

ORIGINAL ARTICLE

Common polygenic variation and risk for childhood-onset schizophrenia

K Ahn¹, SS An², YY Shugart³ and JL Rapoport¹

Childhood-onset schizophrenia (COS) is a rare and severe form of the disorder, with more striking abnormalities with respect to prepsychotic developmental disorders and abnormalities in the brain development compared with later-onset schizophrenia. We previously documented that COS patients, compared with their healthy siblings and with adult-onset patients (AOS), carry significantly more rare chromosomal copy number variations, spanning large genomic regions (> 100 kb) (Ahn *et al.* 2014). Here, we interrogated the contribution of common polygenic variation to the genetic susceptibility for schizophrenia. We examined the association between a direct measure of genetic risk of schizophrenia in 130 COS probands and 103 healthy siblings. Using data from the schizophrenia and autism GWAS of the Psychiatric Genomic Consortia, we selected three risk-related sets of single nucleotide polymorphisms from which we conducted polygenic risk score comparisons for COS probands and their healthy siblings. COS probands had higher genetic risk scores of both schizophrenia and autism than their siblings ($P < 0.05$). Given the small sample size, these findings suggest that COS patients have more salient genetic risk than do AOS.

Molecular Psychiatry (2016) **21**, 94–96; doi:10.1038/mp.2014.158; published online 16 December 2014

INTRODUCTION

With recent advances in genomic microarrays, identifying common genetic variants for complex traits have become a reality through genome-wide association studies (GWAS).¹ For psychiatric disorders, identification of causal gene signatures has remained unclear; however, because each genetic variant or single nucleotide polymorphism (SNP) accounts for such a small proportion of the variance, necessitating very large samples.^{2–4} The Psychiatric Genomic Consortium (PGC) was created in order to conduct large-scale mega-analyses of GWAS data for five major psychiatric disorders, including ADHD, autism, bipolar, major depressive disorder and schizophrenia.^{5–9} Meta-analyses for these psychiatric disorders, on the one hand, confirmed many previous GWAS findings and, on the other hand, added previously unrecognized risk common variants to the list.^{5–9} However, the statistical stringency of multiple test corrections necessary for GWAS has limited studies of smaller samples.

Childhood-onset schizophrenia (COS), defined by the onset of psychotic symptoms before age 13, is a rare and severe form of the disorder. Stratifying by age of onset has been useful across all of medicine and, in particular, for identifying causal genetic variants.^{10,11} We previously showed that COS patients carry a higher rate of large (> 100 kb) and rare (< 0.1% in controls) copy number variations that interrupt genes in pathways of neurodevelopment and regulation than do their healthy siblings or adult-onset patients (AOS).^{12,13} In addition, we showed an unexpected relationship. In order to further understand the genetic susceptibility for COS, here we examined the contribution of common polygenic variation in our COS samples. In addition, we evaluated polygenic risk scores and their association with published GWAS of schizophrenia in the PGC. Finally, because COS shows a high

rate of prepsychotic neurodevelopmental dysfunction,¹⁴ we also examined polygenic risk for autism in our COS sample.

MATERIALS AND METHODS

Patient population

Participants were recruited as part of an ongoing National Institute of Mental Health study of COS. Patients meeting DSM-III-R/DSM-IV criteria for schizophrenia with documented onset of psychosis before age 13 were recruited nationally. Patients and their available first-degree relatives were interviewed for lifetime and current psychiatric disorders using structured psychiatric interviews and Autism Symptom Questionnaire.^{15,16} Diagnosis was confirmed with inpatient medication-free observation. A total of 600 patients were screened from which 361 were admitted for further observation. This study was approved by the Institutional Review Board of The National Institute of Mental Health. All participants provided written assent/consent with written informed consent from a parent or legal guardian for minors. A more detailed description of the inclusion protocol has been described elsewhere.¹³

SNP genotyping and quality control

Details of the sample collection, processing and quality controls have been described elsewhere.¹³ After quality control process, the remained SNP set was then used to calculate principle components through EIGENSOFT¹ (www.hsph.harvard.edu/alkes-price/software) to assess for population stratification and adjust for this in the later analysis. The data were then imputed to HapMap version 3 using the MACH software (Center for Statistical Genetics, University of Michigan; <http://www.sph.umich.edu/csg/abecasis/MACH>)¹⁷ with the default setting and a quality control ($R > 0.8$).

