Schizophrenia Research xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Schizophrenia Research

journal homepage: www.elsevier.com/locate/schres

Evidence for schizophrenia susceptibility alleles in the Indian population: An association of neurodevelopmental genes in case–control and familial samples

Ajay Jajodia [a], Harpreet Kaur [a], Kalpana Kumari [b], Meenal Gupta [a], Ruchi Baghel [a], Ankit Srivastava [a], Mamta Sood [b], Rakesh Kumar Chadda [b], Sanjeev Jain [c], Ritushree Kukreti [a,*]

[a] Genomics and Molecular Medicine, CSIR-Institute of Genomics and Integrative Biology, Mall Road, Delhi 110007, India
[b] Department of Psychiatry, All India Institute of Medical Sciences, Ansari Nagar, New Delhi 110029, India
[c] Molecular Genetic Laboratory, Department of Psychiatry, National Institute of Mental Health and Neuro Sciences, Hosur Road, Bengaluru 560029, India

ARTICLE INFO

Article history:
Received 20 August 2014
Received in revised form 26 November 2014
Accepted 21 December 2014
Available online xxxx

Keywords:
PGC
Schizophrenia
rs1062613
rs548181
Meta-analysis
TDT

ABSTRACT

Schizophrenia is a severe psychiatric disorder with lifetime prevalence of ~1% worldwide. A genotyping study was conducted using a custom panel of Illumina 1536 SNPs in 840 schizophrenia cases and 876 controls (351 patients and 385 controls from North India; and 436 patients, 401 controls and 143 familial samples with 53 probands containing 37 complete and 16 incomplete trios from South India). Meta-analysis of this population of Indo-European and Dravidian ancestry identified three strongly associated variants with schizophrenia: *STT3A* (rs548181, $p = 1.47 \times 10^{-5}$), *NRG1* (rs17603876, $p = 8.66 \times 10^{-5}$) and *GRM7* (rs3864075, $p = 4.06 \times 10^{-3}$). Finally, a meta-analysis was conducted comparing our data with data from the Schizophrenia Psychiatric Genome-Wide Association Study Consortium (PGC-SCZ) that supported rs548181 ($p = 1.39 \times 10^{-7}$). In addition, combined analysis of sporadic case–control association and a transmission disequilibrium test in familial samples from South Indian population identified three associations: rs1062613 ($p = 3.12 \times 10^{-3}$), a functional promoter variant of *HTR3A*; rs6710782 ($p = 3.50 \times 10^{-3}$), an intronic variant of *ERBB4*; and rs891903 ($p = 1.05 \times 10^{-2}$), an intronic variant of *EBF1*. The results support the risk variants observed in the earlier published work and suggest a potential role of neurodevelopmental genes in the schizophrenia pathogenesis.

1. Introduction

Schizophrenia is a debilitating neuropsychiatric disorder that affects 1% of the world population. In spite of low fecundity among inflicted patients that acts as strong selective disadvantage for schizophrenia, the disorder continues to show a high prevalence. Clinical characteristics include positive symptoms (hallucinations, delusions), negative symptoms (avolition, social withdrawal, reduced affective expression) and altered cognitive functioning (Karam et al., 2010).

Genetic and gene–environment interactions account for over 80% susceptibility in the development of schizophrenia (Tandon et al., 2008). De-novo mutations contribute toward the constant replenishment of the pathogenic alleles involved in the disease development (Hatzimanolis et al., 2013). Mutations in genes and pathways controlling the neurodevelopment process includes NRG1-ERBB signaling (Chen et al., 2006; Mei and Xiong, 2008; Shamir et al., 2012), glutamate signaling (Goff and Coyle, 2001; Moghaddam, 2003), active-ligand receptor (Putnam et al., 2011; Adkins et al., 2012; Ayalew et al., 2012; Liu et al., 2013) and their downstream cascade (Harrison and Owen, 2003; Emamian et al., 2004; Karam et al., 2010) which can significantly affect their co-ordinated action (Walsh et al., 2008). Both common and rare variants may play a role in the development of schizophrenia. In the recent past, a number of genome wide association studies (GWASs) have been performed in schizophrenia and other psychiatric disorders (Bergen and Petryshen, 2012). Some significant signals have been identified and validated in the large replication samples.

