

2015

Association of SNPs in EGR3 and ARC with schizophrenia supports a biological pathway for schizophrenia risk

Matthew J. Huentelman
Translational Genomics Research Institute, Phoenix

Leela Muppana
University of Arizona College of Medicine-Phoenix

Jason J. Corneveaux
Translational Genomics Research Institute, Phoenix

Valentin Dinu
Arizona State University

Jeremy J. Pruzin
Translational Genomics Research Institute, Phoenix

See next page for additional authors

Follow this and additional works at: https://digitalcommons.wustl.edu/open_access_pubs

Please let us know how this document benefits you.

Recommended Citation

Huentelman, Matthew J.; Muppana, Leela; Corneveaux, Jason J.; Dinu, Valentin; Pruzin, Jeremy J.; Reiman, Rebecca; Borish, Cassie N.; De Both, Matt; Ahmed, Amber; Todorov, Alexandre; Cloninger, C. Robert; Zhang, Rui; Ma, Jie; and Gallitano, Amelia L., "Association of SNPs in EGR3 and ARC with schizophrenia supports a biological pathway for schizophrenia risk." *PLoS One*. 10, 10. e0135076. (2015). https://digitalcommons.wustl.edu/open_access_pubs/5054

This Open Access Publication is brought to you for free and open access by Digital Commons@Becker. It has been accepted for inclusion in Open Access Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact vanam@wustl.edu.

Authors

Matthew J. Huentelman, Leela Muppana, Jason J. Corneveaux, Valentin Dinu, Jeremy J. Pruzin, Rebecca Reiman, Cassie N. Borish, Matt De Both, Amber Ahmed, Alexandre Todorov, C. Robert Cloninger, Rui Zhang, Jie Ma, and Amelia L. Gallitano

RESEARCH ARTICLE

Association of SNPs in *EGR3* and *ARC* with Schizophrenia Supports a Biological Pathway for Schizophrenia Risk

Matthew J. Huentelman¹, Leela Muppana², Jason J. Corneveaux^{1†}, Valentin Dinu³, Jeremy J. Pruzin¹, Rebecca Reiman¹, Cassie N. Borish¹, Matt De Both¹, Amber Ahmed², Alexandre Todorov⁴, C. Robert Cloninger⁴, Rui Zhang⁵, Jie Ma⁶, Amelia L. Gallitano^{2*}

1 Neurogenomics Division, Translational Genomics Research Institute, Phoenix, Arizona, United States of America, **2** Department of Basic Medical Sciences, University of Arizona College of Medicine—Phoenix, Phoenix, Arizona, United States of America, **3** Department of Biomedical Informatics, Arizona State University, Scottsdale, Arizona, United States of America, **4** Department of Psychiatry, Washington University School of Medicine, St. Louis, Missouri, United States of America, **5** Xi'an Hong Hui Hospital, the Affiliated Hospital of Xi'an Jiaotong University School of Medicine, Xi'an, Shaanxi, China, **6** Department of Biochemistry and Molecular Biology, Xi'an Jiaotong University School of Medicine, Xi'an, Shaanxi, China

† Deceased.

* amelia@email.arizona.edu



OPEN ACCESS

Citation: Huentelman MJ, Muppana L, Corneveaux JJ, Dinu V, Pruzin JJ, Reiman R, et al. (2015) Association of SNPs in *EGR3* and *ARC* with Schizophrenia Supports a Biological Pathway for Schizophrenia Risk. PLoS ONE 10(10): e0135076. doi:10.1371/journal.pone.0135076

Editor: Yong-Gang Yao, Kunming Institute of Zoology, Chinese Academy of Sciences, CHINA

Received: February 13, 2015

Accepted: July 17, 2015

Published: October 16, 2015

Copyright: © 2015 Huentelman et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are available from: <http://dx.doi.org/10.6084/m9.figshare.1562452>.

Funding: The collection of data and biomaterials for this study was supported by funding from NIH U01 MH046276 (to CRC). This work was supported by a Pfizer Fellowship in Biological Psychiatry, a National Alliance for Research on Schizophrenia and Depression/Sidney R Baer Jr. Foundation Young Investigator Award, an Arizona Biomedical Research Collaborative grant, NIH R01 MH097803 (to ALG), NIH R01 MH-083823 (to MJH), NIH R01 MH-083823

Abstract

We have previously hypothesized a biological pathway of activity-dependent synaptic plasticity proteins that addresses the dual genetic and environmental contributions to schizophrenia. Accordingly, variations in the immediate early gene *EGR3*, and its target *ARC*, should influence schizophrenia susceptibility. We used a pooled Next-Generation Sequencing approach to identify variants across these genes in U.S. populations of European (EU) and African (AA) descent. Three *EGR3* and one *ARC* SNP were selected and genotyped for validation, and three SNPs were tested for association in a replication cohort. In the EU group of 386 schizophrenia cases and 150 controls *EGR3* SNP rs1877670 and *ARC* SNP rs35900184 showed significant associations ($p = 0.0078$ and $p = 0.0275$, respectively). In the AA group of 185 cases and 50 controls, only the *ARC* SNP revealed significant association ($p = 0.0448$). The *ARC* SNP did not show association in the Han Chinese (CH) population. However, combining the EU, AA, and CH groups revealed a highly significant association of *ARC* SNP rs35900184 ($p = 2.353 \times 10^{-7}$; OR [95% CI] = 1.54 [1.310–1.820]). These findings support previously reported associations between *EGR3* and schizophrenia. Moreover, this is the first report associating an *ARC* SNP with schizophrenia and supports recent large-scale GWAS findings implicating the *ARC* complex in schizophrenia risk. These results support the need for further investigation of the proposed pathway of environmentally responsive, synaptic plasticity-related, schizophrenia genes.

(to AT), NIH R01 MH-067921 (PI-Todd RD), an Institute of Mental Health Research award (to ALG and MJH), and the State of Arizona Department of Health Services. Additional support was from the Fundamental Research Funds for the Central Universities (No.2012JDGZ07, No.XJJ2012124) and the Ministry of Education, PRC (No.20090201120062) (to JM and RZ). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: This research was supported in part by the 2004 Pfizer Fellowship in Biological Psychiatry, funded by Pfizer U.S. Pharmaceuticals, awarded to Amelia Gallitano. Dr. Gallitano was not an employee or consultant of Pfizer, and the science was performed independently without influence or direction from the company and did not involve product development for, or marketing of products of, the company. There are no restrictions on the sharing of data or materials. This does not alter the authors' adherence to PLOS ONE policies on sharing data and materials.

Introduction

Schizophrenia is a severe mental illness that affects approximately 1% of the world's population and is determined by both genetic and environmental factors [1, 2]. The genetic contribution to schizophrenia risk ranges from 50–80% [1, 3]. The remaining influence may be attributed to environmental factors [2]. However, *how* environmental exposures interact with genetic variations to influence schizophrenia susceptibility remains unclear.

Immediate early gene (IEG) transcription factors are rapidly activated in the brain in response to environmental stimuli and regulate downstream neuronal gene expression. The early growth response gene (*EGR*) family of IEG transcription factors is involved in numerous neural processes, including regulation of synaptic proteins and synaptic plasticity, dysfunction of which have been hypothesized to play a role in schizophrenia pathogenesis [4–10].

Numerous findings suggest a role for the IEG *EGR3* in schizophrenia susceptibility. Single nucleotide polymorphisms (SNPs) in *EGR3* are associated with schizophrenia in Japanese, Korean, and Han Chinese (CH) populations [11–13], and levels of *EGR3* mRNA are reduced in post-mortem brain tissue from schizophrenia patients [11, 14]. We have previously reported that *Egr3*-deficient (*Egr3*KO) mice display schizophrenia-like behavioral abnormalities that are reversible with antipsychotic treatment [4, 15]. *EGR3* is regulated downstream of several key schizophrenia candidate proteins, including Neuregulin 1, N-methyl D-Aspartate Receptors (NMDARs) and calcineurin (CN) [11, 16–21], and *Egr3*KO mice share key phenotypes with NMDARKO and CNKO mice, including memory dysfunction and deficits in the form of hippocampal synaptic plasticity long-term depression (LTD) [4, 22, 23]. These shared characteristics suggested to us that these proteins may form a biological pathway of genes which, when disrupted may increase risk for schizophrenia. If so, then dysfunction of the gene activity-regulated cytoskeleton-associated protein (ARC) [24], which is a downstream target of *EGR3* and is similarly required for memory formation and hippocampal LTD in mice [25–27], may also confer risk for schizophrenia.

