Brain Bank at Neuroscience Research Australia, and these studies were approved by the University of Pennsylvania Institutional Review Board. All reported age, sex and ethnicity data are based on associated medical records. Samples included 63 individuals diagnosed with schizophrenia (44 male and 19 female; Age 50.4 ± 15.8 , mean \pm SD; 60 European, 1 African, 1 Hispanic, 1 Asian) and 63 control individuals who died of non-central nervous system-related reasons and had no known neuropsychiatric diseases (50 male and 13 female; Age 55.4 ± 17.8 , mean \pm SD; 57 European, 1 Asian, 5 Unknown). Restriction enzyme-based enriched L1Hs sequencing (REBELseq) is capable of reliably identifying

L1s in the human genome,²² and was used to identify and annotate transcribed active (Ta) subfamily human specific L1 (L1Hs) retrotransposon insertions in neuronal DNA from all participants, with empirically determined quality control cutoffs of mean MapQ minus 2 standard deviations and ≥ 100 average sequencing reads per person. L1 retrotransposons passing bioinformatic cutoffs were categorized based on whether they annotated to known reference L1s (KR L1s) in build hg19 of the human reference genome or if they were known non-reference L1s that were previously detected by other studies (KNR L1s) or previously unknown L1s (UNK L1s). Putative L1 retrotransposon insertions were annotated for occurring in genes, exons, the 1000 Genome project data, hg19 L1Hs, and hg19 non-L1Hs repetitive elements as we have previously described^{22,23} ([supplementary table 1](#)).

Confirmatory PCR

Primers for confirmatory PCR of putative L1 retrotransposon insertions identified by REBELseq were designed as previously described.²² Primer sequences can be found in [supplementary table 2](#). Confirmatory PCR reactions for five L1 insertions were performed on all 126 individuals to prevent both false positive results (ie, individuals who were predicted by REBELseq to have the L1 insertion, but do not upon confirmation) and false negative results (ie, individuals who were not predicted to have the insertion, but do upon confirmation). 25 μ l confirmatory PCR reactions contained 1 \times Go-Taq hotstart colorless master mix (#M5133, Promega), 1 ng of gDNA, 0.2 μ M L1HsACA primer ([supplementary table 2](#)), and 0.2 μ M L1 insertion specific downstream primer. PCR reactions were thermally cycled and visualized by gel electrophoresis as previously described.²²

Population Genotyping by Confirmatory PCR

Population genotyping utilized DNA from schizophrenia, bipolar, and control individuals of European-American ethnicity with no known history of neuropsychiatric disorders ($n = 2268$ each), provided by the NIMH Rutgers University Cell and DNA Repository. 25 μ l PCR reactions contained 1 \times Go-Taq hotstart colorless master mix, 20 ng of DNA, 0.2 μ M L1HsACA primer and 0.2 μ M L1 insertion specific downstream primer. Reactions were run as previously described.²²

Sanger Sequencing

L1 insertions were amplified in a 25 μ l PCR reaction using 1 \times Go-Taq hotstart green master mix (#M5123, Promega), 20 ng of DNA from repository samples of individuals known to be carriers of the L1 insertion of interest, 0.2 μ M L1 insertion specific downstream primer and 0.2 μ M L1 insertion specific upstream primer ([supplementary table 2](#)). Amplification products

encompassing the entire L1 insertion were visualized with ethidium bromide using 1.2% agarose gel electrophoresis. Bands were excised, gel purified (#28606, Qiagen), cloned into pCRII-TOPO plasmids(#45–0640, Invitrogen) and transformed into *E. coli* (#C404010, Invitrogen). Plasmid DNA containing the L1 insertions was Sanger sequenced using big-dye chemistry at the Children's Hospital of Philadelphia Nucleic Acid PCR Core facility and data was analyzed with Sequencher (v5.4.6).

Statistics

Statistical calculations were performed with SPSS (v24, IBM). L1 counts are presented as mean \pm standard error. For population genotyping, Benjamini–Hochberg correction with a false discovery rate (FDR) of 0.1 was utilized for multiple testing correction. PANTHER analysis (PANTHER v14.0; accessed June 6, 2019) was done using the binomial statistical overrepresentation test, with a Bonferroni corrected *P* value < 0.05 considered as significant.

Data Availability

Sequencing and phenotype data associated with this manuscript are available in dbGaP accession number phs001968.

