UNIVERSITY OF NOVA GORICA
GRADUATE SCHOOL

GENETICS BIOMARKERS AT MULTIPLE SCLEROSIS

DISSERTATION

Polona Lavtar

Prof. Borut Peterlin, MD, PhD

Nova Gorica, 2013

TABLE OF CONTENTS

1.INTRODUCTION 1

1.1. Clinical characteristics of multiple sclerosis (MS) 6

1.1.1. Multiple sclerosis diagnostic criteria 7

1.1.2. The course of MS 8

1.1.3. Expanded Disability Status Scale in multiple sclerosis 10

1.2. Biomarkers 11

1.3. Genetic biomarkers in multiple sclerosis 12

1.3.1 Family studies 13

1.3.2. Genetic association studies 13

1.3.2.1. Genome-wide association studies 14

1.3.2.2. Candidate gene approach 15

1.3.3. Other “omic” studies 16

1.3.3.1. Transcriptomic studies 16

1.3.3.2. Proteomic studies 17

1.3.3.3. Integration of global “omic” data 18

1.4. The signal transducers and activators of transcription genes (STAT genes, STAT3 and STAT5A) 19

1.5. Circadian rhythm genes – the ARNTL and CLOCK genes 22

2. EXPERIMENTAL 24

2.1. Subjects 24

2.2. Ethics statement 25

2.3. Integratomic approach 25

2.4. DNA and RNA extraction 27

2.4.1 DNA extraction 28

2.4.2 RNA extraction 28

2.5. Association studies 29
2.6. Genotyping 30
2.6.1. Genotyping of the STAT3 and STAT5A genes 30
2.6.2. Genotyping of the CLOCK and ARNTL genes 31
2.7. Gene expression 31
2.8. Statistical analyses 33

3.RESULTS AND DISCUSSION 33
3.1. Results of the STAT3 and STAT5A genes 35
3.1.1 Association studies in the STAT3 and STAT5A genes 35
3.1.2. Gene expression in the STAT3 and STAT5A genes 40
3.1.3. Correlation between SNPs and blood expression in the STAT3 and STAT5A genes 41
3.2. Discussion on the STAT3 and STAT5A genes 42
3.3. Results of the CLOCK and ARNTL genes 46
3.3.1. Association studies on the CLOCK and ARNTL genes 46
3.3.2. Gene expression in the CLOCK and ARNTL genes 54
3.3.3. Correlation between SNPs and blood expression in the CLOCK and ARNTL genes 55
3.4. Discussion on the CLOCK and ARNTL genes 55

4. CONCLUSIONS 59

5. SUMMARY 60
5. POVZETEK 62

6. REFFERENCES 64

LIST OF TABLES

Table 1: McDonald's MS diagnostic criteria 8

Table 2: Patients' disease-related characteristics for the studied group in association studies 34

Table 3: Patients' disease-related characteristics for the studied group in gene expression in blood 34

Table 4: Genotype and allelic distribution of the STAT3 and STAT5A polymorphisms in MS patients and healthy controls 35

Table 5: Genotype and allelic distribution of the STAT3 and STAT5A polymorphisms in MS patients and healthy controls in terms of gender 36

Table 6: Genotype and allelic distribution of the STAT3 and STAT5A polymorphisms in MS patients and healthy controls in terms of the course of RRMS and SPMS 37

Table 7: Genotype and allelic distribution of the STAT3 and STAT5A polymorphisms in PPMS patients and healthy controls 38

Table 8: Haplotypes in the STAT3 and STAT5A genes 39

Table 9: Genotype and allelic distribution of the ARNTL and CLOCK polymorphisms in MS patients and healthy controls 47

Table 10: Genotype and allelic distribution of the ARNTL and CLOCK polymorphisms in male MS patients and healthy male controls 48

Table 11: Genotype and allelic distribution of the ARNTL and CLOCK polymorphisms in the female population of MS patients and healthy female controls 49

Table 12: Genotype and allelic distribution of the ARNTL and CLOCK polymorphisms in RRMS patients and healthy controls 50

Table 13: Genotype and allelic distribution of the ARNTL and CLOCK polymorphisms in SPMS patients and healthy controls 51

Table 14: Genotype and allelic distribution of the ARNTL and CLOCK polymorphisms in PP MS patients and healthy controls 52

Table15: Haplotypes in the ARNTL and CLOCK genes 53

LIST OF FIGURES

Figure 1: Signal Transducer and Activator of Transcription pathway 20

Figure 2: Molecular components of the circadian clock network 23

LIST OF CHARTS

Chart 1: Differential expression of genes in MS in comparison to controls in the STAT3 and STAT5A genes in all participants and separated by gender 40

Chart 2: Differential expression of genes in MS in comparison to controls in the STAT3 and STAT5A genes in all participants according to the disease course....41

Chart 3: Correlation between SNP rs963987A/G and expression of STAT3 and Correlation between SNP rs6503695 C/T and expression of STAT5A....42

Chart 4: Differential gene expression of blood in the CLOCK and ARNTL genes in all participants and separated by gender....54

1. INTRODUCTION

Multiple sclerosis (MS) is a multifactorial chronic neuroinflammatory disease with unknown etiology and variable clinical evolution [1]. It affects twice as many women as men, usually with the age of onset ranging from 18 to 40. On a global scale, there are about two to three million MS patients who can mainly be found in Europe and in countries with Caucasian immigration, such as USA, Australia and Northern Asia [2, 3]. The prevalence of MS is influenced by family history, ethnicity, age and geography [4].

Neuroinflammation is a descriptive denominator in MS, based on histopathological observations. Clinical symptoms of relapses, remissions and progressive paralysis result from the losses of myelin and neurons [5, 6]. People who have been diagnosed with clinically definite MS have been through several diagnostic stages. Uncertainty over the cause, course and control of MS means that there is a constant high level of demand for information relating to the disease.

The complex and delicate nature of the disease poses a challenge to clinical diagnosis, especially in its early stages. We have to take in consideration that a large number of diseases can mimic the white matter changes occurring in MS [7].

The etiology of MS remains unknown to this date, as do the mechanisms associated with its unexpected onset, the unpredictable clinical course spanning decades and the different rates of progression leading to disability over time [8]. The available data suggest a complex multifactor etiology, including interactions of genetic and environmental factors [9].

While evidence indicates a genetic component in the pathogenesis of MS, a simple model of inheritance is unlikely [10]. A nonlinear decrease in disease risk in families with increasing genetic distance from the index case has been observed. Recurrence risk estimates in twin data and multiplex families predict that a genotype predisposing to the development of MS results from multiple independent and/or interacting polymorphic genes, each of which may exert a minor or moderate effect [10].

The risk of developing MS is higher for the relatives of a person with the disease than for the general population, especially in the case of siblings, parents and children [11]. The disease has an overall familial recurrence rate of 20%. With monozygotic twins, concordance occurs only in about 35% of the cases, while dropping to around 5% in the case of siblings and even lower in half-siblings [11]. This indicates susceptibility is partly polygenically driven [11].