Based on these previous observations, we hypothesized the involvement of neurodevelopment genes in the deficits associated with schizophrenia. We report the results of candidate gene association study in schizophrenia, conducted in two independent populations (North India and South India) with a total of 840 schizophrenia cases and 876 controls. Samples from these two independent populations were evaluated for any association with schizophrenia and the resultant data were compared to the Schizophrenia Psychiatric GWAS Consortium (PGC-SCZ) data (Schizophrenia Psychiatric Genome-Wide Association Study Consortium, 2011), which has shown a same direction of effect. They reported a mega-analysis of GWAS data sets in Caucasian population comprising 21,856 subjects (9394 cases and 12,462 controls) in stage I and

* Corresponding author. Tel.: +91 11 27662202, +91 11 27666156; fax: +91 11 27667471.
E-mail address: ritus@igib.res.in (R. Kukreti).

http://dx.doi.org/10.1016/j.schres.2014.12.031
0920-9964/

follow-up of top association ($p < 2 \times 10^{-5}$) in stage II independent subjects (8442 cases and 21,397 controls). Additionally, we investigated familial association; followed by its combined analysis with case–control samples from the South Indian population.

2. Materials and methods

2.1. Sample

A total of 736 individuals of North Indian ethnicity (351 cases and 385 controls) and 980 individuals of South Indian origin (436 cases, 401 controls, and 143 familial samples with 53 probands containing 37 complete and 16 incomplete trios) were recruited from Clinical Services of Outpatients Department of Psychiatry Services, All India Institute of Medical Sciences, New Delhi and National Institute of Medical Health and Allied Sciences, Bengaluru, respectively (Table 1). Patients were excluded if they had i) past or family history of any other mental illness or neurological disorder; ii) substance dependence; iii) history of head injury; or iv) pregnancy at the time of enrollment. North Indian samples belong to Indo-European ancestry, whereas South Indian samples were of Dravidian origin. Voluntary informed consent was obtained after a complete description of the study to the participants. This study was reviewed and approved by the review boards of the hospitals and the Institute Ethics Committee.

2.2. Genotyping

The main criterion for selection of candidate gene/pathway were based on their involvement in disease and therapeutic NRG1-ERBB signaling pathway, neuroactive ligand receptor interaction, glutamate signaling, downstream signaling, and for follow-up from previous GWAS conducted in patients with schizophrenia (Need et al., 2009; Schizophrenia Psychiatric Genome-Wide Association Study Consortium, 2011; Shi et al., 2011). We utilized a custom set of 1536 SNPs, of which 984 were prioritized from 40 candidate genes based on gene coverage, functional prediction, and prior reports. Also, 552 neutral markers (neither localize to linkage region and implicated with any brain disorder) were selected for a homogeneity test (Jha et al., 2012).

Genotyping was performed at Genotyping and Sequencing Facility of CSIR-IGIB using Illumina Genomestation using GoldenGate technology according to manufacturer's guidelines. 88 samples of DNA were served as duplicates for quality controlling terms of clustering and reproducibility. Primary genotyping data analyses were performed with GenomeStudio software and followed by visual inspection and assessment of data quality and clustering. To rule out any genotyping error, 15% of the total samples were re-genotyped randomly using primer extension reaction in the MALDI-TOF mass spectrometry platform (Sequenom™, Inc., San Diego, CA).

2.3. Statistical analysis

GenomeStudio Genotyping Module v1.0 was used to analyze SNPs of the Illumina custom panel for DNA sample validation. All data management and quality control assessments were performed with PLINK 1.07 (Purcell et al., 2007). There is a "north–south" gradient in the Indian population (IGVC, 2005). Therefore, analysis was performed separately for both data sets. Results from the two study data sets and the PGC-SCZ study were compared by meta-analysis using PLINK. Heterogeneity was evaluated with the Cochran Q statistic test in the contributing data sets. Odds ratios (ORs) and 95% confidence intervals (CIs) were calculated using R script. A transmission disequilibrium test (TDT) analysis was performed in South Indian trios using PLINK and overall estimates of allelic effect were calculated by combined analysis recommended in published literature (Kazeem and Farrall, 2005).