Results

Knowing the literature evidence suggested a possible role for both individual increases in neuronal L1 burden and a subset of inherited L1 retrotransposons being associated with risk for developing schizophrenia, we sought to investigate these dual hypotheses. REBELseq was used to identify L1 insertions in the neuronal genomes of individuals with schizophrenia and control individuals (supplementary Table 1). After stringent bioinformatic cutoffs, we retained 2751 unique L1 insertions across all genomes. An ANCOVA was run to determine the effect of schizophrenia on the total number of L1s per individual and the distribution of L1 subtypes. After adjustment for the effects of age, sex, and ethnicity, there was not a statistically significant difference in the total number of L1s per individual (schizophrenia 583.0 ± 12.2 ; control 599.4 ± 11.0 ; $F(1,121) = 0.391$, $P = 0.533$, partial $\eta^2 = .003$), the number of KR L1s (schizophrenia 445.0 ± 9.7 ; control 459.6 ± 9.0 ; $F(1,121) = 0.454$, $P = 0.502$, partial $\eta^2 = .004$), the number of KNR L1s (schizophrenia 66.7 ± 1.8 ; control 69.3 ± 2.0 ; $F(1,121) = 0.229$, $P = 0.633$, partial $\eta^2 = .002$), or the number of UNK L1s (schizophrenia 71.3 ± 1.6 ; control 70.5 ± 1.1 ; $F(1,121) = 0.004$, $P = 0.950$, partial $\eta^2 < .001$). Taken together, these data suggest that there is no increase in L1 retrotransposon burden in the neuronal genome of individuals with schizophrenia.

PANTHER Pathway Analysis

Of the 2751 unique L1 insertions retained after bioinformatic cutoffs; 1150 insertions occurred in genes, with 914 unique genes containing L1 insertions in persons with schizophrenia, 862 unique genes containing L1 insertions in control individuals, and 626 of the L1 insertions occurring in both groups. These lists of unique genes were analyzed with PANTHER, to identify gene ontologies for biological processes, molecular functions, and cellular components that are overrepresented or underrepresented in L1 insertions from either schizophrenia or control persons, but not both. We identified 15 gene ontologies where L1s were overrepresented or underrepresented in schizophrenia or control individuals (supplementary table 3), including multiple ontologies related to glutamate or calcium signaling, areas of interest in the context of schizophrenia.

Confirmation of REBELseq L1s and Population Genotyping

Knowing that some inherited L1s have been associated with risk for developing schizophrenia,^{19,21} we selected five L1 insertions occurring in or near genes previously associated with schizophrenia^{24–39} (supplementary table 2) and validated their presence in the genomes of our samples by confirmatory PCR, utilizing the L1 3' untranslated region (UTR) targeted L1HsACA primer (supplementary table 2) and a L1 insertion specific downstream reverse primer (supplementary table 2), to ensure that the L1 insertions were true retrotransposon insertions and not chimeric amplification products from REBELseq. All five L1 insertions were validated successfully (supplementary table 4). Knowing that our small discovery sample may not accurately represent the carrier frequencies of the five validated L1 insertions in larger populations and that inherited L1s can be identified from peripheral sources of genomic material, we next determined the frequency of these five L1 insertions in a larger set of blood derived DNA samples from Caucasian European-American persons with schizophrenia and controls ($n = 2268$ each). The carrier frequency of the L1 insertions in *ARHGAP24* ($\chi^2 P = 0.029$, odds ratio (OR) 1.206) and *SNTG2* ($\chi^2 P = 0.008$, OR 1.182) were significantly increased in persons with schizophrenia, when compared to controls, suggesting that the presence of these L1 retrotransposons is associated with risk for the development of schizophrenia (Table 1).

Since GWAS results have shown that schizophrenia and bipolar disorder have a large number of overlapping genetic risk loci,¹⁰ we next sought to investigate if the L1 insertions in *ARHGAP24* and *SNTG2* might also be associated with risk for developing bipolar disorder. The L1 insertions were genotyped in DNA samples from persons with bipolar disorder ($n = 2268$), and a significant increase in the carrier frequencies of the *ARHGAP24* (χ^2

Table 1. The Number of Individuals Carrying the Genotyped L1 Insertions (Out of $n = 2268$ Each), Carrier Frequency (%), and the Odds Ratio (OR) of Each Insertion