In support of this hypothesis, several large-scale studies have identified associations between schizophrenia and genes whose proteins complex with ARC in the post-synaptic density of neurons [28–30]. However, to date there have been no reports of genetic association between SNPs in the *ARC* gene itself and schizophrenia. In addition, significant associations between SNPs in *EGR3* and schizophrenia have not been reported in populations of European or African descent. We therefore carried out the current studies to test the hypothesis that *EGR3* and *ARC* are schizophrenia risk associated genes.

Materials and Methods

DNA Samples

DNAs from EU and AA patients diagnosed with schizophrenia based on DSM-IV classification were obtained from the National Institute of Mental Health Schizophrenia Genetics Initiative (NIMH-SGI) Release 1.0 and Release 3.0. Control samples were obtained from the NIMH-GI Control Database Release 3.0 and the Environmental Catchment Area ADHD Study. Additional control genotypes were obtained from the 1,000 genomes published findings [31]. Written informed consent was obtained from subjects in all studies. These samples were obtained in a de-identified manner and were thus exempt from approval by the University of Arizona Institutional Review Board (IRB).

The study involving CH samples was approved by the genetic research ethics committee (equivalent to the IRB) of Xi'an Jiaotong University School of Medicine and written informed consent was obtained from all subjects. A total of 982 subjects included 491 schizophrenics

(253 males, mean age = 34.49 ± 11.85 , age of onset = 24.30 ± 7.22 ; 238 females, mean age = 31.77 ± 13.50 , age of onset = 24.42 ± 8.64) and 491 normal controls (272 males, mean age = 28.92 ± 13.78 ; 219 females, mean age = 28.88 ± 13.83). All patients were diagnosed by the psychiatrists of the First Affiliated Hospital of Xi'an Jiaotong University School of Medicine according to *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV) criteria for schizophrenia using a combination of examination of psychiatric case records and clinical interview. The diagnosis was checked and verified by two independent senior psychiatrists who reviewed the psychiatric case records.

CH controls were drawn from a combination of local volunteers or blood transfusion donors. Subjects with current or past evidence of mental illness, or with a first-degree relative with mental illness, were not included in the cohort. In addition, in China only healthy people not taking medication can donate blood, hence the evaluated controls were medication-free. Overall, these controls are unlikely to include a significant number of psychotic individuals, if indeed any. All subjects were *Han* Chinese in origin and from Northwest China.

Next-Generation Sequencing (NGS)

Amplification and sample preparation. Pooled DNA samples were PCR amplified using RainStorm (RainDance Technologies, Inc. Billerica, MA) essentially as described in [32]. Briefly, individual DNA samples were quantitated in triplicate with PicoGreen reagent (Life Technologies, Grand Island, NY) and equimolar pools were created based on sex, race, and diagnosis. Individual DNAs were fragmented to 2–4 kb lengths and merged with a panel of primers designed to create overlapping amplicons spanning the entire coding and noncoding sequence of *EGR3* and *ARC*, from 7 Kb upstream of the 5'UTR to 3 Kb downstream of the 3'UTR. See [S1 Table](#) for primer sequences. RainStorm PCR was performed and amplification products were concatenated using the New England Biolabs (NEB) Quick blunting and ligation kit (Ipswich, MA).

Amplicons were sonicated for 2.5 minutes to produce 100–400 bp DNA fragments, treated with Klenow enzyme Fragment for 30 minutes, blunt-ended with T4 DNA polymerase and T4 polynucleotide kinase, and purified using the Ampure SPRI bead purification system (Beckman Coulter, CA). Fragments were 3' end-labeled with an untemplated A base (Klenow 3'-5' exo- and dATP for 30') followed by Ampure purification and PicoGreen quantification (Invitrogen, Inc., NY). NGS universal adapters were modified with unique DNA bar-codes as previously described [33]. Ligation was performed using 10:1 molar ratio of bar-coded adapters: DNA fragments using high concentration T4 DNA ligase for 15'. The ligation product underwent gel purification (run on a 3% agarose gel at 100 v for 2 h, the 300 bp band was excised and purified using Freeze'N Squeeze (Bio-Rad Laboratories, Inc., CA) columns) and PCR amplification. A second round of gel purification isolated the 200bp–350bp fragment to form a DNA library bar-coded for each individual DNA sample.

NGS of pooled bar-coded DNAs. Individual bar-coded DNA libraries were combined into pools and added to a sequencing flowcell to generate clusters that were subjected to NGS. The molecular bar-coding approach facilitated sequencing of DNA samples from multiple pools on each lane of the flowcell.

NGS data analysis and variant detection. Data from the Illumina Genome Analyzer was converted to text-based sequence with the Illumina Pipeline 1.4 and CASAVA 1.0 software using standard Illumina protocols. Image cluster identification, sharpening, background subtraction, and intensity extraction were performed with the Firecrest module. Base calling and phasing/prephasing effects were corrected for using the Bustard module. Using the pass/fail filter generated by Bustard, Sanger Phred-33 FASTQ files were created with a custom perl script

in preparation for alignment. FASTQ format was chosen because each sequence read is described with the sequence and individual base quality scores, metadata from the instrument, cluster position and paired-read info, and many of the commonly used aligners use the FASTQ format.

NGS data were aligned with the reference genome using the BWA aligner [34], the tool used by the 1000 Genomes project, which uses SAM (Sequence Alignment/Map) format [35]. The SAM format contains information on mapping quality, mapping coordinates, mismatches, and fields for user specified metadata. BWA alignment was carried out using default modes for 36 bp single and paired read (72 bp total) runs. Reads were mapped to the reference human genome used by the 1000 Genomes project [31]. After alignment, variants were called with `VarScan.v2.2.jar pileup2snp` with the options `—min-coverage 250—min-reads2 10—min-avg-qual 20—min-var-freq 0.02—p-value 0.01`.

Genotyping of selected SNPs

DNA samples from the NGS discovery cohort were genotyped for three *EGR3* SNPs and one *ARC* SNP for the validation phase association analyses. Additional schizophrenia case DNAs for the replication association study were obtained from NIMH-SGI release 3.0. DNAs were quantitated by nanodrop spectrophotometry, diluted in sterile Tris buffer, 10mM, pH 7.4, to a concentration of 4 ng/ul, and confirmed using PicoGreen (Invitrogen, Inc., NY) quantification, according to the manufacturer's protocol. Samples were plated on to 96 well plates at 20 ng/well, and dried down overnight at room temperature in the dark. DNAs were genotyped using individual SNP assays from Applied Biosystems Inc., (ABI, Foster City, CA) on an ABI 7500 Fast QRT-PCR instrument following the manufacturer's protocol.

Genotyping of CH samples was performed using direct DNA sequence analysis. The PCR primers for amplification of the rs35900184 were as follows: forward primer 5'- CGCCTGGA GAAGAATCAGAG-3' and reverse primer 5'- AAAGACTTTGTGGGAACCTTGA-3'. The PCR reaction was performed in 25 μ l of standard PCR buffer containing 1.5mM MgCl₂, 0.2mM of each dNTP, 0.5 μ l of each primer, 1 unit of Taq DNA polymerase, and 25 ng of human genomic DNA. The program was one cycle of 2min for denaturation at 95°C, 35 cycles of 30s at °C, 35s at 58°C, 45s at 72°C, and one 7min extension step at 72°C. Purified PCR products were sequenced bi-directionally using PCR primers as sequencing primers and the Applied Biosystems Prism BigDye terminator cycle sequencing reaction kit. The products were evaluated on an ABI 3730 Genetic Analyzer (Applied Biosystems).

Individual Genotype Analysis

Genotype data were analyzed using PLINK v1.07 [36], and the Fisher test with 95% confidence intervals.

Linkage Disequilibrium (LD) Analyses

Haploview [37] was used to estimate LD between SNPs across the *ARC* region in the CEPH population (Utah residents with ancestry from northern and western Europe, abbreviated CEU), Hapmap [38] data release 28. The *ARC* region (chromosome 8, from 143,689,412 to 143,692,835) included 18 SNPs in the CEU Hapmap individuals. Hapmap Genome Browser release 28 (<http://hapmap.ncbi.nlm.nih.gov/>) with Copy Number Variation Region Track enabled, was used to identify known CNVs in the *ARC* region on chromosome 8.

Power Analyses

Calculating power for genetic studies requires knowledge of values such as the prevalence of the disease, the relative risk of the disease risk allele (the ratio of the probability of disease occurrence in the risk-allele carriers to the probability of disease occurrence in individuals that are not carriers of the risk-allele) and the frequency of the disease allele. Since these parameters are unknown for the current study, estimates were used, and the disease risk allele was assumed to be included in our sample. Assuming values of 0.01 for disease prevalence, relative-risk of 1.5 (a relatively optimistic value for a complex disease such as schizophrenia), alpha-level of 0.05 and a standard 1df allelic test, then the estimated power to detect a significant association ranges from 0.05 to 0.47 (for disease allele frequencies ranging from 0.1 to 0.9) for a sample size of 386 cases and 150 controls, as was our EU cohort. For our AA cohort (sample size of 185 and 50 controls) the power range is lower, from 0.05 to 0.21 only. These parameter ranges indicate that our sample size of fewer than 200 cases and 50 controls is underpowered to detect an association in the SNPs in the AA cohort, should one truly exist. Increasing the number of cases and controls to $N = 1,200$, and 300, respectively, would increase the power to detect a positive association to 0.80, for disease allele frequencies of 0.2 (which approximates the MAF for rs1877670 in AA). The Genetic Power Calculator was used to generate these power estimates [39].