Selection of SNPs and risk score calculation

Two sets of SNPs were generated from published GWAS of schizophrenia (PGC-SCZ) and autism spectrum disorder (PGC-ASD) in the Psychiatric Genetic Consortium.⁵ The first set consisted of top 80 significant SNPs

¹Childhood Psychiatry Branch, National Institute of Mental Health, National Institutes of Health, Bethesda, MD, USA; ²Department of Environmental Health Sciences, Johns Hopkins University, Bloomberg School of Public Health, Baltimore, MD, USA and ³Unit of Statistical Genomics, National Institute of Mental Health, National Institutes of Health, Bethesda, MD, USA. Correspondence: Dr K Ahn, Child Psychiatry Branch, National Institute of Mental Health, National Institutes of Health, 10 Center Drive, Room 3N202/MSC 1600, Bethesda, MD 20892, USA.

E-mail: kwangmi.ahn@nih.gov

Received 18 August 2014; revised 2 October 2014; accepted 9 October 2014; published online 16 December 2014

($P < 4 \times 10^{-5}$) from PGC-SCZ for the association test. Another three lists of SNPs were derived from risk markers at significance thresholds of $P < 0.4$, $P < 0.2$ and $P < 0.1$. Polygenic scores were calculated using Plink V. 1.07 (<http://pngu.mgh.harvard.edu/~purcell/plink>)¹⁸ using the method described by Purcell *et al.*¹⁹ In addition, we scored our samples using the most recent 108 schizophrenia risk loci from the PGC.²⁰ Briefly, estimates of the log of the odds ratios of association tests were obtained from the PGC-SCZ or PGC-ASD, and for each SNP, the log of odds ratio of the an allele was multiplied by (0, 1 or 2) depending on the number of reference alleles than an individual carries. The total polygenic score is a sum of across SNPs. Analyses were performed separately for score alleles derived from PGC-SCZ and PGC-ASD data sets.

Statistical analyses

Family-based association test was used to examine the association between selected SNPs and COS using PBAT software (<http://www.hsph.harvard.edu/fbat/default.html>).²¹

For the comparisons of siblings and probands' polygenic scores, we also used linear mixed models in SAS version 9.3 (SAS Institute, Cary, NC, USA), adjusting for family membership and for the first five principal components in order to take into account possible population stratification. Nagelkerke's pseudo R^2 was used to assess the variance explained. We investigated if a polygenic score was associated with COS status. A one-tailed test was applied under a directional hypothesis that expects higher scores to be associated with an increased risk of disease.

RESULTS

A total of 130 COS probands, 210 of their parents and 103 of their full healthy siblings passed all quality control. Demographic characteristics of COS probands are shown in Table 1. Due to the diverse ethnicity background, we performed analyses of all possible subjects, and only white Caucasians, separately.

Selected SNP analysis

Only COS cases, their biological parents, and full siblings (443 individuals) were included in this analysis. This small sample size had under 5% of power to detect any association (effect size = 2) in 1 M genome-wide scan, calculated in PBAT. Due to this lack of statistical power to conduct genome-wide association tests, we selected the top 80 significant SNPs nominated by PGC-SCZ in order to test any association with COS. This design allowed at most 50% (effect size = 2) of statistical power. Table 2a showed six significant SNPs ($P < 0.05$) before the Bonferroni correction for the multiple tests using all of possible subjects. Two markers, rs9662700, on chromosome 1 and rs17512836, on chromosome 18 survived after correction for multiple comparisons ($P < 6.25 \times 10^{-4}$). As shown in Table 2b, from an analysis using only Caucasian (208 individuals), there were eight markers before the correction, but only one marker, rs17512836, remained significant after correction. The SNP rs17595731 is intronic for

transcription factor 4 (*TCF4*) gene. Interestingly, *TCF4* has also been reported as a risk gene for bipolar disorder²² and schizophrenia,^{5,8,23} and common *TCF4* variants are involved in psychosis pathology, probably related to abnormal neurodevelopment.²⁴

