Literature Data Mining and Enrichment Analysis on Top 235 Genes for Attention Deficit Hyperactivity Disorder

Shunan Li; , Karan Kapoor; , Lydia C Manor

Journal of Psychiatry and Brain Science 2016,1(2);3
DOI:10.20900/jpbs.20160008; June 25, 2016

ABSTRACT

Background: Attention deficit hyperactivity disorder (ADHD) is a psychiatric disorder of the neuro-developmental type, marked by an ongoing pattern of inattention or hyperactivity/impulsivity, which interferes with functioning or development. The disorder affects approximately 5-7 % children and 2-5 % of adults worldwide. Numerous studies have indicated that genetic factors predominate the causes for ADHD. Nevertheless, no systematic study has summarized these findings and provided an objective and complete list of genes with a reported association to ADHD.

Methods: Literature and enrichment metrics analyses were used to discover genes of specific significance associated with ADHD. We conducted a literature data mining (LDM) of over 2,410 articles covering publications from Jan. 1988 to Apr. 2016, where 235 genes were reported to be associated with the disease. Then we performed a gene set enrichment analysis (GSEA) and a sub-network enrichment analysis (SNEA) to study the functional profile and pathogenic significance of these genes associated with ADHD. Lastly, we performed a network connectivity analysis (NCA) to study the associations between the reported genes.

Results: 181/235 genes enriched 100 pathways (p<1.1e-007), demonstrating multiple associations with ADHD. Twelve genes were discovered to be associated with ADHD, in terms of both functional diversity and replication frequency, including SLC6A3, DRD4, BDNF, DRD2, HTR2A, DBH, HTR1B, DRD5, GRM7, DRD3, TH and GRIN2A. In addition, one novel gene, SHANK2, was suggested worthy of further study. Moreover, SNEA and NCA results indicated that many of these genes form a functional network, playing roles in the pathogenesis of other ADHD related disorders.

Conclusion: Our results suggest that the genetic causes of ADHD are linked to a genetic and functional network composed of a large group of genes. The gene lists, together with the literature and enrichment metrics provided in this study, could serve as roundwork for further biological/genetic studies in the field.