Results

Since the frequencies of specific SNPs, as well as their potential association with an illness, vary among ethnic and racial groups, it can be difficult to select the SNPs that have a high likelihood of showing association within the population of study. In studies with limited sample sizes limiting the number of SNPs queried helps to maintain statistical power. To address these issues we employed a method using NGS to identify all variations across the *EGR3* and *ARC* loci from a discovery cohort of pooled DNA samples. SNPs in which one allele demonstrates a high degree of difference in frequency between cases and controls, particularly in two separate populations, may be more likely to show association with illness.

NGS of the *EGR3* and *ARC* loci

To identify variations within the *EGR3* and *ARC* loci that have a high likelihood for association with schizophrenia risk in two different racial groups, we used a two-step approach. For the initial step we used a “Discovery” cohort of DNA samples from schizophrenia patients (cases) and non-psychiatrically ill individuals (controls) that were matched for sex and race (EU and AA). [Table 1](#) shows how these samples were pooled into 16 groups of up to 25 samples per group. NGS across the *EGR3* and *ARC* loci was performed on each pool and results were compared among pools for quality control.

Allele frequencies were estimated for each variant along the entire sequenced region in each pooled group using CRISP [40]. The results were averaged among pools according to case/control status and race (sex was not maintained as a variable). For each variant the difference in the average frequency of the minor allele between cases and controls, termed the “delta Minor Allele Frequency” (Δ MAF), was calculated by subtracting the average MAF for schizophrenia patients from that for controls of the same race. [S2](#) and [S3](#) Tables show the allele frequency results for all variations identified in the *EGR3* region, ranked from highest to lowest Δ MAF, in the EU and AA groups. [S4](#) and [S5](#) Tables show the same analysis for the sequenced *ARC* region in the respective EU and AA groups.

The Δ MAF results for the sequenced regions of the EU (in blue) and AA (in red) populations were plotted onto the genome map in the regions of *EGR3* ([Fig 1A](#)) and *ARC* ([Fig 1B](#))

Table 1. Characteristics of DNA Pools Used for Next Generation Sequencing.

Race	Sex	Affected Status	No. of Subjects/Pool	No. of Pools	Total No. of Subjects
EU	F	Cases	25	2	50
	F	Controls	25	2	50
	M	Cases	23–25	4	95
	M	Controls	25	4	100
AA	F	Cases	18	1	18
	F	Controls	25	1	25
	M	Cases	24	1	24
	M	Controls	25	1	25

doi:10.1371/journal.pone.0135076.t001

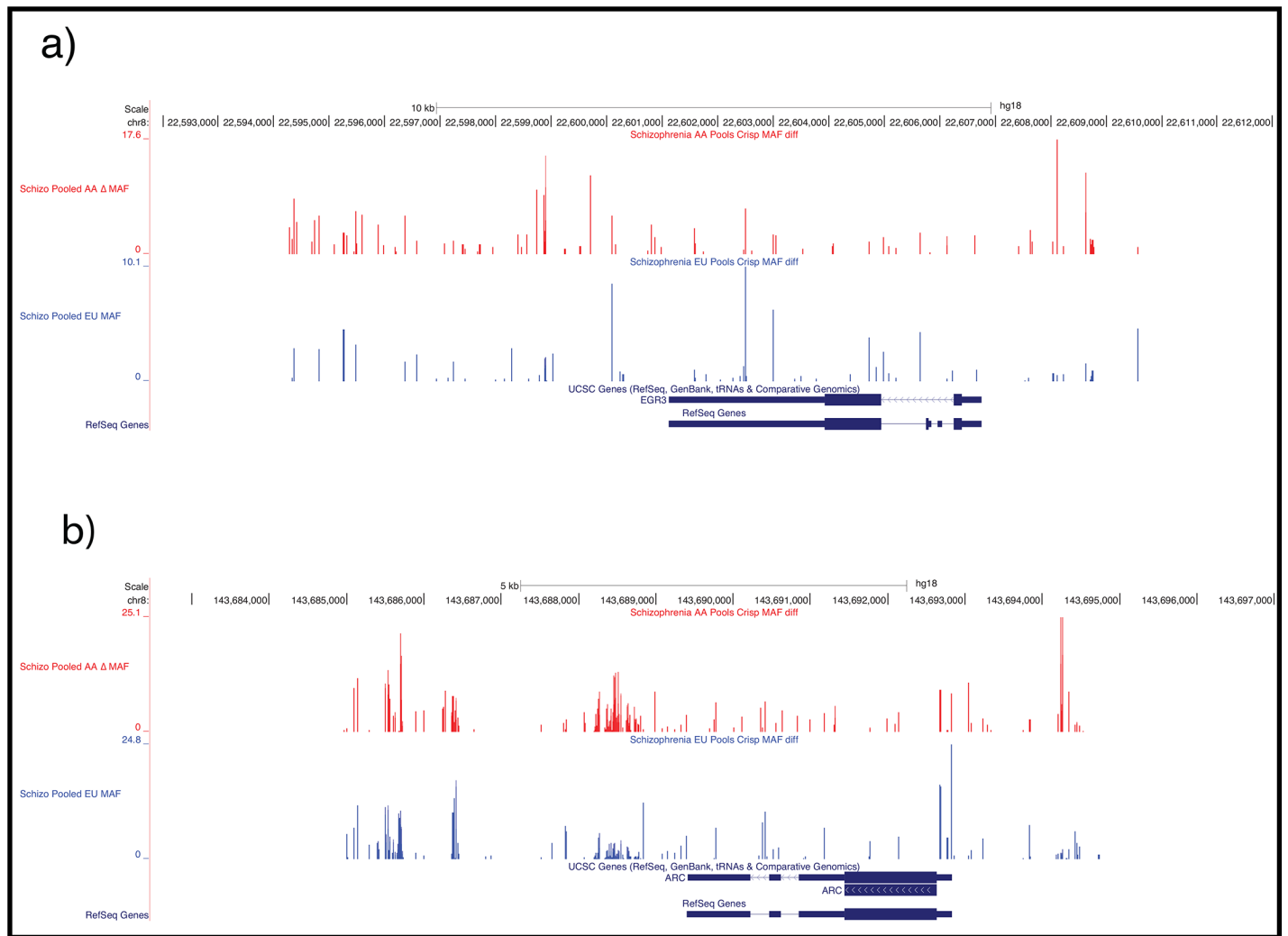


Fig 1. Minor Allele Frequency Differences in EU and AA. Results of polymorphisms identified by NGS are mapped to the reference human genome used by the 1000 Genomes project (hg18). Vertical lines indicate the Δ MAF (difference between the minor allele frequency in cases versus controls) for each base pair of the genome region investigated. Higher peaks indicate a greater difference in prevalence of that variation between cases and controls, and thus a higher potential for association with schizophrenia risk. The red graph indicates data for Whites and the Blue graph indicates data for African Americans. (A) Δ MAF map for the *EGR3* locus, hg18 coordinates chr8:22,591,791–22,612,066. (B) Δ MAF map for the *ARC* locus, hg18 coordinates chr8:143,684,474–143,697,026

doi:10.1371/journal.pone.0135076.g001

using the UCSC genome browser. Individual variations that produced high peaks (i.e. greater Δ MAF) in both racial groups were of greatest interest.

Case-Control Genetic Association Analyses

Several SNPs were selected for validation of the NGS results by genotyping of the individual DNA samples from the discovery cohort. Variations for genotyping validation were selected based on a combination of criteria, including high Δ MAF values in both race cohorts and/or having been reported as associated with schizophrenia in other populations. Three SNPs from the *EGR3* region were selected. The SNP **rs1877670** displayed the highest Δ MAFs in our EU cohort, and ranked eighth in the AA cohort, with a Δ MAF of 7.0%. This SNP was also reported in a haplotype associated with schizophrenia in a CH population [41]. **rs1996147** demonstrated Δ MAFs of 8.6% in EUs and 5.9% in AAs in our NGS discovery analysis, and showed nominally significant association with bipolar disorder in a study of circadian rhythm genes [42]. The SNP **rs10095121** showed a Δ MAFs of 2.9% in EUs and 8.4% in AAs in our NGS results, and revealed nominally significant association with the diagnosis of child bipolar I disorder in our prior family based association study [43].

