Neurobiology of Disease 37 (2010) 613–621

Contents lists available at ScienceDirect

Neurobiology of Disease

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

Gene expression profiling in multiple sclerosis: A disease of the central nervous system, but with relapses triggered in the periphery?

Boel Brynedal [a,*], Mohsen Khademi [b], Erik Wallström [b], Jan Hillert [a], Tomas Olsson [b], Kristina Duvefelt [c]

[a] The Multiple Sclerosis Research Group, Centre for Molecular Medicine, Department of Clinical Neuroscience, Karolinska Institutet, CMM L8:00, 171 76 Stockholm, Sweden
[b] Neuroimmunology Unit, Centre for Molecular Medicine, Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden
[c] Mutation Analysis Facility, Clinical Research Centre, Karolinska University Hospital, Stockholm, Sweden

ARTICLE INFO

Article history:
Received 14 August 2009
Revised 29 October 2009
Accepted 18 November 2009
Available online 26 November 2009

Keywords:
Gene expression profiling
Microarray
Multiple sclerosis
Relapse
CSF
PBMC

ABSTRACT

The aetiology of multiple sclerosis (MS), an autoimmune demyelinating disease of the central nervous system (CNS), includes both genetic and environmental factors, but the pathogenesis is still incompletely known. We performed gene expression profiling on paired cerebrospinal fluid (CSF) and peripheral blood mononuclear cells (PBMCs) samples from 26 MS patients without immunomodulatory treatment, sampled in relapse or remission, and 18 controls using Human Genome U133 plus 2.0 arrays (Affymetrix). In the CSF, 939 probe sets detected differential expression in MS patients compared to controls, but none in PBMCs, confirming that CSF cells might mirror the disease processes. The regulation of selected transcripts in CSF of MS patients was confirmed by quantitative PCR. Unexpectedly however, when comparing MS patients in relapse to those in remission, 266 probe sets detected differential expression in PBMCs, but not in CSF cells, indicating the importance of events outside of the CNS in the triggering of relapse.

Introduction

Multiple sclerosis (MS) is a complex disease affecting the central nervous system (CNS) where both genetic and environmental factors influence susceptibility. The disease is characterised by acute focal inflammation, demyelination and axonal loss with limited remyelination, resulting in sclerotic plaques. The initial disease course is usually characterized by relapses (disease bouts) and remission (periods of recovery) and eventually most patients will enter a secondary progressive phase. A small proportion of cases (~5% in Sweden) show a primary progressive (PP) course.

MS has both autoimmune and neurodegenerative features, but the pathophysiological processes that may occur both within and outside of the CNS are neither defined nor uniformly distributed within the MS population. Expression profiling enables an unbiased survey of the transcriptome for genes with altered transcript levels in disease states, with the purpose of formulating mechanistic hypotheses for further investigation. In addition, the regulation of transcripts within a pathway, or attempts to investigate the transcription factors that drive the differential expression, might decipher the ongoing biological processes.

So far, a limited number studies has been published regarding gene expression profiling in MS, either focusing on the effect of treatment on gene expression (Kauffman et al., 2009; Sellebjerg et al., 2008; Singh et al., 2007) or on the global gene expression in blood cells from a few MS patients (Achiron et al., 2007; Arthur et al., 2008; Avasarala et al., 2008; Bomprezzi et al., 2003; Malmestrom et al., 2008; Sarkijarvi et al., 2006; Satoh et al., 2008; Satoh et al., 2006).

Ideally, an adequate study of the transcriptome should reflect the target organ, i.e., the central nervous system. However, CNS samples from MS patients are scarce and limited by methodological constraints. Post mortem samples poorly reflect the stage of disease where immunomodulation is likely to be efficient, and brain biopsies are, by definition, obtained from patients with odd clinical presentations. Therefore, cerebrospinal fluid (CSF) cells have long been used as a surrogate for the target organ itself in CNS disorders. In the CSF, on the other hand, material is scarce, with the possibility to obtain only relatively few cells. Up till now therefore, methodological constraints have made transcriptome studies difficult. Here, we made use of improved purification procedures and a new amplification methodology allowing the analysis of global gene expression in scarce materials such as CSF cells.

Importantly, we possess a unique CSF bank from several hundred newly diagnosed MS patients, all with matched blood samples, allowing a careful selection of cases for optimal comparisons. Therefore, we were able to investigate the possible differential regulation of gene expression both centrally (CSF) and peripherally (peripheral blood mononuclear cells (PBMCs)). Apart from detecting single differentially expressed transcript, our scope was to investigate regulation within pathways, enrichment of transcription factor binding sites and the connectivity between differentially expressed transcripts. Through these efforts we were able to show that the

* Corresponding author. Fax: +46 8 517 73909.
E-mail address: boel.brynedal@ki.se (B. Brynedal).
Available online on ScienceDirect (www.sciencedirect.com).

0969-9961/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.nbd.2009.11.014

expression in CSF cells distinguishes samples from MS patients from those received from controls, and that PBMCs could differentiate samples drawn during a relapse.