To date, the best-established genetic association with susceptibility to MS is mapped to the chromosomal region 6p21.3 [12–16]. Within this region, the major histocompatibility complex (MHC), class II, DR beta 1."(HLA DRB1) locus has the strongest and most consistent association with MS in both northern European and North American populations [12–16]. Polymorphisms in HLA class II antigens seem to be decisive in attributing the genetic burden for MS. Initial studies found positive correlation between the DRB1∗1501- DRB5∗0101-DQA1∗0102-DQB1∗0602 haplotype and disease frequency [17–22]. Nevertheless, it is striking that, even for individuals who carry the HLA DRB1*1501 allele, only a very small fraction of them ever develops MS [23]. This observation seems to implicate the critical presence of other susceptibility alleles at different locations. Genetic susceptibility to MS seems to be determined by the involvement of multiple genetic loci scattered throughout the genome [24].

In general, genes showing variable expression include mainly immunological and inflammatory genes, as well as stress and antioxidant genes. In addition, metabolic and

central nervous system markers were investigated [25]. Due to the clinical complexity of the disease and large amounts of genetic data, it is difficult to interpret the available information. Several genome-wide association screens and “omic” approaches undertaken in recent times have substantially lengthened the list of genes implicated in MS. Although the etiology and the pathogenesis of MS have been extensively investigated, no single pathway, reliable biomarker, diagnostic test, or specific treatment has yet been identified for all MS patients. Uncertainty over the cause, course and control of MS means that there is a constant high level of demand for information relating to the disease and for the continuous search for new reliable genetic biomarkers. It is therefore of primary importance to distinguish reliable biomarkers for MS, especially in genomic regions and genes related to MS ethiopathogenesis.

In order to facilitate the discovery of biologically significant genetic alterations based on datasets from various types of studies in MS (linkage, genome-wide association, transcriptomic and proteomic studies data), we used integration-based search to find genomic regions and genes related to MS ethiopathogenesis. By integrating different genomic data, we have identified Top gene MS candidates. We decided to carry out tests for association and gene expression in blood, the signal transducers and activators of transcription genes (STAT genes), the STAT3 and STAT5A genes.

We also used a specific epidemiological characteristic of MS to generate an inovative hypothesis. There is namely a north-to-south gradient in the Northern Hemisphere and a south-to-north gradient in the Southern Hemisphere, with MS being much less common among people living near the Equator [11, 26]. Climate and sunlight have been investigated as possible causes of the disease [27]. The suggestion that sun exposure protects against MS by increasing vitamin D3 levels has become a common explanation for the latitude effect on the disease [27], but vitamin D is not the only factor affected by exposure to sunlight that has the capacity to modify the immune function. Nevertheless, the understanding of geoepidemiology of MS can represent a valuable source of environmental and genetic etiological clues [28, 29]. We hypothesize that the difference

in geoepidemiology of MS might be associated with chronobiology. Chronobiology examines periodic phenomena in living organisms and their adaptation to solar- and lunar-related rhythms, which include alternations of day and night (day length, light intensity), changes in the environment temperature and changes of seasons (photoperiod) [30]. External circadian rhythms arise as a direct result of environmental stimuli such as light, while a master clock is responsible for the interior circadian rhythms and is regulated by means of circadian rhythm genes. We have therefore decided to test two master key circadian rhythms genes, the Circadian Locomotor Output Cycles Kaput (CLOCK) and the Aryl hydrocarbon receptor nuclear translocator-like (ARNTL) gene.

We expect our findings to contribute to the better understanding of the mechanisms of this disease, as well as to improve the basic knowledge enabling the understanding of ethiopathogenesis of MS, and to discover novel potential genetic biomarkers for MS in blood.

In order to identify new genetic biomarkers for MS, we decided to test the following two hypotheses:

Hypothesis 1: The differences in gene variability in STAT genes (the STAT3 and STAT5A genes) might be associated with multiple sclerosis.

Hypothesis 2: The differences in epidemiology of MS related to the geographical latitude effect might be associated with chronobiology. Thus, gene variability in key regulators of circadian rhythm genes (the CLOCK and ARNTL genes) might be associated with multiple sclerosis.

To test our two hypotheses, we performed a retrospective cross-sectional case-control genetic association and gene expression studies on the STAT3 and STAT5A as well as the CLOCK and ARNTL genes on a population of patients with multiple sclerosis and healthy controls.

LITERATURE OVERVIEW

1.1. Clinical characteristics of multiple sclerosis (MS)

MS is characterized by the development of plaques within the central nervous system (CNS) [25]. Although some people with MS experience little disability during their lifetime, up to 60% are no longer fully ambulatory 20 years after the onset, with major implications for their quality of life and the financial cost to society [25]. The etiology of MS remains unknown, as do the mechanisms associated with its unexpected onset, the unpredictable clinical course spanning decades and the different rates of progression leading to disability over time [8].

Clinically, the complexity of this disease involves the location, size and duration of lesions which are variable and unpredictable [11, 31]. This lack of predictability can lead to a vast spectrum of symptoms which can vary among attacks or episodes of the disease throughout the progression [11, 31].

Symptoms of multiple sclerosis involve several disturbances of the visual, motor and sensory systems, as well as disruptions in coordination and balance, bowel/bladder/sexual functioning and cognition. Vision problems may include double vision, blurred vision and blindness in one or both eyes, eye pain and jerky eye movements. Motor problems may include partial or full paralysis, muscle weakness, stiffness, slurred speech and twitching muscles or tremors. Some individuals may experience sensory discomfort such as numbness (especially in the extremities), loss of awareness, facial pain, electric shock-like sensations, sensitivity to heat, and tightness around the torso or stomach. Ataxia, nausea, vertigo, stuttering and loss of the ability to produce rapidly alternating movements are symptoms of impaired coordination and balance. Bladder and/or bowel problems include the urgency to urinate/defecate, incontinence, retention and sexual impotence. Cognitive difficulties include depression,

short- or long-term memory loss, dementia, mood swings and anxiety. Other symptoms include fatigue, sleeping disorders and epileptic seizures [11, 32].

The complex and delicate nature of the disease poses a challenge to clinical diagnosis, especially in its early stages. We have to take in consideration that a large number of diseases can mimic the white matter changes occurring in MS. Pathological conditions that cause non-specific white matter lesions on magnetic resonance imaging (MRI) include acute disseminated encephalomyelitis (ADEM), neurosarcoidosis, leukodystrophies, systemic lupus erythematosus, primary anti-phospholipid antibody syndrome, migraine, cerebrovascular disease, Behçet's syndrome, ageing, decompression sickness, high altitude illness, fat embolism, HIV encephalitis, HTLV1-associated myelopathy, hydrocephalus, irradiation, mitochondrial encephalopathy, motor neurone disease, phenylketonuria, progressive multifocal leukoencephalopathy (PML) and subacute sclerosing panencephalitis [7].

