Short Review Article   **Open Access**

Commentary: Integration of gene expression and GWAS results supports involvement of calcium signaling in Schizophrenia

L. Hertzberg[1,2]*, E. Domany[1]

[1]Department of Physics of Complex Systems, Weizmann Institute of Science, Rehovot, Israel
[2]The Emotion-Cognition Research Center, Shalvata Mental Health Center, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel

Article Info

*Correspondence:
Shalvata Mental Health Center, Aliat Hanoar St. Hod Hasharon 45100, Israel; E-mail: libi.hertzberg@gmail.com; Telephone: +972-9-7478-554; Fax: +972-9-7478-623

ABSTRACT

The number of Genome Wide Association Studies (GWAS) of schizophrenia, as well as that of subjects studied have increased dramatically over the last years. Nevertheless, biologically meaningful interpretation of GWAS results remains very difficult due to the complexity and heterogeneity of both clinical and genetic aspects of schizophrenia, and the small contribution of each single gene to the disease. Our study presented a methodology that integrated GWAS results with gene expression data, and applied it to schizophrenia. Integration of two types of information, of both DNA-level and of messenger RNA (mRNA) level, increases the validity and reliability of the results and improves our understanding of the biological meaning of the GWAS results. We comment here on the methodology used and the main result of our study.

Methodology

The aim of GWAS studies is to identify Single Nucleotide Polymorphisms (SNPs) that are associated with a disease or phenotype of interest. Association of a SNP with a disease implicates genes that either reside near the genomic location of the SNP, or are regulated by a genetic factor located there. Since one believes that the contribution of each single gene to such a complex disease is small, it is natural to look for pathways and biological processes to which the implicated genes belong; altered functionality of these pathways may then be a causal factor in the development of the disease. One of the standard ways to go from single genes to pathways is by Pathway Enrichment Analysis (PEA)[1], for which a sufficient number of implicated genes is essential. Notably, the number of GWAS studies of schizophrenia has risen over the years, and more importantly, the numbers of participating subjects have witnessed a dramatic increase (from 479 in[2] to more than 40,000[3]), raising the number of reliably implicated loci.

Indeed, the GWAS of schizophrenia that we have analyzed[4] pointed at more than 100 loci associated with the disorder (and 30 more were added in 2017[3]); nevertheless, identifying the causal genes located near these loci remains a major challenge. An important reason is that the risk loci implicated by GWAS often span large genomic regions. For example, the MHC region showed the most significant association with schizophrenia[4]. However, this region contains a large number of highly linked SNPs spanning multiple genes (over 100). Due to the complex linkage disequilibrium between the implicated SNPs, it is difficult to identify

the genes that are most likely to play a causal role (see, for example,[5], that used the MAGMA tool[6] in order to translate GWAS results from SNP-level into gene-level). Focusing on loci that contain a single gene reduces the complexity inherent in dealing with multiple genes, some of which may not be associated with the disease, removes false positives and increases the fraction of true risk genes that contribute to the development of schizophrenia. Clearly this step may remove some risk genes, increasing the false negative rate. Moreover, the remaining disease-associated genes are reduced from more than 100 to few tens - too small a number to perform meaningful PEA. The number of implicated genes can be increased by choosing a lower threshold on GWAS-based identification; this, however, will also introduce many false positives, and hence derivation of robust and statistically meaningful results from PEA remains hindered.

In order to overcome this problem we integrated GWAS-derived loci that contain a single gene, with gene expression data, allowing us to obtain robust and statistically significant pathway analysis. Gene expression measurements were done on post-mortem human brain samples of subjects with schizophrenia and healthy controls[7]. The two types of data were integrated using the following steps:

1) GWAS-based disease associated genes were extracted from[4], including only associated loci that contain a single gene.

2) Clusters of co-expressed genes were identified on the basis of correlated expression profiles (as measured over all subjects).