Since no studies had been published investigating *ARC* SNPs for association with schizophrenia, selection of a single *ARC* SNP was made based only on high Δ MAF values in both race cohorts. The SNP rs28420666 was initially selected for genotyping as it demonstrated Δ MAF of 24.8% in the EU cohort 8.3% in the AA cohort. However our attempts to create a genotyping assay for this SNP were unsuccessful. Therefore SNP **rs35900184**, which showed Δ MAFs of 10.2% in EUs and 6.6% in AAs, was selected.

Tables 2 and 3 show the results of the validation phase. Of note, the pooled sequencing approach yielded MAFs that were similar to the confirmed individual-based results (e.g. comparing MAF Controls versus MAF Pooled Controls columns in Tables 2 and 3). In the EU group *EGR3* SNP rs1877670 showed a trend toward association ($p = 0.07$), and the *ARC* SNP rs35900184 revealed a significant association with the schizophrenia diagnosis ($p = 0.02$). In the AA cohort, there were no significant associations between the three *EGR3* or the single *ARC* SNP tested, though the *ARC* SNP rs1996147 showed a trend toward significance ($p = 0.07$). We selected the three SNPs that showed significance or trended toward significance for further evaluation in a replication group.

A replication cohort was generated by inclusion of additional cases to the samples from the discovery cohort. No additional controls were included due to lack of availability. The results of association analyses for the two *EGR3* and one *ARC* SNP genotyped for the replication cohort are shown in Tables 4 and 5. In the EU replication study *EGR3* SNP rs1877670 showed a significant association with schizophrenia ($p = 0.0078$). The *ARC* SNP rs35900184 maintained a significant association with the schizophrenia diagnosis ($p = 0.0275$) in the replication study. In the AA cohort only the *ARC* SNP rs35900184 demonstrated significance ($p = 0.045$).

An additional investigation of “Population Controls” from the publicly available 1,000 genomes data was performed [31]. This demonstrated the allele frequency of the four SNPs of interest in presumed controls denoted as “GMAF” in Tables 2–8. These data show that the population controls are in approximately the same allele frequency range as the clinical controls, suggesting that these variants are indeed significantly enriched in the white schizophrenia cases.

ARC Association Study Han Chinese Sample

To test our hypothesis that *ARC*, a direct gene target of *EGR3*, is also a schizophrenia risk gene, we genotyped rs35900184 in a CH population of 491 cases and 491 controls in which *EGR3* had previously demonstrated association with schizophrenia [13]. Table 6 shows that the

Table 2. Validation Phase: Comparison of Pooled Sample NGS with Individual Genotyping Results and Genetic Association Analyses in EU cohort. Pooled EU.

SNP	Chr	Hg18 Position	A1	A2	MAF Controls	MAF Cases	MAF Pooled Controls	MAF Pooled Cases	GMAF	P-value	OR [95% CI]
rs10095121	8	22,538,426	C	T	33.6%	32.8%	32.6%	29.8%	28.4%	0.8566	0.970 [0.680–1.380]
rs1877670	8	22,546,561	C	T	45.8%	54.1%	42.3%	47.6%	44.3%	0.0685	1.390 [0.990–1.960]
rs1996147	8	22,544,158	G	A	35.5%	32.7%	33.5%	25.0%	41.2%	0.5307	0.890 [0.620–1.260]
rs35900184	8	143,693,411	T	C	22.9%	32.3%	18.0%	28.1%	26.5%	0.0215	1.610 [1.090–2.390]

145 cases, 150 controls; 195 males, 100 females.

doi:10.1371/journal.pone.0135076.t002

minor allele of rs35900184 showed no significant difference in frequency between cases and controls ($p = 0.293$) in this CH cohort.

Combined analysis of ARC and EGR3

To increase power and to investigate the role of our candidate variants independent of ethnicity we performed a combined analysis of the SNPs (Tables 7 and 8). The EGR3 SNPs were not significant in the combined group of EU and AA samples (the CH samples were not included as EGR3 has previously been investigated for associations in this group [13]). However, the ARC SNP, rs35900184, was significantly associated with schizophrenia in the combined group of EU, AA, and CH samples ($p = 2.5 \times 10^{-7}$), with a suggested odds ratio of 1.54 (95% confidence interval = 1.31–1.82) (Table 8). We also assessed the heterogeneity of our three candidate SNPs across the study populations using a Breslow-Day test. The results of this indicated significant heterogeneity at SNPs rs1877670 ($p = 0.024$) and rs35900184 ($p = 4.5 \times 10^{-3}$) (see S6 Table). To address any potential confounds from this heterogeneity we performed a meta-analysis using a random effects model in PLINK (see S7 Table). This resulted in a significant association noted in the ARC SNP only (rs35900184, $p = 1.9 \times 10^{-4}$).

Variations in the ARC Region

A 30 Kb region of chromosome 8 spanning the ARC locus was interrogated using HapMap and NCBI dbSNP resources to identify known SNPs and analyze LD in this region. (A similar

Table 3. Validation Phase: Comparison of Pooled Sample NGS with Individual Genotyping Results and Genetic Association Analyses in AA cohort. Pooled AA.

SNP	Chr	Hg18 Position	A1	A2	MAF Controls	MAF Cases	MAF Pooled Controls	MAF Pooled Cases	GMAF	P-value	OR [95% CI]
rs10095121	8	22,538,426	C	T	26.1%	24.4%	31.8%	20.0%	28.4%	0.8621	0.910 [0.460–1.820]
rs1877670	8	22,546,561	C	T	22.3%	14.6%	29.0%	19.0%	44.3%	0.2459	0.600 [0.270–1.300]
rs1996147	8	22,544,158	G	A	29.3%	17.1%	22.6%	15.3%	41.2%	0.0734	0.500 [0.240–1.030]
rs35900184	8	143,693,411	T	C	31.9%	41.2%	25.1%	34.0%	26.5%	0.2101	1.500 [0.800–2.790]

42 cases, 50 controls; 49 males, 43 females.

doi:10.1371/journal.pone.0135076.t003

Table 4. Replication Phase: Genetic Association Between Schizophrenia and SNPs in EGR3 and ARC in EU Population. Replication EU.

SNP	Chr	Hg18 Position	A1	A2	MAF Controls	MAF Cases	MAF Pooled Controls	MAF Pooled Cases	GMAF	P-value	OR [95% CI]
rs1877670	8	22,546,561	C	T	45.8%	55.6%	42.3%	47.6%	44.3%	0.0078	1.480 [1.110–1.960]
rs1996147	8	22,544,158	G	A	35.5%	32.5%	33.5%	25.0%	41.2%	0.3745	0.880 [0.660–1.170]
rs35900184	8	143,693,411	T	C	22.9%	30.0%	18.0%	28.1%	26.5%	0.0275	1.450 [1.050–2.00]

386 cases, 150 controls; 363 males, 173 females.

doi:10.1371/journal.pone.0135076.t004

analysis of the *EGR3* region has previously been published [43].) The *ARC* SNP rs35900184 is not included in either of these public resources. The HapMap project includes 18 SNPs in this region, which have been genotyped in 11 populations (S1A Fig). Genotype data from the CEPH population (Utah residents with ancestry from northern and western Europe, abbreviated CEU) was used to create a linkage disequilibrium map in Haploview, which revealed one haplotype block in this region (S1B Fig). The *ARC* SNP rs35900184 is located outside of the haplotype block in a region that is poorly genotyped in HapMap. Analysis of copy number variation (CNV) domains revealed that *ARC* resides within a CNV region (S1C Fig).

Discussion

While numerous studies have identified candidate genes and genomic variations linked to risk for psychiatric illnesses, few of them provide an explanation for the role of environment. We have previously proposed that proteins acting in a biological cascade required for memory formation and hippocampal LTD, comprise a pathway that, when disrupted, increases risk for schizophrenia [4, 15, 44]. This pathway, triggered by activity-dependent opening of NMDARs, allows neuronal calcium influx, which acts on CN, a Ca⁺⁺ activated phosphatase, to activate expression of the IEG transcription factor *EGR3* [19, 21]. The gene *ARC* is a direct target of *EGR3* protein [24], and is also required for memory formation and LTD [4]. We therefore hypothesized that, like the preceding proteins in this putative pathway, which have each been implicated in schizophrenia risk, *ARC* should also be a schizophrenia susceptibility gene. Such a collection of risk genes into one functional pathway can address both the polygenic nature of, and environmental contribution to, schizophrenia risk.

In the current study we asked the whether genomic variations in the environmentally-activated IEGs *EGR3* and *ARC* are associated with schizophrenia in two racial populations, EU

Table 5. Replication Phase: Genetic Association Between Schizophrenia and SNPs in EGR3 and ARC in AA Population. Replication AA.