1.1.1. Multiple sclerosis diagnostic criteria

There are different clinical manifestations of multiple sclerosis. In some ways, each person with multiple sclerosis lives with a different illness. Although nerve damage is always involved, the pattern is unique for each individual suffering from MS.

Physicians, particularly neurologists, take a detailed history and perform complete physical and neurological examinations. Collectively, three tests (MRI, electrophysiological test and cerebrospinal fluid exam (spinal tap, lumbar puncture)) are of help in confirming the diagnosis of multiple sclerosis. For definitive diagnosis, dissemination in time (at least two separate symptomatic events or changes on MRI over time) and in anatomical space (at least two separate locations within the central nervous system, which can be demonstrated by an MRI or a neurological exam) is usually required.

The most recent McDonald's criteria [33] incorporate MRI findings for determining whether lesions on the brain MRI are abnormal and whether these abnormalities satisfy the criteria for dissemination in time and space. The McDonald's criteria simplify the diagnostic outcome to being 'MS', 'possible MS' or 'not MS'. The McDonald's MS diagnostic criteria are presented in Table 1.

Table 1: McDonald's MS diagnostic criteria [33]

Clinical (attacks)	**Lesions**	**Additional criteria to make diagnosis**
2 or more	Objective clinical evidence of ≥2 lesions or objective clinical evidence of 1 lesion with reasonable historical evidence of a prior attack	None. Clinical evidence alone will suffice; additional evidence is desirable but it must be consistent with MS.
2 or more	Objective clinical evidence of 1 lesion	DIS; Or await further clinical attack implicating a different CNS site
1	Objective clinical evidence of ≥ 2 lesions	DIT; Or await second clinical attack
1	Objective clinical evidence of 1 lesion	DIS; Or await further clinical attack implicating a different CNS site and DIT; Or await second clinical attack
0 (progression from onset)		One year of disease progression (retrospective or prospective) and at least two of: DIS in the brain based on ≥1 T2 lesion in periventricular, juxtacortical or infratentorial regions; DIS in the spinal cord based on ≥2 T2 lesions; or positive CSF

DIT = lesion dissemination in time, DIS = lesion dissemination in space, CNS = central nervous system

1.1.2. The course of MS

Doctors and researchers have identified several major types of MS. The categories are important because they help predict disease severity and response to treatment. The most common type is relapsing-remitting multiple sclerosis (RRMS), which occurs in

approximately 55% of the cases [34]. Clinically, it is characterized by relapses or attacks followed by either partial or total recovery of symptoms. Biologically, it is characterized as focal areas of inflammation and demyelination, which resolve themselves over time leading to recovery. Thus, the damage caused by inflammation is at least partially reversible. Secondary progressive MS (SPMS) is the second most common type of MS accounting for about 30% of the cases. SPMS is characterized by initial relapses of RRMS, which over the course of the disease are replaced by progressive disability. Individuals with SPMS begin with reversible disability but, for unknown reasons, axonal degeneration occurs, leading to irreversible damage, which presents itself clinically as progressive disability [34]. Primary progressive MS (PPMS) occurs in approximately 10% of MS cases. It is characterized by progressive disability with no remitting stages. Thus, individuals with this clinical sub-type are plagued with irreversible damage that causes a slow or gradual progression to increased disability with little to no symptomatic relief [34]. Progressive relapsing MS (PRMS) occurs in nearly 5% of the cases. This clinical sub-type is characterized by progressive disability from the onset of symptoms, similar to PPMS, but also involves acute attacks or relapses. The reasons for this neurological variability remain unknown [34].

It is clear that the course of the disease in multiple sclerosis is characterised by a wide range of progression rates. A clinically isolated syndrome (CIS) is an individual's first neurological episode caused by inflammation or demyelination of the nerve tissue. An episode may be monofocal, in which symptoms occur at a single site in the central nervous system, or multifocal, in which symptoms are manifested at multiple sites [35]. In 85% of young adults with multiple sclerosis, the onset occurs in the form of a subacute clinically isolated syndrome of the optic nerves, brainstem or spinal cord [35]. Methods of assessing the prognosis for patients who display a CIS have been sought, as only 30–70% of the patients with a CIS develop MS [35]. When clinically silent brain lesions are seen on MRI, the likelihood of developing MS is high. MS can be diagnosed within 3 months of CIS presentation with certain MRI and CSF criteria. Disability from MS is less likely in patients with a CIS of optic neuritis or sensory symptoms only, a few or no MRI lesions, and those who have experienced a long period leading up to the first

relapse and no disability after the first 5 years [36]. The mildest form of clinically apparent multiple sclerosis is benign multiple sclerosis [37]. In an international survey on defining the clinical course of multiple sclerosis, there was disagreement concerning the criteria defining the benign disease, the consensus being that patients should be "fully functional" more than 15 years from the onset [37].

Much of the biological process is clinically silent. Moreover, despite intuitive reasoning, the lesion burden does not necessarily correlate with the scope or intensity of the disability [32, 38 and 39].

1.1.3. Expanded Disability Status Scale in multiple sclerosis

The Expanded Disability Status Scale (EDSS) is a rating system that is frequently used to classify and standardize the condition of people with multiple sclerosis.

The EDSS score is based upon neurological testing and examination of functional systems (FS), which are areas of the central nervous system that control body functions. The functional systems are [40] (i) pyramidal (ability to walk), (ii) cerebellar (coordination), (iii) brain stem (speech and swallowing), (iv) sensory (touch and pain), (v) bowel and bladder functions, (vi) visual, (vii) mental and (viii) other (include any other neurological findings due to MS).

The EDSS scale [40] is graded from 1.0 to 10.0. The EDSS steps 1.0 to 4.5 refer to patients who are fully ambulatory, and the precise step number is defined by the Functional System score(s). The EDSS steps 5.0 to 9.5 are defined by impairment to ambulation, and usual equivalents in Functional Systems scores are provided. EDSS should not change by 1.0 step unless there is a change in the same direction of at least one step in at least one FS.

1.2. Biomarkers

Biomarkers are indicators of normal and abnormal biological processes in the human body. It is an objectively measured and evaluated characteristic [41]. For a biomarker to be clinically useful, non-invasive detection is desirable. In neurological diseases affecting the central nervous system (CNS), the availability of the affected tissue is limited. Cerebrospinal fluid (CSF) collection is an invasive, although a relatively benign, procedure and can therefore be used only in selected patient population. It is quite difficult to obtain CSF from healthy controls, especially for research purposes. Studies in the brain are mostly limited to post-mortem tissue because biopsies for analytical purposes are only rarely taken. All studies that include brain tissue are based on a small number of cases, most of which were in a progressive disease stage. This is in contrast to studies using peripheral blood, where large numbers of patients at different disease stages can be sampled [42].