Gene	Control Carriers (Frequency)	Schizophrenia Carriers (Frequency)	<i>P</i> Value	OR	OR 95% Confidence Interval
ARHGAP24	292 (12.9%)	344 (15.2%)	0.029	1.206	1.019–1.426
SNTG2	728 (32.2%)	813 (36.0%)	0.008	1.182	1.045–1.337
LINC01899	381 (16.9%)	378 (16.7%)	0.905	0.991	0.848–1.158
UNC5D	201 (8.9%)	205 (9.1%)	0.876	1.022	0.833–1.253
RGS6	41 (1.8%)	54 (2.4%)	0.178	1.325	0.879–1.997

Table 2. The Number of Individuals Carrying the Genotyped L1 Insertions (Out of $n = 2268$ Each), the Carrier Frequency (%), and the Odds Ratio (OR) of Each Insertion Compared to Controls

Gene	Bipolar Carriers (Frequency)	<i>P</i> Value	OR	OR 95% Confidence Interval
ARHGAP24	347 (15.4%)	0.021	1.218	1.030–1.441
SNTG2	826 (36.5%)	0.002	1.212	1.072–1.370

$P = 0.021$, OR 1.218) and *SNTG2* ($\chi^2 P = 0.002$, OR 1.212) L1 insertions in persons with bipolar disorder, compared to controls, was detected (Table 2). This suggests that the *ARHGAP24* and *SNTG2* L1 retrotransposon insertions are associated with risk for both schizophrenia and bipolar disorder.

L1 Sanger Sequencing

Since the L1 insertions in *ARHGAP24* and *SNTG2* are associated with risk for the development of schizophrenia and bipolar disorder and inherited L1 insertions are frequently mutated or truncated,⁴⁰ we next sought to elucidate the exact size and genetic sequence of these L1 insertions. Utilizing DNA from individuals known to be carriers of the *ARHGAP24* and *SNTG2* L1 insertions, we performed PCR reactions using the upstream and downstream L1 insertion specific primers (supplementary table 2), which allowed for the amplification, cloning and Sanger sequencing of the entire *ARHGAP24* and *SNTG2* L1 insertions.

The inherited L1 insertion in *ARHGAP24* that is associated with risk for schizophrenia and bipolar disorder is a KNR L1 that is 448bp in length, inserted at chr4:86,607,955-86,607,956 (annotated in euL1db as MRIP 5707). It is a 3' fragment of an L1Hs, containing both the diagnostic L1HsACA sequence and L1HsG sequence^{22,41,42} in what would have been the 3' UTR of a full length L1Hs (see supplemental for sequence).

The L1 insertion in *SNTG2* is a KR L1 that is 342bp in length, inserted at chr2:1,248,710-1,249,051. Initial attempts at Sanger sequencing the L1 insertion showed heterozygosity within the L1HsACA sequence (data not shown). To better understand this observation, we Sanger

Genotyped Haplotype		# alleles in carriers	# alleles in non-carriers
GATGACAC		12	0
GATGACAT		8	14
Alternative Haplotypes	AATGACGA	0	5
	AATGACGG	0	1

$\begin{matrix} \text{T} & & \text{T} \\ \text{rs28652617} & & \text{rs74169643} \\ & & \text{rs28538796} \end{matrix}$

Fig. 1. Haplotypes of the *SNTG2* L1 retrotransposon insertion. 10 carriers of L1 insertion and 10 non-carriers were sequenced and the two alleles from each individual were identified. The haplotype labeled as the genotyped haplotype was present in all the carriers and none of the non-carriers. The three alternative haplotypes are shown. Counts of the number of each of the four haplotypes alleles are presented. The SNP accession numbers are shown for reference. The “GAC” haplotype of rs28652617, rs28538796, and rs74169643 is associated with risk for developing schizophrenia and bipolar disorder.

sequenced clones ($n = 5$ /person) from 10 people who were called as carriers for the L1 insertion and 10 people who were called as non-carriers during genotyping. Three SNPs (rs28652617, rs28538796, and rs74169643) were identified in the binding site of the L1HsACA primer. We ascertained four haplotypes in this group of three SNPs, one of which matched the L1HsACA primer sequence (Genotyped Haplotype; figure 1). This haplotype occurred in all individuals who were genotyped as carriers for the *SNTG2* L1 insertion and none of the individuals who were genotyped as non-carriers. To confirm that our population genotyping results were not a byproduct of truncated primers or 3'–5' taq exonuclease activity, we repeated our population genotyping for the 20 sequenced individuals using either the L1HsACA primer previously used for population genotyping or a HPLC-purified L1HsACA primer that contained a phosphorothioate bond of the 3' A nucleotide to block exonucleolytic activity, ensuring that all primer was full length. We observed the same genotyping results with both versions of the L1HsACA primer (data not shown), confirming that what we genotyped for the *SNTG2* L1 insertion was a specific haplotype block of the three SNPs. The “GA” haplotype at rs28652617 and rs28538796 occurs in the context of both the “C” and “T” alleles at rs74169643

(figure 1), while the “C” allele at rs74169643 only occurs in the context of the genotyped haplotype (figure 1). Therefore, it is the “C” allele at rs74169643 that is associated with risk for developing both schizophrenia and bipolar disorder.