SNP	Chr	Hg18 Position	A1	A2	MAF Controls	MAF Cases	MAF Pooled Controls	MAF Pooled Cases	GMAF	P-value	OR [95% CI]
rs1877670	8	22,546,561	C	T	22.3%	20.5%	29.0%	19.0%	44.3%	0.6716	0.900 [0.520–1.550]
rs1996147	8	22,544,158	G	A	29.3%	25.1%	22.6%	15.3%	41.2%	0.4251	0.810 [0.490–1.340]
rs35900184	8	143,693,411	T	C	31.9%	43.7%	25.1%	34.0%	26.5%	0.0448	1.650 [1.020–2.680]

185 cases, 50 controls; 127 males, 108 females.

doi:10.1371/journal.pone.0135076.t005

Table 6. Replication of ARC SNP Association with Schizophrenia in Han Chinese Population. Replication CH ARC.

SNP	Chr	Hg18 Position	A1	A2	MAF Controls	MAF Cases	GMAF	P-value	OR [95% CI]
rs35900184	8	143,693,411	T	C	17.3%	19.2%	26.5%	0.2933	1.140 [0.910–1.430]

491 cases, 491 controls; 525 males, 457 females.

doi:10.1371/journal.pone.0135076.t006

and AA. We used NGS to comprehensively evaluate the frequency of variations present at the *EGR3* and *ARC* loci in both cases and controls. This allowed identification of multiple types of polymorphisms, including rare variants, providing a much greater depth of investigation than a classical SNP genotyping approach, which is limited to detection of SNPs occurring at >10% frequency in the population. We hypothesized that this would be a more effective way to select variations for genotyping than choosing known SNPs randomly across the genomic region for an association study. It also allowed rapid and broad comparison of variation frequencies among racial populations. We indicate the concordance in MAF of our candidate variants in Tables 2 and 3. This demonstrates that the pooled sequencing approach is highly comparable to the results obtained on the same individuals via genotyping.

Our major findings were that *EGR3* SNP rs1877670 was associated with schizophrenia in our population of EU cases and controls ($p = 0.0078$; Table 4). And the *ARC* SNP rs35900184 displayed association with schizophrenia risk in both the EU and the AA populations ($p = 0.0275$ and $p = 0.0448$, respectively; Tables 4 and 5). Further, although *ARC* SNP rs35900184 did not show statistically significant association with schizophrenia in the CH population alone (Table 6), when all three racial groups were combined to increase statistical power, rs35900184 revealed a strongly significant association with schizophrenia ($p = 2.353 \times 10^{-7}$; OR [95% CI] = 1.54 [1.310–1.820]; Table 8). These findings support our hypothesis of a biological pathway involving memory, hippocampal LTD, and schizophrenia susceptibility [4, 15, 44], by demonstrating that *ARC*, the downstream target of proteins in the pathway that are associated with risk for schizophrenia, is itself associated with risk for this illness.

To our knowledge, this is the first study examining association between a SNP in *ARC* and schizophrenia. It is also the first study reporting associations between SNPs in *EGR3* and schizophrenia in an AA population. Recent high-profile studies of rare copy number variations (CNVs) and *de novo* variations enriched in schizophrenia patients have identified the importance of proteins that complex with *ARC* protein in the post-synaptic density in the risk for schizophrenia [28–30, 45]. Small *de novo* mutations in genes encoding proteins that associate with *ARC* are overrepresented in schizophrenia patients versus controls [30]. And exome sequencing has identified rare disruptive mutations involving proteins that associate with *ARC* are more prevalent in schizophrenia patients than controls [29]. However, these large-scale genome-wide studies did not identify associated variations in the *ARC* gene itself. Notably, some of these were exome studies, so would not have included the regulatory region in which the *ARC* SNP rs35900184 resides.

Table 7. EGR3 SNP Association Analyses in All Populations Combined. All Samples Combined *EGR3*.

SNP	Chr	Hg18 Position	A1	A2	MAF Controls	MAF Cases	GMAF	P-value	OR [95% CI]
rs1877670	8	22,546,561	C	T	39.7%	44.1%	44.3%	0.1405	1.200 [0.940–1.530]
rs1996147	8	22,544,158	G	A	34.0%	30.1%	41.2%	0.1748	0.840 [0.0650–1.080]

571 cases, 200 controls; 490 males, 281 females.

doi:10.1371/journal.pone.0135076.t007

Table 8. ARC SNP Association Analysis in All Populations Combined. All Samples Combined ARC.

SNP	Chr	Hg18 Position	A1	A2	MAF Controls	MAF Cases	GMAF	P-value	OR [95% CI]
rs35900184	8	143,693,411	T	C	19.5%	27.2%	26.5%	2.353e-07	1.540 [1.310–1.820]

1062 cases, 691 controls; 1015 males, 735 females.

doi:10.1371/journal.pone.0135076.t008

To date, three case-control and one family-based study have identified significant associations between SNPs in *EGR3* and schizophrenia [11–13, 41]. An additional family study revealed nominally significant transmission disequilibrium for *EGR3* [11]. All of these studies were in Asian populations, including Korean, Japanese, and CH. In all three case-control studies the same SNP, rs35201266, demonstrated association with schizophrenia [11–13]. In the CH family study rs1996147 and rs3750192 showed association with schizophrenia, as did the haplotype rs3750192-rs35201266 [41]. Other case-control studies in the Japanese and CH populations failed to identify significant associations between SNPs in *EGR3* and schizophrenia [46–48], though only one of these three studies included rs35201266 [46].

EGR3 SNP rs35201266 has also been associated with prefrontal hemodynamic response to a verbal fluency test in both control and schizophrenia subjects of Japanese ethnicity, suggesting a relationship between *EGR3* and cognitive function in a brain region implicated in schizophrenia [49]. rs35201266 resides in the single intron of *EGR3*, suggesting a possible functional role in *EGR3* mRNA expression or post-transcriptional processing. *In vitro* analysis of this SNP demonstrated a potential functional effect on transcription [11]. However, we did not genotype rs35201266 in our samples as our NGS results demonstrated a Δ MAF of 4.3% in the EU group and 3.2% in the AA group for this SNP, which did not meet our cut-off of scoring $> 5\%$ in at least one group.

Only one study examining *EGR3* associations in schizophrenia in a EU population has been reported. Although that study, which focused on circadian rhythm-related genes, failed to identify an association between SNPs in *EGR3* and schizophrenia, it did report a suggestive association between *EGR3* and bipolar I disorder ($p = 0.017$), though this did not survive control for multiple comparisons [42]. Nominal associations were also found between two linked *EGR3* SNPs, rs10104039 and rs10095121, and child bipolar I disorder in a family based association study of primarily EU participants (88% EU; $p = 0.027$ and $p = 0.028$) [43]. Numerous findings, including reports from genome wide association studies (GWAS), indicating shared genetic influences on both bipolar disorder and schizophrenia, highlight the potential relevance of these findings for our study [50–52].

In addition to the genetics findings, post mortem studies have reported reduced *EGR3* mRNA levels in schizophrenia patients' brains [11, 14], and a biomedical informatics study identified *EGR3* as the central gene in a network of microRNAs and transcription factors implicated in schizophrenia pathogenesis [53]. *EGR3* also resides in a copy number variation (CNV)-containing region of the chromosome [43], and CNVs have demonstrated a high correlation with schizophrenia and other psychiatric illnesses [54].

The 8p chromosomal region, and indeed the specific location of *EGR3* at 8p21.3, have been repeatedly associated with schizophrenia susceptibility, specifically in the AA population [55–58]. The importance of this region in schizophrenia is also supported by meta-analysis of 32 Genome Wide Linkage studies [59] and rank-based genome scan meta-analysis from 20 schizophrenia genome scans [60]. Although our study did not demonstrate association between schizophrenia and *EGR3* SNPs rs1877670 or rs1996147 in the AA cohort, the limited number of available AA DNA samples led to the study being insufficiently powered to detect

variations of low effect size. Interestingly, rates of schizophrenia have been reported to be up to three times higher in AA than EU individuals, which may be due to differences in both genetic and environmental causes [61].

The major limitation of our study was the number of samples available, particularly in the AA group. Because of this, the DNA samples used for the discovery NGS study were included in the validation phase. This was done for two reasons. First, the initial discovery phase used a pooling approach that allowed us to cost-effectively sequence the region with a very high copy number (providing a high level of accuracy), but did not discern individual genotype results (only an average for the pool). Follow-up genotyping was therefore necessary to conduct association analyses. Second, the number of additional available samples was insufficient to power an independent study. Inclusion of the original discovery set of DNAs in the genotype analysis allowed us to increase our number of samples, which was still of a size that limited the power of the study.

