Modeling the Cumulative Genetic Risk for Multiple Sclerosis from Genome Wide Association Data.

ABSTRACT: BACKGROUND: Multiple Sclerosis (MS) is the most common cause of chronic neurologic disability beginning in early to middle adult life. Results from recent genome-wide association studies (GWAS) have substantially lengthened the list of disease loci and provide convincing evidence supporting a multifactorial and polygenic model of inheritance. Nevertheless, the knowledge of MS genetics remains incomplete, with many risk alleles still to be revealed. METHODS: We used a discovery GWAS dataset (8,844 samples, 2,124 cases and 6,720 controls) and a multi-step logistic regression protocol to identify novel genetic associations. The emerging genetic profile included 350 independent markers and was used to calculate and estimate the cumulative genetic risk in an independent validation dataset (3,606 samples). Analysis of Covariance (ANCOVA) was implemented to compare clinical characteristics of individuals with various degree of genetic risk. Gene ontology and pathway enrichment analysis was done using the DAVID functional annotation tool, the GO Tree Machine, and the Pathway-Express profiling tool. RESULTS: In the discovery dataset, the median cumulative genetic risk (P-Hat) was 0.903 and 0.007 in the case and control groups, together with 79.9% classification sensitivity and 95.8% specificity. The identified profile shows a significant enrichment of genes involved in the immune response, cell adhesion, cell communication/signaling, nervous system development, and neuronal signaling, including ionotropic glutamate receptors, which have been implicated in the pathological mechanism driving neurodegeneration. In the validation dataset, the median cumulative genetic risk was 0.59 and 0.32 in the case and control groups, with classification sensitivity 62.3%, and the specificity 75.9%. No differences in disease progression or T2-lesion volumes were observed among four levels of predicted genetic risk groups (High, Medium, Low, Misclassified). On the other hand, a significant difference (F=2.75, p=0.04) was detected for age of disease onset between the affected misclassified as controls (mean=36 yrs.) and the other three groups (High: 33.5 yrs.; Medium: 33.4 yrs.; Low: 33.1 yrs.). CONCLUSION: The results are consistent with the polygenic model of inheritance. The cumulative genetic risk established using currently available genome-wide association data provides important insights into disease heterogeneity and completeness of current knowledge in MS genetics.

ISSN/ISBN 1756-994X