Evaluation of WGCNA and NERI methods for prioritization of pathways associated to schizophrenia spectrum disorders

Arthur Sant'anna Feltrin[1], Ana Carolina Tahira[2], Sérgio Nery Simões[3], Helena Brentani[2], David Correa Martins Jr[1]

1 UFABC
2 USP
3 INSTITUTO FEDERAL DO ESPÍRITO SANTO

Abstract

Using two gene expression data related to schizophrenia, we proposed a new approach consisting of combining the results of two network analysis algorithms: Weighted Gene Correlation Network Analysis (WGCNA) and Network-Medicine Relative Importance (NERI). Considering the differences between the two methods, our hypothesis is that both are capable of producing compelling results related to different aspects of schizophrenia's biological pathways; therefore, are complementary to each other. For that, we used replication and enrichment analysis using public databases. WGCNA uses gene expression from two groups to build co-expression pairwise correlation matrices, using connectivity parameters for evaluation of the network. NERI also uses expression data, but its network construction is based on the integration of PPI databases, gene expression, and a previously chosen seed genes list; the network analysis are based on shortest ranking path and relative importance calculation. We conducted an enrichment analysis using DAVID for the identification of partial biological function of each result, as well a replication and MSET analysis (for GWAS, transcriptome, methylation and de novo mutation databases related to schizophrenia) to appraise the replication and accuracy of our new approach when compared with each method in separate. The WGCNA module represents a final network of 435 and 300 genes on BAHN and KATO expression data. The enrichment analysis of this group using ppi modules leads to 88 genes across 10 hyper-represented human modules (adj.p<0.05), mostly involving immunological process. By using NERI, the final gene list was 150 genes for both BAHN and KATO with the enrichment analysis leading to modules related to glutamate receptor signaling, apoptotic process, neurotrophin and MAPK pathways. Both methods achieved statistical relevant replication results (p<0.05), but with one gene shared between both methods results. In the MSET analysis, NERI was capable for achieve meaningful results for the methylation and de novo mutation databases; whether our proposal of combining both results achieved better results for these two databases and for transcriptome, also increasing the number of candidate genes of each list. Our study suggests that using both methods in combination could be a promising approach for establishing a group of modules and pathways related to schizophrenia (or any complex disease).