Lastly, it is important to note that the identified variants do not immediately suggest any potential biological function; therefore, it is possible that they are closely linked with neighboring variants that may influence function of the underlying proteins. For example, rs1877670 falls within the 3' UTR of *EGR3*, but doesn't overlap with any known regulatory regions in the gene, rs1996147 is located ~143kb downstream of *EGR3* in the current build of the human genome (GRCh37), and rs35900184 falls within an intron between exons 2 and 3 of *ARC* and doesn't overlap any known regulatory regions. Therefore the putative functional role of these variants is in need of additional investigation.

Conclusions

We report that the *EGR3* SNP rs1877670 is associated with schizophrenia in an EU population. In addition, the *ARC* SNP rs1877670 shows association in both the EU and AA populations, and is highly significant in a combined group of EU, AA, and CH populations. These findings support our hypothesis that dysfunction of genes in the biological pathway of NMDA-receptors to *EGR3* regulation of *ARC* may increase schizophrenia susceptibility. This pathway represents a molecular link between environment and the post-synaptic density proteins recently implicated in schizophrenia risk by CNV and GWAS studies [28–30], and may provide insight into the cognitive deficits characteristic of this illness.

Supporting Information

S1 Fig. NCBI and HapMap Data on Variations in the ARC Region. A 30 Kb region of chromosome 8, spanning from 143676124 to 143706123, which contains the *ARC* gene, was interrogated. **a.** HapMap data version 28 includes 18 SNPs genotyped in this region. Graphics below the chromosome line indicate percentages of each allele for the SNP (indicated in blue versus red) for each racial-ethnic population genotyped; rs numbers below the HapMap results show the 126 SNPs reported in this region on NCBI B36, dbSNP. They do not include rs35900184. Yellow highlight indicates the location of the *ARC* gene. **b.** Linkage disequilibrium map created in Haploview using the CEU genotype data for SNPs in this region identified a single haplotype block. *ARC* (chr8: 143,689,412 to 143,692,835) is located between SNP 16 (rs13260813 A/C chr8 143,689,397) and SNP 18 (rs28473387 C/T chr8 143,704,475), in a region that is poorly genotyped in HapMap CEU data. *ARC* SNP: rs35900184 (8:143,690,413), is located between SNP 16 (rs13260813 A/C; chr8 143,689,397) and SNP 17 (rs10097505 A/G; chr8 143,691,186, an *ARC* SNP) in the middle of a region of poor LD. **c.** The *ARC* gene resides in a copy number variation region.

(PDF)

S1 Table. Primer Sequences Used to Generate PCR Amplicons from Pooled DNAs for Next Generation Sequencing.

(XLSX)

S2 Table. Variations Identified in *EGR3* by Next Generation Sequencing from EU DNA Pools. Bold type indicates SNPs selected for follow-up genetic association analysis.

(XLSX)

S3 Table. Variations Identified in *EGR3* by Next Generation Sequencing from AA DNA Pools. Bold type indicates SNPs selected for follow-up genetic association analysis.

(XLSX)

S4 Table. Variations Identified in *ARC* by Next Generation Sequencing from EU DNA Pools. Bold type indicates SNPs selected for follow-up genetic association analysis.

(XLSX)

S5 Table. Variations Identified in *ARC* by Next Generation Sequencing from AA DNA Pools. Bold type indicates SNPs selected for follow-up genetic association analysis.

(XLSX)

S6 Table. Breslow-Day test for heterogeneity between AA, EU, and CH populations. These results suggest that rs1877670 and rs35900184 SNPs are significantly heterogeneous according to this test. Significant p-values are indicated in bold red text.

(XLSX)

S7 Table. Meta-analysis of combined populations. Meta-analyses were conducted under both a fixed- and random-effect model. Based on the Breslow-Day results in [S6 Table](#), the fixed-effect model is not the most appropriate test for these data. The random-effect model results are presented in the column headed "P(R)". Significant p-value is indicated in bold red text.

(XLSX)

Acknowledgments

The authors thank NIMH for its foresight in initiating its Molecular Genetics Initiative in 1989 to support the collection of the biomaterials used in this study. They are deeply grateful for the participation of all subjects contributing to this research. They also thank research team at the study sites of the Molecular Genetics of Schizophrenia Consortium. This manuscript is dedicated to the memory of Richard Todd, Ph.D., M.D., without whose remarkable encouragement and support this work would not have been possible

Author Contributions

Conceived and designed the experiments: ALG MJH. Performed the experiments: LM JJC JJP RR CNB AA VD RZ JM. Analyzed the data: ALG MJH LM JJC VD RZ JM AT MD. Contributed reagents/materials/analysis tools: ALG CRC RZ JM. Wrote the paper: ALG MJH LM JJC VD JM. Obtained permission for use of DNA Samples: CRC. Intellectual contribution to study design and provided editorial feedback: CRC.

References

1. Rao DC, Morton NE, Gottesman II, Lew R. Path analysis of qualitative data on pairs of relatives: application to schizophrenia. *Human heredity*. 1981; 31(6):325–33. PMID: [7333621](#).

2. Corcoran C, Gallitano A, Leitman D, Malaspina D. The neurobiology of the stress cascade and its potential relevance for schizophrenia. *Journal of psychiatric practice*. 2001; 7(1):3–14. Epub 2005/07/02. PMID: [15990497](#).
3. Sullivan PF, Kendler KS, Neale MC. Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. *Archives of general psychiatry*. 2003; 60(12):1187–92. doi: [10.1001/archpsyc.60.12.1187](#) PMID: [14662550](#).
4. Gallitano-Mendel A, Izumi Y, Tokuda K, Zorumski CF, Howell MP, Muglia LJ, et al. The immediate early gene early growth response gene 3 mediates adaptation to stress and novelty. *Neuroscience*. 2007; 148(3):633–43. Epub 2007/08/19. doi: [10.1016/j.neuroscience.2007.05.050](#) PMID: [17692471](#); PubMed Central PMCID: PMC2597331.
5. Fahmy RG, Dass CR, Sun LQ, Chesterman CN, Khachigian LM. Transcription factor Egr-1 supports FGF-dependent angiogenesis during neovascularization and tumor growth. *Nature medicine*. 2003; 9(8):1026–32. doi: [10.1038/nm905](#) PMID: [12872165](#).
6. Fahmy RG, Khachigian LM. Suppression of growth factor expression and human vascular smooth muscle cell growth by small interfering RNA targeting EGR-1. *Journal of cellular biochemistry*. 2007; 100(6):1526–35. doi: [10.1002/jcb.21145](#) PMID: [17171647](#).
7. Milbrandt J. A nerve growth factor-induced gene encodes a possible transcriptional regulatory factor. *Science*. 1987; 238(4828):797–9. PMID: [3672127](#).
8. Mechtcheriakova D, Wlachos A, Holzmüller H, Binder BR, Hofer E. Vascular endothelial cell growth factor-induced tissue factor expression in endothelial cells is mediated by EGR-1. *Blood*. 1999; 93(11):3811–23. PMID: [10339488](#).
9. Jessen KR, Mirsky R. Signals that determine Schwann cell identity. *Journal of anatomy*. 2002; 200(4):367–76. PMID: [12090403](#); PubMed Central PMCID: PMC1570691.
10. Jones MW, Errington ML, French PJ, Fine A, Bliss TV, Garel S, et al. A requirement for the immediate early gene Zif268 in the expression of late LTP and long-term memories. *Nature neuroscience*. 2001; 4(3):289–96. doi: [10.1038/85138](#) PMID: [11224546](#).
11. Yamada K, Gerber DJ, Iwayama Y, Ohnishi T, Ohba H, Toyota T, et al. Genetic analysis of the calcineurin pathway identifies members of the EGR gene family, specifically EGR3, as potential susceptibility candidates in schizophrenia. *Proceedings of the National Academy of Sciences of the United States of America*. 2007; 104(8):2815–20. Epub 2007/03/16. doi: [10.1073/pnas.0610765104](#) PMID: [17360599](#); PubMed Central PMCID: PMC1815264.
12. Kim SH, Song JY, Joo EJ, Lee KY, Ahn YM, Kim YS. EGR3 as a potential susceptibility gene for schizophrenia in Korea. *American journal of medical genetics Part B, Neuropsychiatric genetics: the official publication of the International Society of Psychiatric Genetics*. 2010; 153B(7):1355–60. Epub 2010/08/06. doi: [10.1002/ajmg.b.31115](#) PMID: [20687139](#).
13. Zhang R, Lu S, Meng L, Min Z, Tian J, Valenzuela RK, et al. Genetic evidence for the association between the early growth response 3 (EGR3) gene and schizophrenia. *PloS one*. 2012; 7(1):e30237. Epub 2012/01/26. doi: [10.1371/journal.pone.0030237](#) PMID: [22276163](#); PubMed Central PMCID: PMC3262808.
14. Mexal S, Frank M, Berger R, Adams CE, Ross RG, Freedman R, et al. Differential modulation of gene expression in the NMDA postsynaptic density of schizophrenic and control smokers. *Brain research Molecular brain research*. 2005; 139(2):317–32. doi: [10.1016/j.molbrainres.2005.06.006](#) PMID: [16122832](#).
15. Gallitano-Mendel A, Wozniak DF, Pehek EA, Milbrandt J. Mice lacking the immediate early gene Egr3 respond to the anti-aggressive effects of clozapine yet are relatively resistant to its sedating effects. *Neuropsychopharmacology: official publication of the American College of Neuropsychopharmacology*. 2008; 33(6):1266–75. Epub 2007/07/20. doi: [10.1038/sj.npp.1301505](#) PMID: [17637609](#).
16. Hippenmeyer S, Schneider NA, Birchmeier C, Burden SJ, Jessell TM, Arber S. A role for neuregulin1 signaling in muscle spindle differentiation. *Neuron*. 2002; 36(6):1035–49. PMID: [12495620](#).
17. Jacobson C, Duggan D, Fischbach G. Neuregulin induces the expression of transcription factors and myosin heavy chains typical of muscle spindles in cultured human muscle. *Proceedings of the National Academy of Sciences of the United States of America*. 2004; 101(33):12218–23. doi: [10.1073/pnas.0404240101](#) PMID: [15302938](#); PubMed Central PMCID: PMC514402.
18. Stefansson H, Sigurdsson E, Steinthorsdóttir V, Björnsdóttir S, Sigmundsson T, Ghosh S, et al. Neuregulin 1 and susceptibility to schizophrenia. *American journal of human genetics*. 2002; 71(4):877–92. Epub 2002/07/30. doi: [10.1086/342734](#) PMID: [12145742](#); PubMed Central PMCID: PMC378543.
19. Yamagata K, Kaufmann WE, Lanahan A, Papapavlou M, Barnes CA, Andreasson KI, et al. Egr3/Pilot, a zinc finger transcription factor, is rapidly regulated by activity in brain neurons and colocalizes with Egr1/zif268. *Learning & memory*. 1994; 1(2):140–52. PMID: [10467592](#).