3) We retained only those clusters of co-expressed genes whose size passed a test for statistical significance (estimated using correlated cluster size distribution derived by sampling from random groups of genes). Existence of co-expressed gene clusters of statistically significant sizes strengthens the validity of GWAS results, i.e., that the identified genes are indeed causal and contribute to the development of schizophrenia. It also suggests that there is a biological basis for the observed correlations; plausibly, genes with such correlations participate in common biological pathways.

4) Extend the list of genes that belong to such a cluster, by searching all expressed genes and adding those with highly correlated expression with the average expression profile of the cluster.

5) Perform PEA of the extended list of genes.

Extension of the list of implicated genes increased their number dramatically, way beyond what could be obtained on the basis of the GWAS results alone. This, in turn, generated robust and statistically significant results of the PEA. In addition, the focus on a cluster of co-expressed genes increases the probability that the genes belong to common biological pathways, and thus our method increases the likelihood of getting statistically significant and biologically meaningful results.

This methodology is simple, intuitive, and useful for the study of additional disorders (e.g.[8]), for which both GWAS results and gene expression datasets are available. For example, such data are available for Parkinson's disease[9, 10], Alzheimer's disease[11, 12], major depressive disorder[13, 14] and bipolar disorder[15, 16]. Since GWAS of these diseases were based on smaller patient cohorts, the list of implicated loci and genes is smaller than for schizophrenia. This can be partially compensated by choosing lower thresholds for GWAS.

While differential expression analysis necessitates a large number of samples, as each gene's contribution is small and hard to detect[17], we used correlation analysis, which yields statistically significant results even for a relatively small number of samples. In general, GWAS based implication of a gene suggests a causal (rather than correlative) relationship between the gene and the disease as stated in[18]: "Causality follows from the central dogma of biology (i.e., DNA variations lead to changes in transcription regulation/protein function, which in turn cause variations in disease phenotypes)". Since GWAS results constitute the seed and starting point of our analysis, it suggests that the enriched pathways play a causal role in the pathogenesis of schizophrenia. .

Simplicity, availability and relevance to many disorders are important advantages of the methodology presented above. However, a prominent shortcoming is the fact that the direction of the effect of the enriched biological pathways is not detected (i.e., enhanced or suppressed activity in the diseased state). This weakness limits our ability to point, on the basis of the involved biological pathways, at simple potential treatments (as discussed below).

The main result: support for a central role of calcium-signaling in the pathogenesis of schizophrenia

By combining GWAS results with gene expression analysis we found enrichment of calcium-signaling related pathways. Mounting evidence, mostly genetic (for example,[19]), points at involvement of calcium signaling in schizophrenia. It was proposed that altered calcium signaling may constitute the central unifying molecular pathology in the disorder[20]. We provided further, expression-based evidence for this involvement.

As noted above, the methodology we used does not detect the direction of the effect of the enriched biological pathways. In order to further explore that, we performed

differential expression analysis of calcium signaling genes in a region-specific manner. Preliminary results (unpublished data), however, detected opposite directions of differential expression in different brain regions; in some regions calcium signaling genes were up-regulated, while in others they were down-regulated. Calcium channel blockers are widely used for indications of cardiovascular disorders, and there is at least one approved compound that is thought to act via calcium channel activation[21]. Unfortunately, we could not formulate on the basis of our findings a clear hypothesis regarding the potential efficacy of these agents in schizophrenia. Notably, a few calcium channel blockers have been examined in clinical trials in schizophrenia with mixed results; however, these studies have been small, often poorly controlled, and commonly involved compounds with relatively poor central nervous system permeability[21].

Summary

Schizophrenia is a multifactorial disease, whose pathogenesis probably involves several biological pathways, comprising a complex network. Our analysis implicates calcium signaling and many additional biological pathways, several of which have already been linked to schizophrenia. As suggested by the high correlation of their expression patterns, these pathways might act in concordance in the pathogenesis of the disease. An attractive hypothesis, supported by our results, is that calcium signaling is a key factor in this network. However, the exact nature of the contribution of the detected pathways to schizophrenia and the interplay between them should be studied further. The methodology we applied, of integrating GWAS results with gene expression data, is simple, intuitive and relevant to additional disorders. As we have demonstrated, it has the potential to improve the understanding of the biological meaning of GWAS results.