Discussion

While a subset of inherited polymorphic L1s have strong linkage disequilibrium (LD) to nearby SNPs,^{43,44} it has been demonstrated that common SNP arrays, like the type utilized for GWAS, poorly identify the presence of polymorphic L1s.⁴⁵ This suggests that alternative approaches to GWAS, like REBELseq,²² may be necessary to understand the role L1s play in human disease. Similar to other recent reports,^{19,21} this manuscript offers evidence suggesting that a subset of inherited L1 retrotransposon insertions are associated with risk for the development of psychiatric disorders, with the identification of an inherited L1 retrotransposon insertion in *ARHGAP24* and a quadallelic SNP in a L1 insertion in *SNTG2* that are associated with risk for the development of schizophrenia and bipolar disorder. Additionally, our pathway analyses identified gene ontologies that are differentially affected by L1 retrotransposon burden.

The role of L1 retrotransposons in schizophrenia is an active area of clinical speculation,⁴⁶ and risk for the development of schizophrenia has been previously associated with human endogenous viruses.⁴⁷ Bundo et al.²⁰ described a difference in L1 retrotransposon burden in NeuN+ neuronal DNA from prefrontal cortex of persons with schizophrenia and bipolar disorder, utilizing a real-time PCR assay of L1 opening reading frame (ORF) 2. This assay quantifies total genomic L1 content, including full length L1s, L1s that are 5′ truncated upstream of ORF2 and L1s that are 3′ truncated downstream of ORF2. In contrast, the REBELseq assay is designed to identify L1s that contain the L1HsACA and L1HsG diagnostic nucleotides of Ta subfamily L1Hs in the 3′ UTR,²² downstream of ORF2. These differences in assay system may explain why Bundo et al. observed an increase in total L1 burden when we did not.

The inherited L1 insertion in *ARHGAP24* is in intron 2, similar to a previous report of inherited intronic L1 insertions that were associated with risk for the development of schizophrenia.²¹ *ARHGAP24* is a Rho GTPase activating protein, which antagonizes Rac. GWAS has associated rs111882035 in *ARHGAP24* with a deficit in logical memory (delayed recall) in mild cognitive impairment,⁴⁸ and missense mutation in Rho GTPase activating proteins have been associated with schizophrenia.⁴⁹ *ARHGAP24* functions in cytoskeleton regulation and changes in microtubule regulation are a hallmark of psychiatric disease,⁵⁰ including schizophrenia⁵¹ and bipolar disorder.⁵² Pharmacological alteration of cytoskeletal regulation has even been suggested

as a potential therapeutic target for schizophrenia.⁵³ Intriguingly, *ARHGAP18*, encoding a Rho GTPase that also plays a role in cytoskeletal regulation, contains alleles associated with schizophrenia.⁵⁴ Full length *ARHGAP24* isoform 1 has been shown to be an important regulator of neuronal morphology.⁵⁵ Taken together, this evidence suggests a putative mechanism whereby the inherited L1 insertion in intron 2 of *ARHGAP24* (or a variant in LD) could play a role in schizophrenia and bipolar disorder by dysregulating normal cytoskeletal function. Further supporting this hypothesis is evidence from the *Disc1*-L100P mouse line, which demonstrates that disruption in *Disc1* leads to decreased expression of *Arhgap24*⁵⁶ and deficits in synaptic morphology and plasticity and memory.⁵⁷