20. Olney JW, Newcomer JW, Farber NB. NMDA receptor hypofunction model of schizophrenia. *Journal of psychiatric research*. 1999; 33(6):523–33. PMID: [10628529](#).
21. Mittelstadt PR, Ashwell JD. Cyclosporin A-sensitive transcription factor Egr-3 regulates Fas ligand expression. *Molecular and cellular biology*. 1998; 18(7):3744–51. PMID: [9632757](#); PubMed Central PMCID: PMC108957.
22. Zeng H, Chattarji S, Barbarosie M, Rondi-Reig L, Philpot BD, Miyakawa T, et al. Forebrain-specific calcineurin knockout selectively impairs bidirectional synaptic plasticity and working/episodic-like memory. *Cell*. 2001; 107(5):617–29. PMID: [11733061](#).
23. Jurado S, Biou V, Malenka RC. A calcineurin/AKAP complex is required for NMDA receptor-dependent long-term depression. *Nature neuroscience*. 2010; 13(9):1053–5. doi: [10.1038/nn.2613](#) PMID: [20694001](#); PubMed Central PMCID: PMC2943866.
24. Li L, Carter J, Gao X, Whitehead J, Tourtellotte WG. The neuroplasticity-associated arc gene is a direct transcriptional target of early growth response (Egr) transcription factors. *Molecular and cellular biology*. 2005; 25(23):10286–300. doi: [10.1128/MCB.25.23.10286-10300.2005](#) PMID: [16287845](#); PubMed Central PMCID: PMC1291244.
25. Plath N, Ohana O, Dammermann B, Errington ML, Schmitz D, Gross C, et al. Arc/Arg3.1 is essential for the consolidation of synaptic plasticity and memories. *Neuron*. 2006; 52(3):437–44. doi: [10.1016/j.neuron.2006.08.024](#) PMID: [17088210](#).
26. Rial Verde EM, Lee-Osbourne J, Worley PF, Malinow R, Cline HT. Increased expression of the immediate-early gene arc/arg3.1 reduces AMPA receptor-mediated synaptic transmission. *Neuron*. 2006; 52(3):461–74. doi: [10.1016/j.neuron.2006.09.031](#) PMID: [17088212](#); PubMed Central PMCID: PMC3951199.
27. Tzingounis AV, Nicoll RA. Arc/Arg3.1: linking gene expression to synaptic plasticity and memory. *Neuron*. 2006; 52(3):403–7. doi: [10.1016/j.neuron.2006.10.016](#) PMID: [17088207](#).
28. Kirov G, Pocklington AJ, Holmans P, Ivanov D, Ikeda M, Ruderfer D, et al. De novo CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia. *Molecular psychiatry*. 2012; 17(2):142–53. doi: [10.1038/mp.2011.154](#) PMID: [22083728](#); PubMed Central PMCID: PMC3603134.
29. Purcell SM, Moran JL, Fromer M, Ruderfer D, Solovieff N, Roussos P, et al. A polygenic burden of rare disruptive mutations in schizophrenia. *Nature*. 2014; 506(7487):185–90. doi: [10.1038/nature12975](#) PMID: [24463508](#); PubMed Central PMCID: PMC4136494.
30. Fromer M, Pocklington AJ, Kavanagh DH, Williams HJ, Dwyer S, Gormley P, et al. De novo mutations in schizophrenia implicate synaptic networks. *Nature*. 2014; 506(7487):179–84. doi: [10.1038/nature12929](#) PMID: [24463507](#).
31. Genomes Project C, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, et al. An integrated map of genetic variation from 1,092 human genomes. *Nature*. 2012; 491(7422):56–65. doi: [10.1038/nature11632](#) PMID: [23128226](#); PubMed Central PMCID: PMC3498066.
32. Schrauwen I, Sommen M, Comeveaux JJ, Reiman RA, Hackett NJ, Claes C, et al. A sensitive and specific diagnostic test for hearing loss using a microdroplet PCR-based approach and next generation sequencing. *American journal of medical genetics Part A*. 2013; 161A(1):145–52. doi: [10.1002/ajmg.a.35737](#) PMID: [23208854](#).
33. Craig DW, Pearson JV, Szeling S, Sekar A, Redman M, Comeveaux JJ, et al. Identification of genetic variants using bar-coded multiplexed sequencing. *Nat Methods*. 2008; 5(10):887–93. PMID: [18794863](#). doi: [10.1038/nmeth.1251](#)
34. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009; 25(14):1754–60. doi: [10.1093/bioinformatics/btp324](#) PMID: [19451168](#).
35. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009; 25(16):2078–9. doi: [10.1093/bioinformatics/btp352](#) PMID: [19505943](#).
36. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics*. 2007; 81(3):559–75. doi: [10.1086/519795](#) PMID: [17701901](#); PubMed Central PMCID: PMC1950838.
37. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005; 21(2):263–5. PMID: [15297300](#).
38. Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, et al. A second generation human haplotype map of over 3.1 million SNPs. *Nature*. 2007; 449(7164):851–61. PMID: [17943122](#).
39. Purcell S, Cherny SS, Sham PC. Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics*. 2003; 19(1):149–50. PMID: [12499305](#).