References

1. Hung JH. Gene Set/Pathway Enrichment Analysis. Humana Press, Totowa, NJ; 2013. p. 201–13.

2. O'Donovan MC, Craddock N, Norton N, et al. Identification of loci associated with schizophrenia by genome-wide association and follow-up. Nat Genet. 2008; 40: 1053–5.

3. Li Z, Chen J, Yu H, et al. Genome-wide association analysis identifies 30 new susceptibility loci for schizophrenia. Nat Genet. 2017; 49: 1576–83.

4. Schizophrenia Working Group of the Psychiatric Genomics C. Biological insights from 108 schizophrenia-associated genetic loci. Nature. 2014; 511: 421–7.

5. Schijven D, Kofink D, Tragante V, et al. Comprehensive pathway analyses of schizophrenia risk loci point to dysfunctional postsynaptic signaling. Schizophr Res. 2018.

6. de Leeuw CA, Mooij JM, Heskes T, et al. MAGMA: Generalized Gene-Set Analysis of GWAS Data. PLOS Comput Biol. 2015; 11: e1004219.

7. Hertzberg L, Katsel P, Roussos P, et al. Integration of gene expression and GWAS results supports involvement of calcium signaling in Schizophrenia. Schizophr Res. 2015; 164: 92–9.

8. Belaish S, Domany E, Hertzberg L, et al. Depression: Integration of GWAS results and gene expression data. 2018; Work in progress.

9. Nalls MA, Pankratz N, Lill CM, et al. Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. Nat Genet. 2014; 46: 989–93.

10. PAPAPETROPOULOS S, FFRENCH-MULLEN J, MCCORQUODALE D, et al. Multiregional Gene Expression Profiling Identifies MRPS6 as a Possible Candidate Gene for Parkinson's Disease. Gene Expr. 2006; 13: 205–15.

11. Hokama M, Oka S, Leon J, et al. Altered expression of diabetes-related genes in Alzheimer's disease brains: the Hisayama study. Cereb Cortex. 2014; 24: 2476–88.

12. Lambert JC, Ibrahim-Verbaas CA, Harold D, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet. 2013; 45: 1452–8.

13. Ripke S, Wray NR, Lewis CM, et al. A mega-analysis of genome-wide association studies for major depressive disorder. Mol Psychiatry. 2013; 18: 497–511.

14. Labonté B, Engmann O, Purushothaman I, et al. Sex-specific transcriptional signatures in human depression. Nat Med. 2017; 23: 1102–11.

15. Group PGCBDW, Sklar P, Ripke S, et al. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. Nat Genet. 2011; 43: 977–83.

16. Ryan MM, Lockstone HE, Huffaker SJ, et al. Gene expression analysis of bipolar disorder reveals downregulation of the ubiquitin cycle and alterations in synaptic genes. Mol Psychiatry. 2006; 11: 965–78.

17. Fromer M, Roussos P, Sieberts SK, et al. Gene expression elucidates functional impact of polygenic risk for schizophrenia. Nat Neurosci. 2016; 19: 1442–53.

18. Zhu J, Lum PY, Lamb J, et al. An integrative genomics approach to the reconstruction of gene networks in segregating populations. Cytogenet Genome Res. 2004; 105: 363–74.

19. Li W, Fan CC, Mäki-Marttunen T, et al. A molecule-based genetic association approach implicates a range of voltage-gated calcium channels associated with schizophrenia. Am J Med Genet B Neuropsychiatr Genet. 2018.

20. Lidow MS. Calcium signaling dysfunction in schizophrenia: a unifying approach. Brain Res Brain Res Rev. 2003; 43: 70–84.

21. Lencz T, Malhotra AK. Targeting the schizophrenia genome: a fast track strategy from GWAS to clinic. Mol Psychiatry. 2015; 20: 820–6.