SNTG2 encodes Syntrophin Gamma 2, a membrane protein that binds to the C-terminal domain of dystrophins and dystrophin-related proteins.⁵⁸ Dystrophin has been previously associated with schizophrenia,⁵⁹ and a CNV in *SNTG2* segregates with schizophrenia in a family.³⁶ Additionally, examination of 18 parent–child trios with bipolar disorder demonstrated an association with variants in *SNTG2*.⁶⁰ The haplotype block in the L1 insertion in *SNTG2* shows that rs28652617 “G” allele and rs28538796 “A” allele are in perfect LD. rs74169643 is a quadallelic SNP; however, the “C” allele in our genotyped haplotype structure only occurs in the context of rs28652617 “G” allele and rs28538796 “A” allele, suggesting that it is effectively the rs74169643 “C” allele that was associated with risk for the development of schizophrenia and bipolar disorder. rs74169643 was not included in previous GWAS of schizophrenia,⁶¹ likely due to the quadallelic nature of the variant, explaining why an association between rs74169643 and schizophrenia was not previously identified, despite larger sample size. Our results suggest that genotyping of rs74169643 in large populations of schizophrenia and control individuals is warranted.

Our pathway analysis, which utilized lists of unique genes containing L1s in the genome of our human post-mortem samples, identified many gene ontologies that were overrepresented or underrepresented in the context of schizophrenia or control individuals. Overrepresentation of L1s in the biological process “synaptic transmission, glutamatergic” and the cellular component “ionotropic glutamate receptor complex” may support the glutamate hypothesis of schizophrenia,⁶² whereas underrepresentation of L1s in the biological processes “innate immune response,” “defense response to bacterium,” and “humoral immune response” may be relevant to the long-suspected involvement of immune system regulation in schizophrenia.^{63,64} Genes containing L1 retrotransposons involved in these and other identified gene ontologies may make interesting targets for future studies.

One limitation of this study is the utilization of REBELseq, which identifies Ta subfamily L1Hs

insertions. While it was utilized productively in the context of this manuscript, it fails to identify other families of chronologically older L1 insertions and other types of repetitive genomic elements that may play a role in neuropsychiatric disorders. Further examination of other repetitive element types in the context of schizophrenia and bipolar disorder is likely warranted. Another limitation of this study is the use of relatively small, by psychiatric genetics standards, discovery and validation cohorts, limiting both the identification and independent confirmation of relevant L1 retrotransposons. Additional studies with larger sample sizes will be important for ultimately determining the relevance of the variants identified in this study. The allele frequency of rs74169643 shows considerable global variation, and for this reason all samples in the replication cohort were Caucasian European-Americans, reducing the potential for false positive result caused by population stratification. However, further analysis of samples using a SNP or sequencing based analysis is needed to fully disprove the effects of population stratification. A further limitation of this manuscript is the lack of functional studies, as it will be important to determine the mechanisms that underlie the identified genomic variants.

This study provides further evidence that L1 retrotransposons are associated with risk for the development of schizophrenia and bipolar disorder. However, there is no reason to assume that this association to psychiatric disease is limited to schizophrenia and bipolar disorder. Further studies of the role of L1s in other psychiatric disorders is warranted, as they may identify additional L1 retrotransposons that are associated with risk.

Supplementary Material

Supplementary data are available at *Schizophrenia Bulletin* Open online.

Acknowledgments

The authors would like to thank the brain banks that participated in this study and the individuals who donated their brains to science, without which, this study would not have been possible. BCR was supported by a 2017 NARSAD Young Investigator Award (#26634) from the Brain and Behavior Research Foundation, as the Patrick A. Coffey Investigator, funding generously provided by Ronald and Kathy Chandonais. BR was supported during part of this work by T32 MH014654 (WHB is PI). RCC was supported by grant K01 DA036751. AEW was supported by T32 MH014654 (WHB is PI). This manuscript was supported by R01 MH109260 (WHB is PI). Bio-samples and/or data for this publication were obtained from NIMH Repository & Genomics Resource, a centralized national biorepository for genetic studies of

psychiatric disorders. Detailed acknowledgements for all NIMH studies whose samples were utilized in this manuscript can be found in the supplemental.

Disclosures

The authors declare no conflicts of interest.