Peripheral blood is easily accessible tissue allowing relatively non-invasive screening for a host of medical conditions, and gene expression patterns have been shown to exist in human blood in a wide variety of central nervous system diseases where no obvious clinical phenotype in blood is present [43–47]. Correlation between gene expression in blood and in CNS tissues shares significant gene expression similarities with multiple CNS tissues [48].

During the last decades it has been proven very difficult to establish satisfactory biomarkers for multiple sclerosis due to the clinical and pathophysiological complexities of the disease. Moreover, different mechanisms of inflammation – demyelination, axonal damage – neurodegeneration, gliosis, and remyelintion – repair combine together in various degrees to create a unique clinical result for each patient [17].

An ideal biomarker for MS should represent the use of certain criteria [17]:

- Biological rationale: degree of the correlation of the biomarker with an identified pathogenic mechanism
- Clinical rationale: accuracy in depicting the clinical status
- Predictability of disease initiation, reactivation or progression
- Sensitivity and specificity
- Reproducibility of results
- Practicality of the method in use for the measurement
- Correlation with therapeutic outcome
- Correlation with prognosis and disability

The etiology of MS remains unknown to this date, as do the mechanisms associated with its unexpected onset, the unpredictable clinical course spanning decades and the different rates of progression leading to disability over time [8]. Reliable and clinically useful diagnostic and prognostic markers for MS are still not available. In diseases with a complex pathogenesis, such as multiple sclerosis, an individual biomarker is likely to reflect only one of the many on-going pathogenic processes. Consequently, there is still a substantial gap between the estimated heritability and evidence of contribution of already identified genetic biomarkers. Further research information is essential for the understanding of the pathogenesis of MS. It is therefore of primary importance to distinguish reliable biomarkers for MS, especially in genomic regions and genes related to MS ethiopathogenesis.

1.3. Genetic biomarkers in multiple sclerosis

Complex diseases, as opposed to Mendelian diseases, do not show a clear inheritance pattern in families [4]. The genetic component in MS is suggested by familial aggregation of cases and by the high incidence in some ethnic populations compared

with other populations [4]. A genetic etiology is indicated the most by the recurrence risk in family members of the affected individual.

1.3.1 Family studies

The first familial clustering of multiple sclerosis documented by neurologists has revealed that ~20% of patients have one or more affected relatives [49]. Building on this observation, the last decade has seen definitive studies of recurrence risks in informative groups of family members. These studies represent an enormous amount of field work, and, although some have incorporated flaws in design or analysis, their results can be usefully summarized as follows: compared with a lifetime risk for northern European Caucasians (~1:300), there is an increased relative risk (λ) for identical twins (λ=100–190 [50,51]), full siblings (λ=13 [52]), half siblings (λ=7 [53]), other first-degree (λ=7 [52]) and second-degree (λ=3.5 [53]) relatives and the children of single affected parents (λ=5.5 [52]), and in conjugal parents (λ=60 [54]), but not adoptees (λ=1 [55]). Further evidence for a genetic contribution to causation is provided by a similar age at the onset within multiplex families [56, 57]. These statistics suggest that genetic factors play a major role in multiple sclerosis.

Prior to the introduction of genome-wide association studies (GWAS), the primary method of investigation was through inheritance studies of genetic linkage in families. This approach has proven highly useful with regard to single gene disorders [58], for complex diseases, however, the results of genetic linkage studies proved hard to be reproduced [59, 60]. A suggested alternative to linkage studies was the genetic association study.

1.3.2. Genetic association studies

Genetic association study is a study type that poses the question of whether the allele of a genetic variant is found more often than expected in individuals with the disease being

studied, and it might represent or reflect a piece of the mosaic in susceptibility factors [61].

Two types of genetic association studies are performed at present:

- hypothesis-free studies (GWAS studies) and
- candidate gene approaches.

1.3.2.1. Genome-wide association studies

A genome-wide association study (a GWA study or GWAS), also known as a whole genome association study (a WGA study or WGAS), is an examination of many common genetic variants in different individuals to see if any variant is associated with a trait. GWAS using single nucleotide polymorphism (SNP) maps have revolutionized the mapping of common genetic loci determining susceptibility to a wide range of common, multifactorial disorders [24], particularly in autoimmune diseases [62].

A major challenge of the genome-scale transcription analysis is the very large number of genes compared to a generally small number of measurements [63]. Without appropriate statistical measures used to make corrections in multiple testing, almost any approach will yield significant genes, including many false positives [63]. The creation of large databases in recent years has brought an additional layer of complexity and precautions to take.

To date, 12 GWAS and 16 other large-scale studies have been conducted [64, 65]. So far, GWAS have enabled 57 additional risk loci associated with the disease [66]. Immunological and inflammatory genes, stress and antioxidant genes, as well as metabolic and central nervous system markers were mainly investigated as MS associated genes. 80% of so far identified genes associated with the disease are primarily expressed in immune cells [67]. Polymorphisms of the IL2RA and IL7RA regions seem

to be the most promising at the moment [68]. Increasing evidence also indicates other loci, such as the EVI5, CD58, KIAA0350 and RPL5 genes [69, 70 and 71].

Identification of biomarkers appears desirable for an improved diagnosis of MS as well as for monitoring the disease activity and treatment response. It is of primary importance to distinguish reliable biomarkers for MS, especially in genomic regions and genes related to MS ethiopathogenesis.

1.3.2.2. Candidate gene approach

To date, the best-established genetic association with susceptibility to MS has been mapped to the chromosomal region 6p21.3 [12–16]. Within this region, the major histocompatibility complex (MHC), class II, DR beta 1."(HLA DRB1) locus (and especially the haplotype that includes the DRB1*1501 allele) has the strongest and most consistent association with MS in both northern European and North American populations [12–16].

Polymorphisms in HLA class II antigens seem to be decisive in attributing the genetic burden for MS. Initial studies found a positive correlation between the DRB1∗1501-DRB5∗0101-DQA1∗0102-DQB1∗0602 haplotype and the disease frequency. Multiple recent researches conducted in many MS cohorts made it clear that HLA-DRB1∗1501 is the allele that contributes the most to genetic risk in the MS population [17–22].

Nevertheless, it is striking that, even among individuals who carry the HLA DRB1*1501 allele, only a very small fraction of them ever develops MS [50]. This observation seems to indicate the critical presence of other susceptibility alleles at different locations. Genetic susceptibility to MS seems to be determined by the involvement of multiple genetic loci scattered throughout the genome [24].

The next steps to follow up on these findings are the identification of particular candidate variants and haplotypes, and the investigation of the molecular effects of these genetic variants. To investigate the potential effects of candidate causal variants and haplotypes on gene regulation, researchers have been correlating SNPs with inherited gene expression known as expression quantitative trait loci (eQTLs). The combination of genome-wide genotyping with quantification of RNA transcripts using microarray technology in sufficiently large cohorts has already demonstrated the widespread presence of eQTLs in the human genome [72–75].