Polygenic score analysis

For this analysis, 130 COS probands and 103 their healthy siblings were included. Seventy-one probands have at least one healthy full sibling, and the mixed effect model was used in order to control the random effect due to the family membership. In the analyses based on the PGC-SCZ data, polygenic scores for schizophrenia in COS probands were higher than those in sibling group ($P < 0.001$). The polygenic scores based on 108 risk loci in Schizophrenia Working Group of the Psychiatric Genomics²⁰ also showed that COS patients had significantly higher score than their healthy siblings ($P = 0.025$, $R^2 = 0.0552$).

In the analyses based on the PGC-ASD, however, the scores of COS probands were also significantly higher than healthy sibling group, but at only one GWAS significance threshold ($P_T < 0.4$) (see Table 3).

We also examined the relationship between autism spectrum questionnaire score and polygenetic scores driven by Autism GWAS of PGC as 18.5% of our COS probands ($n = 24$) had autism spectrum disorders and many of rare copy number variation abnormalities overlap as risk for autism and schizophrenia.¹³ We found that the polygenetic score driven by autism risk alleles had no correlation with autism spectrum questionnaire score ($r = -0.06881$, $P = 0.617$). On the basis of pseudo R^2 , the polygenic risk score accounted for 7.5% when $P_T = 0.4$; analyses using only Caucasians showed the same patterns of results.

Table 2a. Significant SNPs of family-based association test in the childhood-onset schizophrenia samples using 80 selected SNPs based on schizophrenia GWAS of PGC (all subjects)

Chr	SNP ID	Reference allele	P-value	Nearby genes (within 50 kb)
1	rs11802925	C	0.002006	EF570048
1	rs9662700*	A	0.0005254	ZSCAN20
4	rs4295265	T	0.04676	BANK1
8	rs10503253	A	0.04806	CSMD1
12	rs2108636	T	0.001599	CACNA1C
18	rs17512836*	C	0.0004442	TCF4

Abbreviations: GWAS, genome-wide association studies; PGC, Psychiatric Genomic Consortium, SNPs, single nucleotide polymorphisms. * $P < 6.25 \times 10^{-4}$.

Table 2b. Significant SNPs of family-based association test in the childhood-onset schizophrenia samples using 80 selected SNPs based on schizophrenia GWAS of PGC (Caucasians only)

Chr	SNP ID	Reference allele	P-value	Nearby genes (within 50 kb)
1	rs9662700	A	0.01543	ZSCAN20
4	rs4295265	T	0.02862	BANK1
7	rs7807744	T	0.03934	INHBA
8	rs983309	A	0.008915	AKO55863
8	rs2162119	G	0.04882	
10	rs11191732	A	0.02096	EF570048
11	rs1044796	T	0.005251	CREB3L1, OASIS, DGKZ
18	rs17512836 ^a	C	0.0005349	TCF4

Abbreviations: GWAS, genome-wide association studies; PGC, Psychiatric Genomic Consortium, SNPs, single nucleotide polymorphisms. ^aSignificant after the multiple comparison corrections.

Table 1. Demographic characteristics of childhood-onset schizophrenia probands

	COS probands (n = 130)	Healthy siblings (n = 103)
Male, n	71	52
Caucasian, n	69	51
Age of onset, median (range)	10 (4–12)	—
<i>Prepsychotic developmental pattern</i>		
PDD, n	85	—
ASD, n	17	—
Comorbid ASD, n	24	—

Abbreviations: ASD, autism spectrum disorder; PDD, pervasive developmental disorder.