References

1. Lichtenstein P, Yip BH, Björk C, *et al.* Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet*. 2009;373(9659):234–239.
2. Hilker R, Helenius D, Fagerlund B, *et al.* Heritability of schizophrenia and schizophrenia spectrum based on the nationwide Danish twin register. *Biol Psychiatry*. 2018;83(6):492–498.
3. Stefansson H, Rujescu D, Cichon S, *et al.*; GROUP. Large recurrent microdeletions associated with schizophrenia. *Nature*. 2008;455(7210):232–236.
4. Walsh T, McClellan JM, McCarthy SE, *et al.* Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science*. 2008;320(5875):539–543.
5. Xu B, Roos JL, Levy S, van Rensburg EJ, Gogos JA, Karayiorgou M. Strong association of de novo copy number mutations with sporadic schizophrenia. *Nat Genet*. 2008;40(7):880–885.
6. Pardiñas AF, Holmans P, Pocklington AJ, *et al.*; GERAD1 Consortium; CRESTAR Consortium. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nat Genet*. 2018;50(3):381–389.
7. Purcell SM, Moran JL, Fromer M, *et al.* A polygenic burden of rare disruptive mutations in schizophrenia. *Nature*. 2014;506(7487):185–190.
8. Nöthen MM, Nieratschker V, Cichon S, Rietschel M. New findings in the genetics of major psychoses. *Dialogues Clin Neurosci*. 2010;12(1):85–93.
9. Woo HJ, Yu C, Kumar K, Reifman J. Large-scale interaction effects reveal missing heritability in schizophrenia, bipolar disorder and posttraumatic stress disorder. *Transl Psychiatry*. 2017;7(4):e1089.
10. Bipolar D, Schizophrenia Working Group of the Psychiatric Genomics Consortium. Genomic dissection of bipolar disorder and schizophrenia, including 28 subphenotypes. *Cell*. 2018;173(7):1705–1715.e1716.
11. Beck CR, Collier P, Macfarlane C, *et al.* LINE-1 retrotransposition activity in human genomes. *Cell*. 2010;141(7):1159–1170.
12. Hancks DC, Kazazian HH Jr. Roles for retrotransposon insertions in human disease. *Mob DNA*. 2016;7:9.
13. Misiak B, Szmida E, Karpiński P, Loska O, Szażdek MM, Frydecka D. Lower LINE-1 methylation in first-episode schizophrenia patients with the history of childhood trauma. *Epigenomics*. 2015;7(8):1275–1285.
14. Rosser JM, An W. L1 expression and regulation in humans and rodents. *Front Biosci (Elite Ed)*. 2012;4:2203–2225.
15. Sanchez-Luque FJ, Kempen MHC, Gerdes P, *et al.* LINE-1 evasion of epigenetic repression in humans. *Mol Cell*. 2019;75(3):590–604 e512.
16. Dong E, Gavin DP, Chen Y, Davis J. Upregulation of TET1 and downregulation of APOBEC3A and APOBEC3C in