1.3.3. Other "omic" studies

The term "omics" refers to a group of emerging novel technologies. These technologies enabled large-scale analyses and identification of candidate biomarkers on multiple levels of cell biology [17]. The following are the most commonly utilized "omics" technologies [17]:

(i) Transcriptomics: Genome-wide studies of RNA sequences. Two main types of transcriptomics technologies in common use are microarrays and next-generation sequencing.

(ii) Proteomics: Large-scale studies of protein distribution

1.3.3.1. Transcriptomic studies

A transcriptome is a set of all RNA molecules which are commonly defined as the complete set of transcripts in a cell for a specific developmental stage or physiological condition [76]. It is important to understand that the transcriptome is essential for interpreting the functional elements of the genome and that it reveals the molecular constituents of cells and tissues. The key aims of transcriptomics are to determine the

transcriptional structure of genes and to quantify the changing expression levels of each transcript during development and under different conditions [76].

Microarray technologies have provided hundreds of valuable datasets in a wide variety of diseases, including multiple sclerosis [63]. This approach has been used to disentangle different aspects of its complex pathogenesis. RNA sequencing (RNA-seq), the current next-generation sequencing approach, is expected to provide similar power as microarrays.

In addition to messenger RNA, miRNAs play an important role in regulating physiological and homeostatic processes. In the human immune system, miRNA molecules modulate innate immune responses. These inflammatory responses of immune cells against invading pathogens must be efficiently and closely regulated [77]. MicroRNAs regulate gene expression in health and disease, during development, immune system activation, neurogenesis, and myelin formation in the CNS. In MS, miRNA profiles are altered within CNS lesions and in the immune system, and affect gene expression in many cell types involved in the disease [77]. A global consideration of miRNA dysregulation and the resulting alterations in gene expression may provide valuable insights into the pathophysiology of MS and reveal new alternatives for early diagnosis and treatment [77].

1.3.3.2. Proteomic studies

Proteomics is an evolving technology platform that has the potential to identify novel proteins involved in key biological processes in the cell that may serve as potential drug targets [78].

Proteomics-based analyses are now starting to make progress in building databases to be utilized in studies on the normal brain and spinal cord, and on neurological and psychiatric disorders [79–81]. Several performed proteomic research approaches play an

important role in identification and quantitation of proteins which undergo disease-related changes, and also in protein expression studies on normal adults. Proteomic alterations in MS were investigated in various tissues such as the human brain, brain lesions, cerebrospinal fluid and the blood serum [79–81].

1.3.3.3. Integration of global "omic" data

Results obtained from the "omic" studies in MS have still not clarified the disease etiology and pathophysiology to a satisfactory degree, and have provided a multiplicity of results that are complex to interpret [82]. Multiple attempts at integration of such heterogeneous data have already been reported; approaches to such prioritization have been based either on candidate genes, functional similarity to a set of predefined training genes with known biological role in a disease under investigation [83], or on experimental data from studies examining specific alterations in the disease state [84, 85].

Methodologies utilizing integration approaches predominantly investigated which genes were overlapping in results from different types of included studies, and thus utilized the gene-centric approach toward integration [83, 85 and 86]. Considerable fractions of meaningful genetic alterations are located outside of genes, between adjustment genes, or spanning over several neighbouring genes: single nucleotide polymorphisms, copy number variations, regions defined by linkage disequilibrium studies as well as epigenetic alterations and others [87]. As it is difficult to map these genetic alterations to appropriate genes unambiguously, a choice of an efficient common denominator for the integration process might be problematic [87]. There might be difficulties due to incompatible reporting of results in publications from various studies. Consequently, this represents notable problems not only when analysing data from multiple study types, but also when using data from a single type of study [87]. Although tools and relation databases enabling conversion of these identifiers exist, this process is often incomplete and results in significant loss of data [87].

Moreover, the etiology and the pathogenesis of MS have been extensively investigated, but no single pathway, reliable biomarker, diagnostic test or specific treatment has yet been identified for all MS patients. Consequently, there is still a substantial gap (missing heritability) between the estimated heritability and evidence of contribution of the already identified genetic biomarkers. Uncertainty over the cause, course and control of MS means that there is a constant high level of demand for information relating to the disease and for the continuous search for new reliable genetic biomarkers. Identification of reliable biomarkers for MS bears the potential for an improved diagnosis of MS, monitoring of the disease activity and progression, and evaluation of treatment responses.

The selection of two candidate genes was based on our integratomic study (linkage, GWAS, transcriptomic and proteomic studies data), where we characterised significant genomic regions associated with MS (details about region selection are described in the section entitled *Experimental*). We therefore decided to conduct tests for association and gene expression in blood, the signal transducers and activators of transcription genes (STAT genes), especially the STAT3 and STAT5A genes as possible genetic biomarkers.

1.4. The signal transducers and activators of transcription genes (STAT genes, STAT3 and STAT5A)

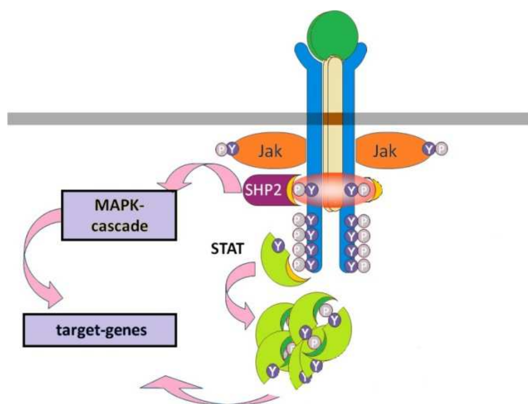
The human STAT gene family consists of seven members (STAT1, STAT2, STAT3, STAT4, STAT5A and STAT5B, STAT6), many of which play highly specific roles in innate and acquired immunity, that are localized on three chromosomal regions (chromosome 2, 12 and 17) [88].

Signal transduction pathways are of essential importance for organisms, enabling them to react to environmental changes [82–85]. Activation of STATs results in expression of genes that control critical cellular functions including cell proliferation, survival,

differentiation and development, as well as specialized cellular functions such as those associated with immune responses [89–92]. Although the STAT family is highly preserved in terms of structure, there are distinct differences both in the primary sequence and function [89–92]. The STATs genes are involved in multiple pathways and functions, including the JAK/STAT pathway, neuron axonal guidance, apoptosis, activation of immune responses, and Th17 cell differentiation [93]. Although the mechanism of JAK/STAT signaling is relatively simple in theory, the biological consequences of pathway activation are complicated by interactions with other signaling pathways [94–96]. The understanding of this cross-talk is only beginning to emerge, but the best characterized interactions of the JAK/STAT pathway are those occurring with the receptor tyrosine kinase (RTK)/Ras/MAPK (mitogen-activated protein kinase) pathway as presented in Figure 1.