Table 3. Thresholds, number of SNPs for polygenic score and summary of results in the comparison of COS probands and healthy siblings

Discovery data	Threshold	Selected SNPs, n	P	R ²
PGC-SCZ	$P_T=0.1$	13 717	< 0.0001	0.0668
PGC-SCZ	$P_T=0.2$	22 968	< 0.0001	0.1030
PGC-SCZ	$P_T=0.4$	37 402	< 0.00001	0.1852
PGC-ASD	$P_T=0.1$	8688	0.3062	3.6E-05
PGC-ASD	$P_T=0.2$	17 404	0.1294	0.0002
PGC-ASD	$P_T=0.4$	34 662	0.0150	0.0648

Abbreviations: COS, childhood-onset schizophrenia; PGC-ASD, autism GWAS in Psychiatric Genetics Consortium; PGC-SCZ, schizophrenia GWAS in Psychiatric Genetics Consortium; SNPs, single nucleotide polymorphisms.

DISCUSSION

In this study, we found that the polygenic risk score for schizophrenia, created by using 80 selected genetic common variants from the schizophrenia GWAS in PGC, effectively predicted COS status. Vorstman *et al.*⁶ failed to show that the polygenic score derived from adult-onset schizophrenia case-control data set could differentiate autism cases ($n=2736$) from controls. This study is also the first demonstration of significance overlap in polygenic susceptibility to autism and early-onset schizophrenia, although the relationship was detectable only using the most liberal significance threshold, $P_T < 0.4$. On the other hand, this observation is consistent with our previous reports that autism and COS share several rare copy number variations as risk factors.¹³ These findings suggest that COS may share additional common risk markers with autism.

We replicated the association between rs17512836 (intron 3 in *TCF4*), located on chromosome 18q21, and COS risk. *TCF4* is highly expressed in the brain, and has a role in neurodevelopment, interacting with class II bHLH transcription factors Math1, HASH1 and neuroD2. The Ca(2+) sensor protein calmodulin interacts with the DNA binding domain of *TCF4*, inhibiting transcriptional activation.^{25,26} Because of its interaction with calmodulin and involvement in calcium signaling, *TCF4* may also have direct functional neuronal effects.²⁶ Although family-based association can be controlled for population stratification, COS samples were ethnically heterogeneous. Moreover, our sample, is necessarily small, given the rarity of this subgroup of patients. A definitive statement regarding risk loci and effect size cannot be made conclusively until replication studies are performed.

