Journal of Autoimmunity 38 (2012) J144–J155

Contents lists available at SciVerse ScienceDirect

Journal of Autoimmunity

journal homepage: www.elsevier.com/locate/jautimm

Gender-based blood transcriptomes and interactomes in multiple sclerosis: Involvement of SP1 dependent gene transcription

Ramesh Menon [a,b,1], Marco Di Dario [a,b,1], Chiara Cordiglieri [a,b], Silvia Musio [b], Loredana La Mantia [b], Clara Milanese [b], Anna Luisa Di Stefano [c], Massimo Crabbio [c], Diego Franciotta [d], Roberto Bergamaschi [c], Rosetta Pedotti [b], Enzo Medico [e], Cinthia Farina [a,b,*]

[a] Institute of Experimental Neurology (INSpe), Division of Neuroscience, San Raffaele Scientific Institute, Milan, Italy
[b] Neuroimmunology and Neuromuscular Disorders Unit, Foundation IRCCS Neurological Institute Carlo Besta, Milan, Italy
[c] Department of Neurological Sciences, Foundation IRCCS Neurological Institute Casimiro Mondino, Pavia, Italy
[d] Laboratory of Neuroimmunology, Foundation IRCCS Neurological Institute Casimiro Mondino, Pavia, Italy
[e] Laboratory of Functional Genomics, Institute for Cancer Research and Treatment (IRCC), University of Turin Medical School, Candiolo, Italy

ARTICLE INFO

Article history:
Received 6 October 2011
Received in revised form
11 November 2011
Accepted 12 November 2011

Keywords:
Multiple sclerosis
Transcriptomics
Gender
Systems biology
SP1 transcription factor

ABSTRACT

In this study we investigated the contribution of gender to global gene expression in peripheral blood mononuclear cells from multiple sclerosis (MS) patients and healthy controls. We observed that, in contrast to the conventional approach, gender-based case-control comparisons resulted in genelists with significantly reduced heterogeneity in human populations. In addition, MS was characterized by significant changes both in the quantity and in the quality of the sex-specific genes. Application of stringent statistics defined gender-based signatures which classified a second independent MS population with high precision. The global unsupervised cluster analyses for 60 subjects showed that 29/31 female and 27/29 male samples were properly identified. Notably, MS was associated in women and in men with distinct gene signatures which however shared several molecular functions, biological processes and interactors. Issues regarding epigenetic control of gene expression appeared as the main common theme for disease, with a central role for the functional modules related to histone deacetylase, NF-kappaB and androgen receptor signaling. Moreover, *in silico* analyses predicted that the differential expression in MS women and men were depending on the transcription factor SP1. Specific targeting of this pathway by the bis-anthracycline WP631 impaired T cell responses *in vitro* and *in vivo*, and reduced the incidence and the severity of experimental autoimmune encephalomyelitis, indicating that SP1 dependent gene transcription sustains neuroinflammation.

Thus, the gender-based approach with its reduced heterogeneity and the systems biology tools with the identification of the molecular and functional networks successfully uncovered the differences but also the commonalities associated to multiple sclerosis in women and men. In conclusion, we propose gender-based systems biology as a novel tool to gain fundamental information on disease-associated functional processes.

1. Introduction

Multiple sclerosis (MS) is a chronic demyelinating disorder of the central nervous system (CNS) characterized by the presence of inflammatory cells and mediators within nervous tissue [1,2]. It is a chronic disease with onset in young adulthood, prevalence of 1:500 in some geographical regions and predominance in women (2:1 female/male ratio) with the relapsing-remitting (RR-MS) form of disease [3]. Diagnosis of MS is difficult due to the variability of clinical symptoms patients experience and to the resemblance with other neurological disorders of CNS. Currently, diagnostic criteria are mainly based on clinical and radiological examinations, as specific laboratory tests are not yet available. Furthermore, there is still no single biomarker that correlates accurately with clinical activity or treatment response [4,5].

Multiple sclerosis is a complex disease determined by both genetic and environmental factors. Intriguingly, many of the susceptibility genes play a role in immune system. For instance, classical linkage

* Corresponding author. Institute of Experimental Neurology (INSpe), DIBIT II - 1st Floor, San Raffaele Scientific Institute, via Olgettina 58, 20132 Milan, Italy. Tel.: +39 2 26436180.
E-mail address: farina.cinthia@hsr.it (C. Farina).
[1] These authors contributed equally to the work.

0896-8411/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jaut.2011.11.004

analysis in large families with MS pointed out mainly the importance of the major histocompatibility complex locus in determining predisposition to the disease [6]. More recently, application of high-throughput screening techniques in genetics association studies [6] identified additional polymorphisms in immune related genes such as interleukin (IL) 7 receptor and IL2 receptor [7]. Although MS affects the CNS, there are evidences of altered immunity in the periphery in MS patients [2]. Further, the most widely used therapeutic drugs in MS are either immunosuppressive or immunomodulatory agents [2,8], indicating that targeting peripheral immune system is beneficial to patients with this CNS disorder. These observations sustain the rationale for employing peripheral blood mononuclear cells (PBMC) as an easily accessible and informative source of biological material in transcriptome studies. Transcriptional profiling aims to comprehend complex molecular interactions and to identify disease related biomarker. Such information may lead to new hypotheses that either would not be considered based on current knowledge or which are too complex to be examined by conventional approaches. Here, we performed microarray-based gene expression profiling on peripheral blood mononuclear cells (PBMC) derived from 23 MS patients and 22 healthy controls (CTRL). To reduce variability, we included patients with the relapsing-remitting clinical course, who were free from immunomodulatory treatments and from any other acute or chronic inflammatory disorder. Distinct from the conventional disease transcriptomics approach, we applied a gender-based approach for selection of differentially expressed genes (DEG) in MS and validated them in a second independent MS population. Further, we extended bioinformatical annotation to the systems biology level and verified the involvement of a novel pathway in the animal model of multiple sclerosis.