In normal cells, these signal transduction mechanisms are closely controlled to prevent unscheduled gene regulation and, consequently, inappropriate biological responses. Uncontrolled STAT activation can result in a variety of immune-mediated disease states such as multiple sclerosis, rheumatoid arthritis disorders of the bone marrow and several types of tumours [78, 97 and 98].

Figure 1*: *Signal Transducer and Activator of Transcription pathway



Modified by [99]

The STAT3 protein is active in tissues throughout the body. It plays an important role in the development and function of several body systems and is essential for life. In the immune system, the STAT3 protein transmits signals that help control the body's response to foreign invaders such as bacteria and fungi. In particular, the protein is involved in the regulation of inflammation, which is one way the immune system responds to infections or injuries. A recent meta-analysis of GWAS in MS listed STAT3 as one of the genes with a suggestive role in inflammatory bowel disease and multiple sclerosis [16].

The STAT5 gene (STAT5A) was first identified as a prolactin-induced transcription factor. Biochemical studies aiming to identify an Interleukin-3 (IL-3)-induced transcription factor and cDNA screens led to a discovery of two closely related sequences of STAT5A and STAT5B [100, 101]. These proteins were found to be encoded by two linked genes, STAT5A and STAT5B, which shared 96% of identity and only diverged at their carboxy termin [102]. STAT5 promotes cell survival and/or proliferation by regulating the expression of genes involved in the control of the cell cycle [103]. Besides the physiological role of STAT5 in hematopoietic cell development, there is increasing evidence suggesting that inappropriate activation of STAT5 may contribute to the development of leukaemias and solid cancers [104, 105].

Activation of STAT5A is often an integral component of redundant cytokine signal cascades involving complex cross-talk and pleiotropic gene regulation. STAT5 has been implicated in cellular functions of proliferation, differentiation and apoptosis with relevance to the processes of hematopoiesis and immunoregulation, reproduction and lipid metabolism [106]. Moreover, diseases like asthma [107], lupus [108], B-cell leukaemia [109], melanoma [110] and breast cancer [111] have been linked to STAT5A.

1.5. Circadian rhythm genes – the ARNTL and CLOCK genes

There are some MS characteristics that are far from adequately explained. A direct correlation exists between latitude and frequency, i.e. the disease is most frequent in northern climes [11, 112]. According to our hypothesis, the difference in geoepidemiology of MS might be associated with chronobiology.

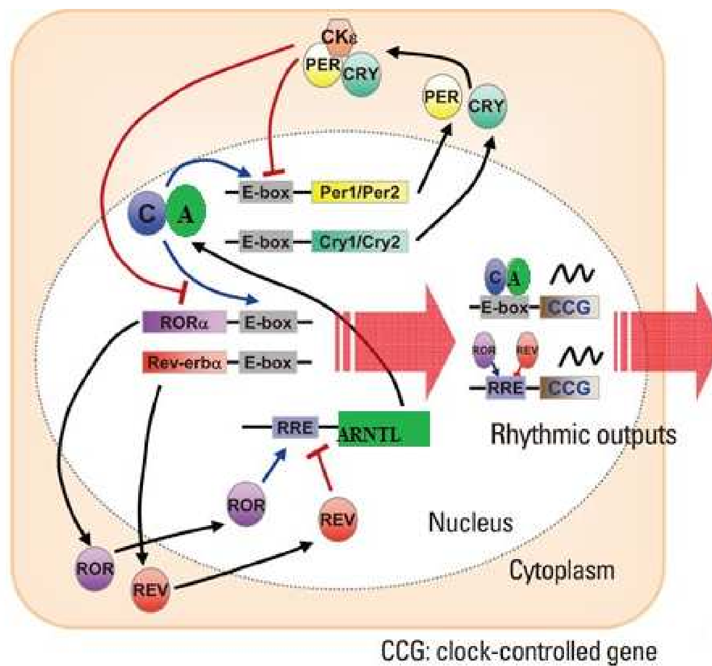
Chronobiology examines periodic (cyclic) phenomena in living organisms and their adaptation to solar- and lunar-related rhythms. These adaptations include alternations of day and night, changes in the environment temperature and changes of the seasons. External circadian rhythms arise as a direct result of environmental stimuli such as light, while a master clock is responsible for the interior circadian rhythms and is regulated by means of circadian rhythm genes. Beyond the potential for underlying genetic predisposition, however, it is possible that geographic differences in the amount of sunlight may prove to be potential risk factors [113].

Although the CNS cells have a central role as timekeepers in mammals, they are not the only cells in the mammalian body that have an internal circadian rhythm or an ability to reset it in response to light [114]. The proper working of circadian clocks represents a fundamental problem in cell biology. Although we do not yet understand all the details, studies in a wide variety of organisms have revealed many of the basic principles and molecular components [114].

Molecular components of the circadian clock network represent positive and negative transcriptional-translational feedback loops of many genes. The core circadian clock genes that have been identified in mammals are Circadian Locomotor Output Cycles Kaput (CLOCK), Aryl hydrocarbon receptor nuclear translocator-like (ARNTL), Cryptohromes (CRY), Period (PER) and Rev-erb α genes. CLOCK and ARNTL are basic helix-loop-helix transcriptional activators that drive the expression of PER and CRY genes as presented in Figure 2 [115]. PER and CRY inhibit their own CLOCK and ARNTL induced transcription, whereas the turnover of PER and CRY allows this cycle

to restart (see Figure 2) [116, 117]. The stability and precision of circadian rhythms is further enhanced by the interplay between different modulators of the core clock genes [118].

Figure 2: Molecular components of the circadian clock network



Modified by [119]

The circadian system represents a complex multioscillatory temporal network [120]. The circadian clock influences hormones, behaviour, cognitive function, metabolism, cell proliferation, apoptosis and responses to genotoxic stress [121]. In an intact body, the CNS cells receive neural cues from the retina, entraining them to the daily cycle of darkness and light, and they send information about the time of day to another brain area, the pineal gland, which relays the time signal to the rest of the body by releasing the hormone melatonin in time with the clock [114].

Absolute sensitivity of the suprachiasmatic nucleus (SCN) to photic stimulation received through the retino-hypothalamic tract changes along the 24-hour day [122, 123]. Pevet et

al. demonstrated that the duration of the SCN phase sensitivity to light is closely related to the length of the night [124]. However, circadian disruption occurs when endogenous rhythms are out of phase with the external environment.

Circadian disruption as consequences of shift work are associated with disturbed melatonin secretion and enhanced pro-inflammatory responses, and may play a role in the onset of MS [121]. Furthermore, disturbed melatonin secretion due to circadian disruption may lead to dysregulation of the immune system [121]. Desynchronization of circadian rhythm genes has been linked to various neurodegenerative disorders, metabolic disorders, insomnia, depression, coronary heart diseases and cancer [117, 125 and 126].