Linkage Disequilibrium (LD) was estimated among SNPs by r^2 value for the genotype data of all markers in healthy controls using Tagger algorithm in Haploview program version 4.1 (Barrett et al., 2005). Identification of tag SNPs was performed using Tagger program (de Bakker et al., 2005). Population structure was investigated by Principal Component (PC) analysis performed for a set of neutral markers with Eigenstrat (Price et al., 2006). Additionally, a stratification test was performed by STRUCTURE 2.1 (Pritchard et al., 2000). We assumed that $K = 2$ and model parameters were burn-in period, number of Markov Chain Monte Carlo (MCMC) repeats after burn-in, and iterations as 500, 5000 and 1000, respectively for Structured Population Association Test (STRAT) program.

Table 1
Descriptive statistics of samples in case–control and trio studies.

	South India		North India	
	Cases	Controls	Cases	Controls
Case–control study				
n	436	401	351	385
% female	39.9	35.32	47	49.6
Age (years)	29.77 ± 7.5	27.41 ± 7.21	32.94 ± 9.27	34.38 ± 11.81
Trio study	53			
No. of pedigree	53	90		
% female	39.62	46.66		
Age (years)	25.36 ± 4.51	53.40 ± 7.31		

3. Results

3.1. Quality control

During the initial quality check (QC), loci were screened and excluded from analysis for the following reasons: poorly defined clusters, GenTrain score < 0.6, more than 60% genotyping failure, significant differences in missingness between cases and controls, excessive Mendelian and replication errors and samples having all the heterozygous calls (Anderson et al., 2010). Among 1536 SNPs, few were discarded due to low GenTrain score (42 SNPs and 43 SNPs), failure in missingness between cases and controls (83 SNPs and 19 SNPs) and low genotyping (7 SNPs and 4 SNPs) in North and South Indian populations, respectively. Genotype concordance was ~98.8% for Illumina and ~94.6% for Sequenom platform.

Hardy–Weinberg Equilibrium (HWE) was assessed for all the SNPs that passed initial QC. SNPs deviating from HWE were checked by a goodness of fit test (df = 1) with a threshold of significance of $p < 0.0001$ in controls. SNPs with minor allele frequency (MAF) < 1% and non-polymorphic nature were not included in further analysis. In North Indian samples, 58 markers failed to withstand HWE and 61 markers had MAF < 1%. In South Indian samples, 53 SNPs did not withstand HWE and 91 markers failed to withstand 1% MAF criteria. Finally, 1287 markers in 336 cases and 379 unrelated controls that passed QC were submitted to further analysis in North Indian samples and; a total of 1329 SNPs in 435 cases and 393 healthy controls and 141 individuals of family samples withstanding QC criteria were made available for South Indian data set. Overall, the total genotyping rate was above 98%. 153 SNPs in North Indian samples and 141 SNPs in South Indian samples were pruned out based on high pair-wise linkage disequilibrium ($r^2 \geq 0.9$) in control samples.

3.2. Population stratification

A major caveat in population based associations is stratification in the sample pool which can serve as a potential confounder leading to false positive results. To rule out the possibility of population stratification, patients and unrelated healthy individuals were recruited from the same ethnicity. Additionally, 552 neutral markers were genotyped in the studied populations. After QC, 436 SNPs in North Indian samples and 486 SNPs in South Indian samples were eligible for the homogeneity test. The stratification test revealed homogeneity between cases and controls

Please cite this article as: Jajodia, A., et al., Evidence for schizophrenia susceptibility alleles in the Indian population: An association of neurodevelopmental genes in case–c..., Schizophr. Res. (2015), http://dx.doi.org/10.1016/j.schres.2014.12.031