- the parietal cortex of psychotic patients. *Transl Psychiatry*. 2012;2:e159.
17. Goodier JL. Restricting retrotransposons: a review. *Mob DNA*. 2016;7:16.
 18. Doyle GA, Crist RC, Karatas ET, *et al*. Analysis of LINE-1 elements in DNA from postmortem brains of individuals with schizophrenia. *Neuropsychopharmacology*. 2017;42:2602–2611.
 19. Zhu X, Zhou B, Pattni R, *et al*; Brain Somatic Mosaicism Network. Machine learning reveals bilateral distribution of somatic L1 insertions in human neurons and glia. *Nat Neurosci*. 2021;24(2):186–196.
 20. Bundo M, Toyoshima M, Okada Y, *et al*. Increased L1 retrotransposition in the neuronal genome in schizophrenia. *Neuron*. 2014;81(2):306–313.
 21. Guffanti G, Gaudi S, Klengel T, *et al*. LINE1 insertions as a genomic risk factor for schizophrenia: preliminary evidence from an affected family. *Am J Med Genet B Neuropsychiatr Genet*. 2016;171(4):534–545.
 22. Reiner BC, Doyle GA, Weller AE, *et al*. Restriction enzyme based enriched L1Hs sequencing (REBELseq): a scalable technique for detection of Ta subfamily L1Hs in the human genome. *G3-GENES GENOM GENET*. 2020;10(5):1647–1655.
 23. Doyle GA, Reiner BC, Crist RC, *et al*. Investigation of long interspersed element-1 retrotransposons as potential risk factors for idiopathic temporal lobe epilepsy. *Epilepsia*. 2021;62(6):1329–1342.
 24. Ahlers KE, Chakravarti B, Fisher RA. RGS6 as a novel therapeutic target in CNS diseases and cancer. *AAPS J*. 2016;18(3):560–572.
 25. Bassett AS, Lowther C, Merico D, *et al*; International 22q11.2DS Brain and Behavior Consortium. Rare genome-wide copy number variation and expression of schizophrenia in 22q11.2 deletion syndrome. *Am J Psychiatry*. 2017;174(11):1054–1063.
 26. Fabbri C, Serretti A. Role of 108 schizophrenia-associated loci in modulating psychopathological dimensions in schizophrenia and bipolar disorder. *Am J Med Genet B Neuropsychiatr Genet*. 2017;174(7):757–764.
 27. Gu Z, Jiang Q, Yan Z. RGS4 modulates serotonin signaling in prefrontal cortex and links to serotonin dysfunction in a rat model of schizophrenia. *Mol Pharmacol*. 2007;71(4):1030–1039.
 28. Li Z, Xiang Y, Chen J, *et al*. Loci with genome-wide associations with schizophrenia in the Han Chinese population. *Br J Psychiatry*. 2015;207(6):490–494.
 29. Liedtke EI, Zhang S, Thompson JA, Sillau S, Gault J. Correlated expression analysis of genes implicated in schizophrenia: identification of putative disease-related pathways. *New Horiz Transl Med*. 2017;3(5):224–232.
 30. Lin JR, Cai Y, Zhang Q, Zhang W, Nogales-Cadenas R, Zhang ZD. Integrated post-GWAS analysis sheds new light on the disease mechanisms of schizophrenia. *Genetics*. 2016;204(4):1587–1600.
 31. Novikova SI, He F, Cutrufello NJ, Lidow MS. Identification of protein biomarkers for schizophrenia and bipolar disorder in the postmortem prefrontal cortex using SELDI-TOF-MS ProteinChip profiling combined with MALDI-TOF-PSD-MS analysis. *Neurobiol Dis*. 2006;23(1):61–76.
 32. Ohi K, Shimada T, Nitta Y, *et al*. Specific gene expression patterns of 108 schizophrenia-associated loci in cortex. *Schizophr Res*. 2016;174(1-3):35–38.
 33. Peng X, Bader JS, Avramopoulos D. Schizophrenia risk alleles often affect the expression of many genes and each gene may have a different effect on the risk: a mediation analysis. *Am J Med Genet B: Neuropsychiatr Genet*. 2021;186(4):251–258.
 34. Radulescu E, Jaffe AE, Straub RE, *et al*. Identification and prioritization of gene sets associated with schizophrenia risk by co-expression network analysis in human brain. *Mol Psychiatry*. 2020;25(4):791–804.
 35. Srikanth P, Lagomarsino VN, Pearse RV 2nd, *et al*. Convergence of independent DISC1 mutations on impaired neurite growth via decreased UNC5D expression. *Transl Psychiatry*. 2018;8(1):245.
 36. Van Den Bossche MJ, Strazisar M, Cammaerts S, *et al*. Identification of rare copy number variants in high burden schizophrenia families. *Am J Med Genet B Neuropsychiatr Genet*. 2013;162B(3):273–282.
 37. Wu C, Pan W. A powerful fine-mapping method for transcriptome-wide association studies. *Hum Genet*. 2020;139(2):199–213.
 38. Yazdani A, Mendez-Giraldez R, Yazdani A, Kosorok MR, Roussos P. Differential gene regulatory pattern in the human brain from schizophrenia using transcriptomic-causal network. *BMC Bioinformatics*. 2020;21(1):469.
 39. Zhu L, Lei J, Devlin B, Roeder K. Testing high-dimensional covariance matrices, with application to detecting schizophrenia risk genes. *Ann Appl Stat*. 2017;11(3):1810–1831.
 40. Ostertag EM, Kazazian HH Jr. Biology of mammalian L1 retrotransposons. *Annu Rev Genet*. 2001;35:501–538.
 41. Boissinot S, Chevret P, Furano AV. L1 (LINE-1) retrotransposon evolution and amplification in recent human history. *Mol Biol Evol*. 2000;17(6):915–928.
 42. Ewing AD, Kazazian HH Jr. Whole-genome resequencing allows detection of many rare LINE-1 insertion alleles in humans. *Genome Res*. 2011;21(6):985–990.
 43. Higashino S, Ohno T, Ishiguro K, Aizawa Y. Polymorphic L1 retrotransposons are frequently in strong linkage disequilibrium with neighboring SNPs. *Gene*. 2014;541(1):55–59.
 44. Wang L, Norris ET, Jordan IK. Human retrotransposon insertion polymorphisms are associated with health and disease via gene regulatory phenotypes. *Front Microbiol*. 2017;8:1418.
 45. Kuhn A, Ong YM, Cheng CY, Wong TY, Quake SR, Burkholder WF. Linkage disequilibrium and signatures of positive selection around LINE-1 retrotransposons in the human genome. *Proc Natl Acad Sci U S A*. 2014;111(22):8131–8136.
 46. Hussein MA, Ross DA. Leaping forward: the surprising role of jumping genes in psychiatric genetics. *Biol Psychiatry*. 2019;86(5):e13–e15.
 47. Sekar A, Bialas AR, de Rivera H, *et al*; Schizophrenia Working Group of the Psychiatric Genomics Consortium. Schizophrenia risk from complex variation of complement component 4. *Nature*. 2016;530(7589):177–183.
 48. Chung J, Wang X, Maruyama T, *et al*; Alzheimer's Disease Neuroimaging Initiative. Genome-wide association study of Alzheimer's disease endophenotypes at prediagnosis stages. *Alzheimers Dement*. 2018;14(5):623–633.
 49. Hashimoto R, Yoshida M, Kunugi H, *et al*. A missense polymorphism (H204R) of a Rho GTPase-activating protein, the chimerin 2 gene, is associated with schizophrenia in men. *Schizophr Res*. 2005;73(2-3):383–385.
 50. Marchisella F, Coffey ET, Hollos P. Microtubule and microtubule associated protein anomalies in psychiatric disease. *Cytoskeleton (Hoboken)*. 2016;73(10):596–611.
 51. Brown AS, Borgmann-Winter K, Hahn CG, *et al*. Increased stability of microtubules in cultured olfactory