2. EXPERIMENTAL

To test our two hypotheses, we performed a retrospective cross-sectional case-control genetic association and gene expression study on the STAT3 and STAT5A and CLOCK and ARNTL genes on a population of patients with multiple sclerosis in comparison with healthy control population.

2.1. Subjects:

The study was performed on Caucasian population from the Central South East Europe. Slovenian subjects diagnosed with multiple sclerosis were recruited with collaboration of the Specialist Outpatient Clinics functioning within the University Medical Centre Ljubljana, the Department of Neurology at the University Medical Centre Ljubljana, and the Department of Family Medicine Rehabilitation for patients with multiple sclerosis in Topolšica, Slovenia. The patients' blood samples from other countries were obtained from the Central South East European Genomic Consortium for Multiple Sclerosis (the

CSEE-GCMS consortium). The research included patients with definitive MS disease who fulfilled the McDonald's criteria for MS [33]. The patients filled out a structured questionnaire about family history and risk factors associated with MS. One thousand and fifty-eight white patients from the Central South East Europe (Slovenia, Croatia and Serbia) with relapse-remitting multiple sclerosis (RRMS), secondary progressive multiple sclerosis (SPMS), primary progressive multiple sclerosis (PPMS), clinically isolated syndrome (CIS) and patients with the benign course of MS were recruited in this study. Patients were excluded from the study if they had a history of any kind of previous neurological symptoms or signs, i.e. clinical, laboratory, MRI, or cerebrospinal fluid findings suggestive of any diagnosis other than multiple sclerosis. The control group consisted of healthy individuals matched in terms of ethnicity, age and gender, with no family history of MS or any other inflammatory demyelinating disease. The control group comprised one thousand and forty individuals. Blood samples of healthy controls from Slovenia were obtained from healthy blood donors and under the research project entitled Family History – A Genetic Tool for Preventive Medicine (J3-3628 (C)). Some of the blood from controls from other countries was obtained through the CSEE-GCMS consortium. For gene expression in blood, we analysed 50 patients and 40 controls. None of the studied patients was undergoing immunomodulatory therapy.

2.2. Ethics statement:

The study was approved by the National Medical Ethics Committee (90/08/12). All participants gave informed written consent to participate in the study.

2.3. Integratomic approach:

We approached the integration of datasets originating from global molecular studies in MS using a position-based integrative approach as described previously by our research

group [87]. The study was commenced by searching, reviewing and selecting the studies reporting results of global molecular profiling in MS, with the aim of inclusion in the dataset for integration. Initially, a search of Medline (www.ncbi.nlm.nih.gov/pubmed/) was performed using the search string: ["Multiple sclerosis" AND (transcriptom* OR proteom* OR "genome-wide" OR "linkage scans" OR microarray OR profiling)]. Afterwards, a search in microarray data repositories was conducted: Gene Expression Omnibus (GEO) repository (http://www.ncbi.nlm.nih.gov/geo/), ArrayExpress (http://www.ebi.ac.uk/arrayexpress/) and Stanford Microarray Database (http://smd.stanford.edu). Datasets resulting from genome-wide association studies were obtained from the dbGAP repository (http://www.ncbi.nlm.nih.gov/gap). Standardized functional characterization of genes was obtained from Kyoto Encyclopedia of Genes and Genomes (KEGG) (http://www.genome.jp/kegg/) and GeneOntology databases (http://www.geneontology.org/), while the coexpression dataset was downloaded from the CoXpresDB project site (http://coxpresdb.jp/), and genotype-phenotype associations were obtained from Online Mendelian Inheritance in Man (OMIM). GWAS revealed many single nucleotide polymorphisms as candidates for genetic burden augmentation in MS. Most of them, however, had only a modest effect on susceptibility in Human Phenotype Ontology Project databases (http://www.ncbi.nlm.nih.gov/omim and http://www.human-phenotype-ontology.org/) and data on MS mouse models from the Mouse Genome Informatics (MGI) site (http://www.informatics.jax.org/). Ultimately, data obtained from 39 studies or data sources and comprising 16 study categories were included in the final data set prior to integration.

In total, 158,520 distinct and significant results discovered on 16 different biological levels were included in the dataset prior to statistical analyses. A custom rank product prioritization approach based on genomic position of included results was utilized for data synthesis, and nested permutation cycling was then employed to determine significant accumulation of most significant results discovered on the most diverse biological levels. Results from different studies were synthesized for each biological layer separately. Human genome assembly was subdivided into 10 kb regions and with a 5000 bp overlap with the neighbouring region (616,108 regions). This overlap was

introduced in order to prevent losing information of signal aggregation at the borders of defined regions. Significant signals from the above-referenced studies were arranged into their corresponding regions according to their genomic position. For each region, rank product values were calculated and significance values were estimated by permuting the original dataset. In total, 381 genomic regions were characterized by a significant accumulation of results, reaching local permutation P-value below 0.001.

Results originating from several different studies conducted on various molecular biological layers have supported the selection of STAT3 and STAT5A as notable candidates for multiple sclerosis. Firstly, data from three different GWAS data sources suggested several SNPs in STAT3-STAT5A genomic region to be associated with MS [16, 127 and 128]. Furthermore, a differential expression of the STAT3 gene was detected in blood samples in a study conducted by [129], and the gene was also singled out in the meta-analysis performed across three different genome-wide expression profiling studies. Additionally, STAT3 [130, 131] and STAT5A [130] represent two of the targets of miRNA molecules that have been found differentially expressed in the blood samples of MS patients. The gene was found to be indirectly related to the genes strongly associated with MS in the OMIM database, either through protein interaction reference information, KEGG pathways or phenotypic similarity in the Human Phenotype ontology datasets. Finally, STAT3 was found to be interacting with proteins that are related to EAE in mouse knock-out models [132].

2.4. DNA and RNA extraction

Blood collections were taken within a 4-hour window (9 am to 1 pm) to minimize the effect of diurnal variation (gene expression studies) [133]; the patients and controls were not using immunomodulatory drugs.

2.4.1 DNA extraction

For extraction of genomic DNA from whole blood samples collected in 3ml tubes with EDTA, we used the FUJIFILM QuickGene DNA whole blood extraction kit and Nucleic Acid Isolation System QuickGene-610L. All blood samples were frozen before extraction. The DNA extraction protocol was carried out in the following way: In an empty 50ml centrifuge tube we added 300μl of protease (dried protease was dissolved with ultra-pure water following the manufacturer's instructions), 2ml of whole blood and 2.5ml of lysis buffer. We mixed this mixture thoroughly by shaking the tube up and down 10 times, which was followed by thorough mixing by means of vortexing at the maximum speed of 2500 rpm for 15 seconds. We incubated the tube by means of a water bath at 56°C for 5 minutes. Next, we added 2.5ml of 99% ethanol. We then mixed the mixture thoroughly by shaking it up and down 10 times, which was followed by thorough mixing by means of vortexing at the maximum speed of 2500 rpm for 15 seconds. Afterwards, we transferred the whole lysate into the cartridge of QuickGene 610L. Genomic DNA was eluted with automated FUJIFILM Nucleic Acid Isolation System QuickGene-610L. The eluted volume of genomic DNA amounted to 500μl. Genomic DNA was then stored at -24°C until the time for genotyping.