2. Materials and methods

2.1. Inclusion criteria for RR-MS patients and healthy controls

Clinical investigations have been conducted according to the principles expressed in the Declaration of Helsinki and peripheral blood was drawn after signing of the institutional informed consent. Twenty-three MS patients were initially enrolled for this study. They were Italian adult subjects comprised of 13 women and 10 men. They had relapsing-remitting course of the disease diagnosed according to the McDonald criteria [9] and all of them were clinically stable at the time of blood sampling. Moreover, they were not suffering from any acute or chronic inflammatory diseases or other autoimmune disorders and they had not yet started any immunomodulatory therapy for MS. Sampling was performed at least 4 weeks after the last clinical attack or steroid treatment. All participants had peripheral blood counts within the reference range. A second MS population (comprised of 10 women and 8 men) was subsequently enrolled according to the same inclusion criteria. Twenty-two healthy subjects (comprised of 10 women and 12 men), who had no acute or chronic inflammatory diseases or autoimmune disorders, were included as controls for this study. All blood samplings were performed between 9 and 12 a.m.

2.2. PBMC isolation and RNA isolation

Peripheral blood mononuclear cells (PBMC) were isolated using a discontinuous density gradient (Lymphoprep, Nycomed, Oslo, Norway). Viable cells were counted by Trypan Blue (Sigma–Aldrich, Milan, Italy) exclusion. Then total RNA was extracted using TriReagent (Ambion, Applied Biosystems, Monza, Italy) and stored at −80 °C.

2.3. The microarray experiment and data processing

Total RNA extracted from PBMC was used for microarray experiments on Illumina Human_Ref-8_V2 arrays (Illumina, Son, Netherlands). Quantification and quality analyses of RNA were performed on a Bioanalyzer 2100 (Agilent, Milan, Italy). Reverse transcription and biotinylated cRNA synthesis were performed using the Illumina TotalPrep RNA Amplification Kit (Ambion), according to the manufacturer's protocol. Hybridization of the cRNAs was carried out on Illumina Human_Ref-8_V2 arrays. These arrays contain about 22,000 probes exploring the transcripts contained in the Refseq database. Array hybridization, washing, staining and scanning in the Beadstation 500 (Illumina) were performed according to standard Illumina protocols. The BeadStudio software (Illumina) was used to analyze raw data grouped by experimental condition. The data were normalized using cubic spline method as implemented in the software. The MIAMI compliant microarray data have been deposited in the EBI ArrayExpress database (Accession no. E-MTAB-380). The Beadstudio software reports a detection p-value, which represents the confidence that a given transcript is expressed above the background defined by negative control probes. This value clues the "absent" or "present" status of transcripts in the array. In this study, we considered a probe to be "present" in the array if at least 2/3 of the samples in both control and disease phenotypes had a detection p-value < 0.05. With this paradigm we removed the potential unfavorable probes from the analysis and proceeded with 10,390 probe sets for the study. In addition, we identified three outlier samples through both the principal component analysis (PCA) and hierarchical sample clustering methods using MeV package 4.3.01 [10], and removed those samples from further analysis. There were no significant differences in age (36.9 ± 11.3 vs. 35.5 ± 11.4 in female and male controls, 36.5 ± 7.5 vs. 36.3 ± 7.8 in female and male patients), disease duration (3.8 ± 3.2 vs. 6.1 ± 5.5 in female and male patients), EDSS score (1.7 ± 1.4 vs. 1.8 ± 1.5 in female and male patients) and annual relapse rate (1.5 ± 1.6 vs. 1.0 ± 0.9 in female and male patients) between the two gender groups.

For the validation group comprising of 10 female and 8 male RR-MS samples (Age: 40.5 ± 14.1 vs 36.5 ± 8.6; EDSS: 1.6 ± 0.7 vs. 1.0 ± 1.0; Disease duration: 6.9 ± 8.4 vs. 8.4 ± 6.3; Annual relapse rate: 1.1 ± 1.2 vs. 0.8 ± 0.4 in female and male patients respectively), the experimental protocols for RNA extraction, labeling and array hybridizations were performed as described above. In order to reduce the technical variation due to hybridizations made at distinct time points, batch correction was performed for the newly recruited samples using the dChip software [11].

2.4. Analysis of differential gene expression

Toward the identification of differentially expressed genes, the samples were divided into MS and healthy control groups, each comprising of 21 samples. We used three parallel statistical methods to identify the differentially expressed genes. A differentially expressed gene was the one that passed at least two tests. The first method was a two sample t-test performed on MS and healthy controls with a p-value threshold of 0.01. The t-test was performed with Welch approximation for unequal variances in the two groups using the MeV package. Secondly, we used *Significant Analysis of Microarrays* [12], a robust permutation based non-parametric method that relies on variance information present in measurements obtained from the probes. Unlike t-test, SAM outputs the q-value which represents the significance of the differential expression for a given gene. In our study, the q-value cut-off was 20% and used the SAM implementation in *Stanford tools for Excel* software, version 1.1. Finally, a bayesian approach to identify differentially