



Published in final edited form as:

Mol Psychiatry. 2014 March ; 19(3): 276–277. doi:10.1038/mp.2013.33.

Functional SNPs are enriched for schizophrenia association signals

S-A Bacanu¹, J Chen¹, J Sun², K Richardson³, C-Q Lai³, Z Zhao², MC O'Donovan⁴, KS Kendler¹, X Chen¹

¹Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, VA, USA

²Departments of Biomedical Informatics and Psychiatry, Vanderbilt University School of Medicine, Nashville, TN, USA

³Nutrition and Genomics Laboratory, Jean Mayer United States Department of Agriculture Human Nutrition Research Center on Aging at Tufts University, Boston, MA, USA

⁴Department of Psychological Medicine, MRC Centre for Neuropsychiatric Genetics and Genomics, School of Medicine, Cardiff University, Cardiff, UK

Recent results suggest that there are likely to be many true association signals in large psychiatric genome-wide association study (GWAS) data sets.¹ The big challenge of the field is the identification of these risk genes of small effect. We attempt to address this issue by using Psychiatric Genetics Consortium (PGC) GWAS data sets²⁻⁴ to test whether functional variants from the GWAS are significantly enriched in low *P*-values when compared with the remainder of the GWAS data set.

We hypothesized that as functional single-nucleotide polymorphisms (SNPs) can directly affect gene expression and function, they are likely to significantly have an impact on disease risk. Therefore, functional SNPs are more likely to have true association signals. To test this hypothesis, we used genomic control- adjusted statistics⁵ from PGC schizophrenia,³ bipolar disorder⁴ and major depression disorder² GWAS data sets to investigate SNPs that (i) affect transcription factor-binding sites in the promoter regions (promoter SNPs), (ii) change gene expression via microRNA binding (microRNA SNPs) and DNA methylation (methylation SNPs), and (iii) correlate with gene expression (eQTL (expression quantitative trait loci) SNPs) in different brain regions.

The experimentally validated promoter SNPs were obtained from dbQSNP (<http://qsnp.gen.kyushu-u.ac.jp/>), whereas the microRNA, methylation and eQTL SNPs were obtained from the literature.^{6,7} We used all SNPs deemed statistically significant by the original studies (Supplementary Table S1). We extracted the association *P*-values for these functional SNPs from the PGC GWAS data sets, and tested if these SNPs were more enriched in association signals than the entire GWAS using the Simes and sum of squares

xchen@vcu.edu.

Supplementary Information accompanies the paper on the Molecular Psychiatry website (<http://www.nature.com/mp>)

tests (SST) (see Supplementary Methods). To account for the linkage disequilibrium amongst selected SNPs, the statistical significance of SST was assessed via 50 000 permutations based on the linkage disequilibrium patterns of the 288 European subjects sequenced by the 1000 Genomes Project.⁸ The results are summarized in Table 1. When applying SST, 6 of the 10 groups of SNPs (promoter SNPs, eQTL SNPs in the frontal cortex and cerebellum, methylation SNPs in the temporal cortex, pons and frontal cortex) showed significant enrichment above GWAS background for association signals for schizophrenia. Promoter SNPs were also significantly enriched in association with bipolar disorder, and showed a trend in major depression. No significant enrichments were detected by the Simes test (Supplementary Table S2).

Intuitively, the Simes test is useful to detect SNP sets having few strong signals, and SST is valuable at detecting SNP sets having many signals of small effect. Thus, our results suggest that there are no SNPs showing stronger associations than that observed in the GWASs. In contrast, SST results indicate that many promoter, eQTL and methylation SNPs seem to have a modest association with both schizophrenia and bipolar disorder. As the tested functional groups overlap, we conducted *post hoc* analyses for the unique SNPs amongst the significant groups. There were 11 868 unique SNPs amongst the methylation SNP groups, and they were significantly enriched for schizophrenia signals ($P<0.0001$). The 4178 unique eQTL SNPs were enriched for similar signals as well ($P=0.0006$). There were 127 and 49 shared SNPs between the promoter/methylation and between the promoter/eQTL groups (Supplementary Table S1), and both shared SNPs were enriched in schizophrenia association signals ($P=0.0014$ and 0.07 , respectively). A pathway analysis of these shared SNPs indicated that they are involved in DNA replication and repair, cell cycle, cellular development and proliferation, system development and function, neurological disease and immune response (Supplementary Tables S3 and S4).

In this study, we found that functional SNPs are more enriched in association signals than the remainder of the PGC GWAS data. The most significant enrichment was found in promoters for schizophrenia and bipolar disorder, and eQTL and methylation SNPs for schizophrenia. We noticed a trend for a decrease of enrichment signals across the disorders, a reflection of the power (sample size) of original studies. We also observed that there are differences in association signals across brain regions, supporting the notion that brain dysfunction in schizophrenia may be region-specific. Overall, our results suggested that for GWAS data sets with reasonable power, systematically selecting and testing functional SNPs may be an effective approach to identify risk genes with small but true effect. This conclusion may be extended to other complex diseases.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

REFERENCES

1. Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF et al. Nature 2009; 460: 748–752. [PubMed: 19571811]

2. Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium. *Mol Psychiatry* 2013; 18: 497–511. [PubMed: 22472876]
3. Ripke S, Sanders AR, Kendler KS, Levinson DF, Sklar P, Holmans PA et al. *Nat Genet* 2011; 43: 969–976. [PubMed: 21926974]
4. Sklar P, Ripke S, Scott LJ, Andreassen OA, Cichon S, Craddock N et al. *Nat Genet* 2011; 43: 977–983. [PubMed: 21926972]
5. Devlin B, Roeder K. *Biometrics* 1999; 55: 997–1004. [PubMed: 11315092]
6. Richardson K, Lai CQ, Parnell LD, Lee YC, Ordovas JM. *BMC Genomics* 2011; 12: 504. [PubMed: 21995669]
7. Gibbs JR, van der Brug MP, Hernandez DG, Traynor BJ, Nalls MA, Lai SL et al. *PLoS Genet* 2010; 6: e1000952. [PubMed: 20485568]
8. 1000 Genomes Project Consortium. *Nature* 2010; 467: 1061–1073. [PubMed: 20981092]

Table 1.

P-values for testing whether functional SNPs are enriched for association signals above PGC GWAS background.

Functional SNP group	Schizophrenia	Bipolar disorder	Major depression
Promoter SNPs	0.0013	<0.0001	0.0611
miRNA SNPs	0.0893	0.0329	0.1968
eQTL SNPs (temporal cortex)	0.0068	0.1687	0.1385
eQTL SNPs (pons)	0.0271	0.6235	0.2217
eQTL SNPs (frontal cortex)	0.0008	0.1866	0.5421
eQTL SNPs (cerebellum)	0.0005	0.1776	0.2114
Methylation SNPs (temporal cortex)	<0.0001	0.1981	0.4226
Methylation SNPs (pons)	<0.0001	0.0690	0.0613
Methylation SNPs (frontal cortex)	0.0003	0.0169	0.1979
Methylation SNPs (cerebellum)	0.0342	0.0372	0.0089

Abbreviations: eQTL, expression quantitative trait loci; miRNA, microRNA; SNP, single-nucleotide polymorphism.

P-values are computed empirically from 50 000 permutations using a sum of squares test. *P*-values significant after adjusting for multiple testing are highlighted in bold.