- neuroepithelial cells from individuals with schizophrenia. *Prog Neuropsychopharmacol Biol Psychiatry*. 2014;48:252–258.
52. Hodgkinson CA, Goldman D, Jaeger J, et al. Disrupted in schizophrenia 1 (DISC1): association with schizophrenia, schizoaffective disorder, and bipolar disorder. *Am J Hum Genet*. 2004;75(5):862–872.
53. Benitez-King G, Ramirez-Rodriguez G, Ortiz L, Meza I. The neuronal cytoskeleton as a potential therapeutic target in neurodegenerative diseases and schizophrenia. *Curr Drug Targets CNS Neurol Disord*. 2004;3(6):515–533.
54. Guo W, Cai Y, Zhang H, et al. Association of ARHGAP18 polymorphisms with schizophrenia in the Chinese-Han population. *PLoS One*. 2017;12(4):e0175209.
55. Nguyen LS, Jolly L, Shoubridge C, et al. Transcriptome profiling of UPF3B/NMD-deficient lymphoblastoid cells from patients with various forms of intellectual disability. *Mol Psychiatry*. 2012;17(11):1103–1115.
56. Lipina TV, Haque FN, McGirr A, et al. Prophylactic valproic acid treatment prevents schizophrenia-related behaviour in Disc1-L100P mutant mice. *PLoS One*. 2012;7(12):e51562.
57. Cui L, Sun W, Yu M, et al. Disrupted-in-schizophrenia 1 (DISC1) L100P mutation alters synaptic transmission and plasticity in the hippocampus and causes recognition memory deficits. *Mol Brain*. 2016;9(1):89.
58. Piluso G, Mirabella M, Ricci E, et al. Gamma1- and gamma2-syntrophins, two novel dystrophin-binding proteins localized in neuronal cells. *J Biol Chem*. 2000;275(21):15851–15860.
59. Lindor NM, Sobell JL, Heston LL, Thibodeau SN, Sommer SS. Screening the dystrophin gene suggests a high rate of polymorphism in general but no exonic deletions in schizophrenics. *Am J Med Genet*. 1994;54(1):1–4.
60. Georgi B, Craig D, Kember RL, et al. Genomic view of bipolar disorder revealed by whole genome sequencing in a genetic isolate. *PLoS Genet*. 2014;10(3):e1004229.
61. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature*. 2014;511(7510):421–427.
62. Hu W, MacDonald ML, Elswick DE, Sweet RA. The glutamate hypothesis of schizophrenia: evidence from human brain tissue studies. *Ann N Y Acad Sci*. 2015;1338:38–57.
63. Müller N, Schwarz MJ. Immune system and schizophrenia. *Curr Immunol Rev*. 2010;6(3):213–220.
64. Patterson PH. Immune involvement in schizophrenia and autism: etiology, pathology and animal models. *Behav Brain Res*. 2009;204(2):313–321.