Materials and methods

Study subjects

Global gene expression in both CSF cells and PBMC samples was assessed in 26 MS patients, diagnosed with MS (McDonald et al., 2001) at the Karolinska University Hospital. Twelve of these individuals were sampled during a disease bout (verified by neurologist at Karolinska University Hospital) initiated no more than 2 weeks earlier (oral information from patient). The remaining 14 MS patients were sampled during remission, at least 6 months after an earlier relapse. The MS patients were selected from a large cohort of newly diagnosed MS patients collected prospectively in the Stockholm area, aiming at an even number of patients sampled in relapse and remission. None of the patients had ever received immunomodulatory drugs. A control population consisting of 18 individuals with other neurological diseases of a non-inflammatory kind was used to assess MS specific expression profiles. The control population was diagnosed with psychosis ($n=4$), vertigo ($n=2$), paresthesia ($n=2$) cervical spinal stenosis, ependymoma, headache, hereditary spastic paresis, migraine, musclepain, neurasthenia, sensory symptoms, vertebral dissection or dissociative motor disorder. All individuals were Scandinavians living in the Stockholm area. Patients sampled during a relapse were between 16 and 68 years old (mean 37.1), while MS patients sampled during remission were between 22 and 47 years (mean 34.4). Female-to-male ratio was 1 in relapse patients and 1.8 among patients in remission. In total, the mean age among MS patients was 39.2 years and in controls 40.6 years (range 26–57). The female-to-male ratio was 1.3 among MS patients and 1.2 in controls. In total, 88 samples were hybridized to expression arrays.

Independent study populations were utilized for the quantitative PCR validation. The CSF samples were collected from 30 MS patients with mean age of 33.2 years (range 20–52), female-to-male ratio of 2.8 where 10 patients were sampled during a relapse, and from 19 controls (non inflammatory OND) with mean age 30.5 years (range 21–44) and a female-to-male ratio of 2.2. The PBMC samples were collected from 12 additional MS patients in relapse and 20 patients in remission with mean age 33.1 years (range 20–52) and a female-to-male ratio of 2.5. All samples were collected between 2002 and 2006.

This study was approved by the ethic committee of Karolinska Institutet and is in compliance with national legislation and the Code of Ethical Principles for Medical Research Involving Human Subjects of the World Medical Association (Declaration of Helsinki). Informed consent was given by each participant.

Cell collection and RNA extraction

From each individual both the total cell population from cerebrospinal fluid (CSF), which contains 80% T lymphocytes, and peripheral blood mononuclear cell (PBMC) samples were obtained at the same scheduled visit at the Karolinska University Hospitals Neurology clinics. Peripheral blood was sampled in sodium citrate-containing cell preparation tubes (Vacutainer CPT; BD Biosciences, San Jose, CA) and CSF in siliconized glass tubes. CSF samples were immediately centrifuged, and the pellet was recovered and stored at -70°C until use. PBMCs were separated by density gradient centrifugation, and cells from the interphase were collected and washed twice with Dulbecco's PBS. Thereafter the cells were pelleted, frozen on dry ice, and stored at -70°C until use. Total RNA was purified according to manufacturer's instructions using PicoPure™ RNA isolation kit (Arcturus, USA) and incubated with DNase according to supplier's instructions (Qiagen RNase free DNase set, Hilden, Germany).

Microarray procedure

Each RNA sample was analyzed using a 2100 bioanalyzer (Agilent, USA) to assess RNA quality. In addition to the strict clinical inclusion criteria all samples from a single individual were obligated to have a 18s/28s ratio of at least 1.3 and a RIN ratio (Schroeder et al., 2006) of at least 7.7. Since cell numbers are sparse in CSF, we conducted an experiment using six low quantity samples ranging from 0.14 to 3.18 ng of RNA to examine the lowest amount of RNA able to produce sufficient amount of cRNA for hybridization in a two-cycle labeling/amplification protocol (Affymetrix).

The CSF samples of our study population had a RNA quantity of 0.8–10 ng, mean 7.1 ng. The quantity of total RNA from PBMC samples ranged from 1.9 to 20, mean 17 ng.

The resulting targets from the labeling were hybridized individually to 88 Human Genome 133 plus 2.0 arrays (Affymetrix, Santa Clara, CA) which examines the expression of over 47,000 known transcripts corresponding to over 39,000 genes.

Pre-processing

All pre-processing were done in the R environment (R Development Core Team, 2007) using packages from the Bioconductor project (Gentleman et al., 2004). All images of CEL files were investigated visually and two arrays with regional biases were excluded (one MS CSF, one control PBMC). The basic Affymetrix quality control measurements were assessed using the Simpleaffy package (Wilson and Miller, 2005) (average background, scale factor, percent present, 3′ to 5′ ratio, data not shown). Samples that showed deviant values in these initial quality assessments were also deviant in primitive clustering and component analysis (data not shown) and were therefore removed from further analysis (three samples: two PBMC MS relapse and one control CSF).

Summarization of probe set intensity, background correction and normalization was done using the Bioconductor implementation of the GCRMA algorithm (Wu et al., 2004), and conducted for each tissue separately. Prior to statistical analyses, genes whose expression levels did not vary in the sample population (inter quartile range<0.5) were removed.

Raw and pre-processed data are available at Array express (E-MTAB-69).

Exploratory analysis

Histograms were created from raw *p*-values in all four comparisons.

Principal component analysis (PCA) was performed on all arrays as well as the final set of 49 MS and 33 control arrays used in the analysis using the stats package in R.

Analysis of regulated transcripts

We assessed the association between transcript amounts and group specificity by *t*-tests with a null distribution created by 1,000,000 permutations as implemented in the Bioconductor package multtest (version 1.18.0). The false discovery rate (FDR) was estimated using the Benjamini and Hochberg method (Benjamini and Hochberg, 1995) as implemented in Bioconductor package multtest in order to avoid a high rate of falsely rejected null hypothesis, where we considered an FDR less or equal to 5% as acceptable. Additionally, probe sets with a fold change greater or equal to 2, or smaller or equal to 0.5 were regarded as significantly differently expressed in order to override possible technical biases affecting the data.