40. Bansal V. A statistical method for the detection of variants from next-generation resequencing of DNA pools. *Bioinformatics*. 2010; 26(12):i318–24. doi: [10.1093/bioinformatics/btq214](https://doi.org/10.1093/bioinformatics/btq214) PMID: [20529923](https://pubmed.ncbi.nlm.nih.gov/20529923/); PubMed Central PMCID: PMC2881398.
41. Ning QL, Ma XD, Jiao LZ, Niu XR, Li JP, Wang B, et al. [A family-based association study of the EGR3 gene polymorphisms and schizophrenia]. *Yi chuan = Hereditas / Zhongguo yi chuan xue hui bian ji*. 2012; 34(3):307–14. Epub 2012/03/20. PMID: [22425949](https://pubmed.ncbi.nlm.nih.gov/22425949/).
42. Mansour HA, Talkowski ME, Wood J, Chowdari KV, McClain L, Prasad K, et al. Association study of 21 circadian genes with bipolar I disorder, schizoaffective disorder, and schizophrenia. *Bipolar disorders*. 2009; 11(7):701–10. Epub 2009/10/21. doi: [10.1111/j.1399-5618.2009.00756.x](https://doi.org/10.1111/j.1399-5618.2009.00756.x) PMID: [19839995](https://pubmed.ncbi.nlm.nih.gov/19839995/); PubMed Central PMCID: PMC3401899.
43. Gallitano AL, Tillman R, Dinu V, Geller B. Family-based association study of early growth response gene 3 with child bipolar I disorder. *Journal of affective disorders*. 2012; 138(3):387–96. Epub 2012/03/01. doi: [10.1016/j.jad.2012.01.011](https://doi.org/10.1016/j.jad.2012.01.011) PMID: [22370066](https://pubmed.ncbi.nlm.nih.gov/22370066/); PubMed Central PMCID: PMC3349283.
44. Gallitano A. Dysfunction of the immediate early gene transcription factor Egr3 is implicated in schizophrenia. Society for Neuroscience; November 17th, 2008; Washington D.C. 2008 Neuroscience Meeting Planner 2008.
45. Zhang W, Wu J, Ward MD, Yang S, Chuang YA, Xiao M, et al. Structural basis of arc binding to synaptic proteins: implications for cognitive disease. *Neuron*. 2015; 86(2):490–500. doi: [10.1016/j.neuron.2015.03.030](https://doi.org/10.1016/j.neuron.2015.03.030) PMID: [25864631](https://pubmed.ncbi.nlm.nih.gov/25864631/); PubMed Central PMCID: PMC4409568.
46. Cheng MC, Chuang YA, Lu CL, Chen YJ, Luu SU, Li JM, et al. Genetic and functional analyses of early growth response (EGR) family genes in schizophrenia. *Progress in neuro-psychopharmacology & biological psychiatry*. 2012; 39(1):149–55. Epub 2012/06/14. doi: [10.1016/j.pnpbp.2012.06.004](https://doi.org/10.1016/j.pnpbp.2012.06.004) PMID: [22691714](https://pubmed.ncbi.nlm.nih.gov/22691714/).
47. Kyogoku C, Yanagi M, Nishimura K, Sugiyama D, Morinobu A, Fukutake M, et al. Association of calcineurin A gamma subunit (PPP3CC) and early growth response 3 (EGR3) gene polymorphisms with susceptibility to schizophrenia in a Japanese population. *Psychiatry research*. 2011; 185(1–2):16–9. Epub 2010/06/12. doi: [10.1016/j.psychres.2009.11.003](https://doi.org/10.1016/j.psychres.2009.11.003) PMID: [20537399](https://pubmed.ncbi.nlm.nih.gov/20537399/).
48. Liu BC, Zhang J, Wang L, Li XW, Wang Y, Ji J, et al. No association between EGR gene family polymorphisms and schizophrenia in the Chinese population. *Progress in neuro-psychopharmacology & biological psychiatry*. 2010; 34(3):506–9. Epub 2010/02/11. doi: [10.1016/j.pnpbp.2010.02.005](https://doi.org/10.1016/j.pnpbp.2010.02.005) PMID: [20144677](https://pubmed.ncbi.nlm.nih.gov/20144677/).
49. Nishimura Y, Takizawa R, Koike S, Kinoshita A, Satomura Y, Kawasaki S, et al. Association of decreased prefrontal hemodynamic response during a verbal fluency task with EGR3 gene polymorphism in patients with schizophrenia and in healthy individuals. *NeuroImage*. 2014; 85 Pt 1:527–34. Epub 2013/08/22. doi: [10.1016/j.neuroimage.2013.08.021](https://doi.org/10.1016/j.neuroimage.2013.08.021) PMID: [23962955](https://pubmed.ncbi.nlm.nih.gov/23962955/).
50. Maier W. Common risk genes for affective and schizophrenic psychoses. *European archives of psychiatry and clinical neuroscience*. 2008; 258 Suppl 2:37–40. Epub 2008/07/03. doi: [10.1007/s00406-008-2008-z](https://doi.org/10.1007/s00406-008-2008-z) PMID: [18516516](https://pubmed.ncbi.nlm.nih.gov/18516516/).
51. Lichtenstein P, Yip BH, Bjork C, Pawitan Y, Cannon TD, Sullivan PF, et al. Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet*. 2009; 373(9659):234–9. Epub 2009/01/20. doi: [10.1016/S0140-6736\(09\)60072-6](https://doi.org/10.1016/S0140-6736(09)60072-6) PMID: [19150704](https://pubmed.ncbi.nlm.nih.gov/19150704/); PubMed Central PMCID: PMC3879718.
52. Cross-Disorder Group of the Psychiatric Genomics C, Genetic Risk Outcome of Psychosis C. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet*. 2013; 381(9875):1371–9. doi: [10.1016/S0140-6736\(12\)62129-1](https://doi.org/10.1016/S0140-6736(12)62129-1) PMID: [23453885](https://pubmed.ncbi.nlm.nih.gov/23453885/); PubMed Central PMCID: PMC3714010.
53. Guo AY, Sun J, Jia P, Zhao Z. A novel microRNA and transcription factor mediated regulatory network in schizophrenia. *BMC systems biology*. 2010; 4:10. Epub 2010/02/17. doi: [10.1186/1752-0509-4-10](https://doi.org/10.1186/1752-0509-4-10) PMID: [20156358](https://pubmed.ncbi.nlm.nih.gov/20156358/); PubMed Central PMCID: PMC2834616.
54. Tam GW, Redon R, Carter NP, Grant SG. The role of DNA copy number variation in schizophrenia. *Biological psychiatry*. 2009; 66(11):1005–12. doi: [10.1016/j.biopsych.2009.07.027](https://doi.org/10.1016/j.biopsych.2009.07.027) PMID: [19748074](https://pubmed.ncbi.nlm.nih.gov/19748074/).
55. Tabares-Seisdedos R, Rubenstein JL. Chromosome 8p as a potential hub for developmental neuro-psychiatric disorders: implications for schizophrenia, autism and cancer. *Molecular psychiatry*. 2009; 14(6):563–89. Epub 2009/02/11. doi: [10.1038/mp.2009.2](https://doi.org/10.1038/mp.2009.2) PMID: [19204725](https://pubmed.ncbi.nlm.nih.gov/19204725/).
56. Takahashi S, Faraone SV, Lasky-Su J, Tsuang MT. Genome-wide scan of homogeneous subtypes of NIMH genetics initiative schizophrenia families. *Psychiatry research*. 2005; 133(2–3):111–22. doi: [10.1016/j.psychres.2004.12.003](https://doi.org/10.1016/j.psychres.2004.12.003) PMID: [15740987](https://pubmed.ncbi.nlm.nih.gov/15740987/).
57. Kaufmann CA, Suarez B, Malaspina D, Pepple J, Svraic D, Markel PD, et al. NIMH Genetics Initiative Millenium Schizophrenia Consortium: linkage analysis of African-American pedigrees. *American journal of medical genetics*. 1998; 81(4):282–9. PMID: [9674972](https://pubmed.ncbi.nlm.nih.gov/9674972/).

58. Holliday EG, Mowry BJ, Nyholt DR. A reanalysis of 409 European-Ancestry and African American schizophrenia pedigrees reveals significant linkage to 8p23.3 with evidence of locus heterogeneity. *American journal of medical genetics Part B, Neuropsychiatric genetics: the official publication of the International Society of Psychiatric Genetics*. 2008; 147B(7):1080–8. doi: [10.1002/ajmg.b.30722](https://doi.org/10.1002/ajmg.b.30722) PMID: [18361422](https://pubmed.ncbi.nlm.nih.gov/18361422/).
59. Ng MY, Levinson DF, Faraone SV, Suarez BK, DeLisi LE, Arinami T, et al. Meta-analysis of 32 genome-wide linkage studies of schizophrenia. *Molecular psychiatry*. 2009; 14(8):774–85. Epub 2009/04/08. doi: [10.1038/mp.2008.135](https://doi.org/10.1038/mp.2008.135) PMID: [19349958](https://pubmed.ncbi.nlm.nih.gov/19349958/); PubMed Central PMCID: PMC2715392.
60. Lewis CM, Levinson DF, Wise LH, DeLisi LE, Straub RE, Hovatta I, et al. Genome scan meta-analysis of schizophrenia and bipolar disorder, part II: Schizophrenia. *American journal of human genetics*. 2003; 73(1):34–48. Epub 2003/06/13. doi: [10.1086/376549](https://doi.org/10.1086/376549) PMID: [12802786](https://pubmed.ncbi.nlm.nih.gov/12802786/); PubMed Central PMCID: PMC1180588.
61. Bresnahan M, Begg MD, Brown A, Schaefer C, Sohler N, Insel B, et al. Race and risk of schizophrenia in a US birth cohort: another example of health disparity? *International journal of epidemiology*. 2007; 36(4):751–8. doi: [10.1093/ije/dym041](https://doi.org/10.1093/ije/dym041) PMID: [17440031](https://pubmed.ncbi.nlm.nih.gov/17440031/).