The estimated variances using R^2 based on the schizophrenia risk variants in this sample (range 5.5 ~ 18.5%) are much stronger than that for the polygenic score representing later-onset schizophrenia (~6%)⁵ or bipolar disorder (~3%).²⁷ These findings suggest that COS patients have a more salient genetic risk with respect to common variants than do adult-onset patients.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006; **38**: 904–909.
- Moonesinghe R, Khoury MJ, Liu T, Ioannidis JP. Required sample size and non-replicability thresholds for heterogeneous genetic associations. *Proc Natl Acad Sci USA* 2008; **105**: 617–622.
- Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. *Am J Hum Genet* 2012; **90**: 7–24.
- Chapman K, Ferreira T, Morris A, Asimit J, Zeggini E. Defining the power limits of genome-wide association scan meta-analyses. *Genet Epidemiol* 2011; **35**: 781–789.
- Schizophrenia Psychiatric Genome-Wide Association Study Consortium. Genome-wide association study identifies five new schizophrenia loci. *Nat Genet* 2011; **43**: 969–976.
- Vorstman JA, Anney RJ, Derks EM, Gallagher L, Gill M, de Jonge MV *et al.* No evidence that common genetic risk variation is shared between schizophrenia and autism. *Am J Med Genet B Neuropsychiatr Genet* 2013; **162B**: 55–60.
- Cross-Disorder Group of the Psychiatric Genomics Consortium. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet* 2013; **45**: 984–994.
- Stefansson H, Ophoff RA, Steinberg S, Andreassen OA, Cichon S, Rujescu D *et al.* Common variants conferring risk of schizophrenia. *Nature* 2009; **460**: 744–747.
- Neale BM, Medland SE, Ripke S, Asherson P, Franke B, Lesch KP *et al.* Meta-analysis of genome-wide association studies of attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry* 2010; **49**: 884–897.
- Hall JM, Lee MK, Newman B, Morrow JE, Anderson LA, Huey B *et al.* Linkage of early-onset familial breast cancer to chromosome 17q21. *Science* 1990; **250**: 1684–1689.
- Murrell J, Farlow M, Ghetti B, Benson MD. A mutation in the amyloid precursor protein associated with hereditary Alzheimer's disease. *Science* 1991; **254**: 97–99.
- Walsh T, McClellan JM, McCarthy SE, Addington AM, Pierce SB, Cooper GM *et al.* Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science* 2008; **320**: 539–543.
- Ahn K, Gotay N, Andersen TM, Anvari AA, Gochman P, Lee Y *et al.* High rate of disease-related copy number variations in childhood onset schizophrenia. *Mol Psychiatry* 2014; **19**: 568–572.
- Addington AM, Rapoport JL. The genetics of childhood-onset schizophrenia: when madness strikes the prepubescent. *Curr Psychiatry Rep* 2009; **11**: 156–161.
- Kumra S, Frazier JA, Jacobsen LK, McKenna K, Gordon CT, Lenane MC *et al.* Childhood-onset schizophrenia. A double-blind clozapine-haloperidol comparison. *Arch Gen Psychiatry* 1996; **53**: 1090–1097.
- Shaw P, Sporn A, Gogtay N, Overman GP, Greenstein D, Gochman P *et al.* Childhood-onset schizophrenia: a double-blind, randomized clozapine-olanzapine comparison. *Arch Gen Psychiatry* 2006; **63**: 721–730.
- Li Y, Willer C, Sanna S, Abecasis G. Genotype imputation. *Annu Rev Genomics Hum Genet* 2009; **10**: 387–406.
- 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. *Am J Hum Genet* 2007; **81**: 559–575.
- Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC *et al.* International Schizophrenia Consortium. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009; **460**: 748–752.
- Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 2014; **511**: 421–427.
- Laird NM, Horvath S, Xu X. Implementing a unified approach to family-based tests of association. *Genet Epidemiol* 2000; **19**: 536–542.
- Del-Favero J, Gestel SV, Borglum AD, Muir W, Ewald H, Mors O *et al.* European combined analysis of the CTG18.1 and the ERDA1 CAG/CTG repeats in bipolar disorder. *Eur J Hum Genet* 2002; **10**: 276–280.
- Steinberg S, de Jong S, Irish Schizophrenia Genomics Consortium, Andreassen OA, Werge T, Borglum AD *et al.* Common variants at VRR2 and *TCF4* conferring risk of schizophrenia. *Hum Mol Genet* 2011; **20**: 4076–4081.
- Wirgenes KV, Sonderby IE, Haukvik UK, Mattingsdal M, Tesli M, Athanasiu L *et al.* *TCF4* sequence variants and mRNA levels are associated with neurodevelopmental characteristics in psychotic disorders. *Trans Psychiatry* 2012; **2**: e112.
- Mudge J, Miller NA, Khrebtkova I, Lindquist IE, May GD, Huntley JJ *et al.* Genomic convergence analysis of schizophrenia: mRNA sequencing reveals altered synaptic vesicular transport in post-mortem cerebellum. *PLoS One* 2008; **3**: e3625.
- Navarrete K, Pedroso I, De Jong S, Stefansson H, Steinberg S, Stefansson K *et al.* *TCF4* (e2-2; ITF2): a schizophrenia-associated gene with pleiotropic effects on human disease. *Am J Med Genet B Neuropsychiatr Genet* 2013; **162B**: 1–16.
- Psychiatric GWAS Consortium Bipolar Disorder Working Group. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near *ODZ4*. *Nat Genet* 2011; **43**: 977–983.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>