IGNG1-IGNG3 locus, its possible role in the multiple sclerosis, and biases SNP allele frequencies for europeoid and non-europeoid populations

Bykadorov P.A[1], Fridman M.V.[1*] , Oparina N.J.[2], Makeev V.J.[1]

[1] – Vavilov Institute of General Genetics, Moscow

[2] - *IMBIM, Uppsala University*

**e-mail: marina-free@mail.ru*

A large amount of GWAS data on multiple sclerosis (MS) has been obtained recently almost exclusively for populations dominated by Caucasians. Unfortunately, SNPs that display significant association with MS may only be linked to those that are mechanistically related to the disease development. In addition, GWAS does not provide any idea on the mechanisms of the SNP influence. Such information can obtained from eQTLs, but an eQTL also usually marks only a fairly large locus. Therefore, to study genome segments functionally associated with disease developement it is relevant to identify a rather long genome segments containing several SNPs and perform enrichment analysis for the features associated with the segment.

The idea of our approach is to take into account SNPs displaying low association with the target feature, and discard the corresponding genome segments, thus reducing the target regions. We have generated a set of loci enriched with SNPs associated with the MS development. Then, we created two SNP lists, statistically associated and non-associated with the development of MS according to all GWAS data. SNPs strongly linked with SNPs associated with MS were added to the target list, whereas SNPs linked with SNPs displaying low MS association were discarded. The second list contained SNPs simultaneously linked with two SNPs, statistically associated with MS according to all GWAS data set, the third list contained SNPs linked with three MS associated SNPs. Surprisingly, all SNPs in the final list were found in one locuse in chromosome 14, containing IGHG1 and IGHG3 genes and in several loci in chromosome 6 (containing HLA genes).

Out of 66 SNPs found in the chr 14 segment 8 are very frequent in Europeans (>0,3) and very rare in Africans and Asians (<0,03). It is noteworthy that out of 18 299 SNPs not associated with MS or linked with non-associated SNPs only 13 displayed such a bias in population frequencies. Conversly, out of 7 524 SNPs associated with MS or linked with associated SNPs 14 displayed such population frequency bias, and of these all but one were found in the locus we identified in

the chr 14 or in its immediate vicinity. The frequency of alternative variants in this locus is 0.43-0.62 for the European population.

We analyzed which SNPs from our lists were local or distant eQTLs for genes expressed in blood and nervous system. SNPs found in chromosome 6 are primary related to HLA genes, so we also discarded these SNPs for not to jam the weaker signal from other SNPs. We created the lists of the dependent genes and analyzed their enrichment for GO functions and pathways. We discovered that antigen processing and presentation with MHC II (and related functions) were overrepresented in both tissue. The list for SNPs linked with three MS associated SNPs, contained eQTLs for all overrepresented functions. The overrepresentation of these functions was primarily associated with the effect on the expression of HLA genes.

As for the mechanism, we propose that IgG increases antigen presentation by interacting with FcRγ-receptors (1), and stimulates B-cell secondary immune responce to IgG synthesis by activating T-helper cells of antigenic presenting complex. Some of these processes may occur subdurally, which is indirectly confirmed by eQTL similarity in the blood and nervous system. More to the point, the given locus is associated with IgG index (the ratio of concentrations of IgG in the cerebrospinal fluid and serum as compared with the same ratio for albumin) (2, 3).

SNPs found in the locus discussed affect expression of genes HLA-DOB, HLA-DQA1, HLA-DQA2, HLA-DQB1 in blood and nervous system. It is shown that the MS protective allele HLA II DQB1 has specific impact on the presentation of myelin basic protein, the most significant for the MS development, without causing comparable impact on other proteins (4).

As it is known, the MS frequency is about ten-fold higher in countries with a predominance of the Caucasian population. In addition, the disease development sometimes is different for non-Europeans. For instance, the abnormal intrathecal synthesis of IgG, reflected by cerebrospinal fluid oligoclonal IgG bands and increased IgG index, is much less frequent in Japanese (5). We assume that the given locus is responsible for the corresponding differences. It is also known that relocation to higher latitudes, especially at the age of 15, can increase the risk of multiple sclerosis (6). In other words, habitation in high latitudes causes similar modification and genetic changes, presumably adaptive, the by-product of which is predisposition to MS.

1. Getahun A. Heyman B (2004) IgG- and IgE-mediated antigen presentation on MHC class II, *Immunol. Lett.,* **92**: 33-38.
2. Goris A., Pauwels I., Gustavsen M.W. et al. (2015) Genetic variants are major determinants of CSF antibody levels in multiple sclerosis, *Brain*, **138**: 632-643.

3. Buck D., Albrecht E., Aslam M. et al. (2013) Genetic variants in the immunoglobulin heavy chain locus are associated with the IgG index in multiple sclerosis, *Ann.Neurol.*, **73:** 86-94.
4. Kumar A., Melis P., Genna V., et al. (2014) Antigenic peptide molecular recognition by the DRB1-DQB1 haplotype modulates multiple sclerosis susceptibility, *Mol.Biosyst.*, **10**: 2043-2054.
5. Yoshimura S., Isobe N., Matsushita T. et al (2014) Genetic and infectious profiles influence cerebrospinal fluid IgG abnormality in Japanese multiple sclerosis patients. *PLoS One,* **9**: e95367.
6. Compston A., Coles A. (2008) Multiple sclerosis. *Lancet* **372**:1502-17.