Association analysis of the multiple sclerosis susceptibility gene TNFSF14 in the Italian population: high density fine mapping and functional analysis of the identified primary associated variant

M. Zuccalà*[1], N. Barizzone*[1], M. Sorosina*[2] R. Bordoni[3], F. Clarelli[2], S. Anand[3], E. Mangano[3], F. Esposito[2], E. Corsetti[1], PROGEMUS, PROGRESSO, IMSGC, D. Vecchio[4], G. Predebon[1], R. Cantello[4], V. Martinelli[2], G. Comi[2], M. Leone[4], G. De Bellis[3], F. Martinelli-Boneschi*[2], S. D'Alfonso*[1];

1. Avogadro University, Novara, Italy,
2. Laboratory of Human Genetics of Neurological Diseases, San Raffaele Scientific Institute, Milano, Italy,
3. National Research Council of Italy, Institute for Biomedical Technologies, Segrate, Milano, Italy, ,
4. MS Centre, SCDU Neurology, AOU Maggiore della Carità, Novara, Italy,

Abstract

We performed a Genome Wide Association Study on Multiple Sclerosis (MS) in the Italian population (1727 MS and 2258 healthy controls (HC)): the strongest non-HLA signal was from an intronic variant in the Tumor Necrosis Factor (ligand) superfamily member 14 (*TNFSF14*) gene (p=5.9e-8), located in a known MS associated region. In order to define the primary associated variant in this region, we sequenced the whole region (17.5kb) using NGS on 588 MS and 408 HC from the Italian population, pooled in groups of 12 samples and performed an association study in the pool according to a method validated by our group.

After QC, we detected: 63 private variants and 50 with an MAF>1%; 6 variants were in the coding region and 15 showed a significant (p<0.05) association with MS.

A representative subset of these variants (N=59, including all common, MAF>1%, significantly associated and coding variants) were individually genotyped in an independent sample set (867 MS and 878 HC) in order to perform a fine-mapping of the gene. We observed a significant association for 7 variants, confirming 5 of the associations observed in pools. The strongest associations were the known MS risk variant in intron 1 (p=6.2e-5) and a synonymous variant (p=4.6e-4).

Conditional analysis on 1677 MS and 1680 HC showed that the intronic variant is the primary associated one.

TNFSF14 mRNA expression was lower in 84 MS compared to 80 HC (p=0.031) and individuals in homozygous for the MS risk allele (p=1.1e-4). This SNP maps in intron 1 in a region defined as "active promoter" (UCSC) and creates a consensus site for the AHR (Aryl hydrocarbon receptor) transcription factor (TRANSFAC). These observations are supported by eQTL data from public databases on lymphoblastoid cell lines from Geuvadis consortium (465 samples, p=1,22E-09), Bioportal (270 samples, p=0,034) and Gtex portal (114 samples, p=0,0081). The same effect was observed on PBMCs (338 samples, p=0,0026) and hippocampus (81 samples, p=0,012) in Gtex portal. Not significant eQTl association was found on Brain eQTL Almanac (Braineac) (data from 133 samples from ten different brain tissues). These data are also consistent with the observation that in hetherozygous individuals the allelic expression is unbalanced in favor of the allele with higher eQTL expression (Wilcoxon paired-samples test: p<0.0001, RNAseq on 97 samples from Geuvadis consortium).

In conclusion, we defined the association in the *TNFSF14* region at the single variant level and propose a functional role of the associated variant.