2.4.2 RNA extraction

For extraction of total RNA from venous whole blood collected in Vacuette Tempus blood RNA tubes (9ml), we used the FUJIFILM QuickGene RNA blood cell kit and Nucleic Acid Isolation System QuickGene-810. All blood samples were frozen before extraction. The RNA extraction protocol was carried out in the following way: Frozen venous whole blood collected in tubes was melted for 2 hours until the RNA isolation (extraction) procedure was started. In an empty 50ml centrifuge tube we added all 9ml of whole blood from the tempus tubes (6ml of stabilizer and 3ml of venous whole blood) and 3ml of PBS (1x). We mixed the mixture thoroughly by vortexing it at the maximum speed of 2500 rpm for 30 seconds, after which we centrifuged the tubes at 4000 rpm for

30 minutes at 4°C. We eliminated pellets by flicking the tube. We added 520μl of the mixture of 2-mercaptoethanol (2-ME) and lysis buffer (LRB) per sample (10μl of 2-ME per 1μl of 1ml of LRB). The mixture was mixed gently by means of vortexing. The samples were kept on ice during the entire RNA extraction protocol. We transferred the samples into 1.5ml bio pure tubes and added 250μl of 99% ethanol, followed by vortexing for 5 minutes at 1700 rpm. We then transferred the samples into the QuickGene 810 cartridge. RNA was eluted by means of automated FUJIFILM Nucleic Acid Isolation System QuickGene-810. The eluted volume of the extracted RNA amounted to 50μl. Samples of RNA were then stored at -80°C until the time for gene expression.

2.5. Association studies:

Genetic variants in STAT genes (STAT3 and STAT5A) and in both key circadian rhythm-regulating genes (CLOCK and ARNTL) were genotyped in this study. Five tagging SNPs were chosen from STAT genes. Of these, three SNPs, rs7211777, rs963987 and rs1053004, were selected in STAT3 and two SNPs, rs6503695 and rs12601982, in the STAT5A gene. Eight tagging single nucleotide polymorphisms (SNPs) of circadian rhythm-regulating genes were chosen from both genes. Of these, four SNPs, rs11932595, rs6811520, rs6850524, and rs13124436, were selected in the CLOCK gene, and four SNPs, rs3789327, rs1481892, rs4757144 and rs12363415, in the ARNTL gene. We based the selection of SNPs on the known genetic linkage in genes according to HapMap Phase 3 (http://www.hapmap.org). The SNPs were regarded as proxies for the neighbouring SNPs when their pairwise r^2 values exceeded 0.80. The set of the most representative tagSNPs for gene regions was obtained using the Tagger algorithm [134] available through the Haploview software (Haploview, version 4.2). To maximize statistical power, only the SNPs with minor allele frequency values exceeding 0.05 were included in the investigated set.

2.6. Genotyping

Genotyping was performed by real-time PCR, using KASPar SNP Genotyping system (KBiosciences, UK). KASPar system employs a competitive allele-specific PCR combined with a FRET quenching reporter oligo. For each SNP genotyping reaction, custom made assays were ordered from KBiosciences, UK. For each assay three oligos are synthesized (two labeled allele specific primers and one common reverse primer) and after amplification using genomic DNA as a template, the fluorophore signals are measured and genotypes are determined. KASPar SNP genotyping system uses FAM and VIC fluorescence dyes for distinguishing between genotypes, and ROX as a passive reference dye. After completion of PCR cycling under conditions specific for each assay, the fluorescence levels were measured using ABI prism 7000 sequence detection system (ABI, Foster City, USA). Genotypes were scored by analyzing data from allele discrimination method using SDS software provided by the ABI (ABI, Foster City, USA). The total reaction volume for SNP in all genotyping reactions was 8µl as recommended by KBioscience. Pipetting of reactions in 96 well plates was performed by automated pipetting system (Precision XS Microplate Pipetting System, BioTek™ , VT, USA). All genotyping was performed at the same location at the Laboratory of the Institute of Medical Genetics, University Medical Centre Ljubljana, Slovenia, and was interpreted independently by three investigators.

2.6.1. Genotyping of the STAT3 and STAT5A genes

The PCR reaction mix for STAT genes consisted of 8μl final volume comprising 3μL of the DNA sample and 5μl of the mixture (4μl of reaction mix (2x), 0.11μl of assay mix, and 0.989μl of nuclease-free water (NFH_2O)). The protocol for PCR amplification in STAT genes consisted of initial denaturation at 94°C for 15 minutes, then 10 cycles of denaturation at 94°C for 10 seconds, annealing started at 61°C for 60 seconds and then touchdown with a drop of 0.6 ^{o}C/cycle and final extension of 36 cycles of denaturation at 94°C for 10 seconds, annealing at 55 ^{o}C for 60 seconds.

2.6.2. Genotyping of the CLOCK and ARNTL

The PCR reaction mix in circadian rhythm genes consisted of the 8μl final volume comprising 3μL of the DNA sample and 5μl of the mixture (4μl of reaction mix (2x), 0.11μl of assay mix, and 0.826 or 0.890μl of NFH_2O (SNP dependent) and 0.064μl of 50mM $MgCl_2$ at rs3789327, rs4757144 and rs12363414 in the ARNTL gene and rs11932595, and rs13124436 in the CLOCK gene). The protocol for amplification of the circadian rhythm genes was carried out in the following way: initial denaturation step at 94°C for 15 minutes, then 20 cycles of: 94°C for 10 seconds, followed by 57°C or 61°C for 5 seconds (SNP dependent), 72°C for 10 seconds and 20 cycles of: 94°C for 10 seconds, 57°C or 61°C for 20 seconds (SNP dependent), and final extension at 72°C for 40 seconds.

2.7. Gene expression

Three quality control steps were performed on isolated RNA samples. The first step determined the quantity, the second the purity and the third the integrity of the RNA that has been isolated. All these quality control steps were performed by means of Nanodrop (Thermo scientific NanoDrop 2000). Equal amounts of RNA samples passing quality checks were subjected to reverse transcription using the SuperScript® VILO™ reverse transcriptase (Invitrogen, Carlsbad, CA, USA) in line with the manufacturer's instructions. Input cDNA samples were diluted so that an amount of cDNA equivalent to 50ng of RNA was used in each reaction. The resulting cDNA was then amplified using specific gene candidate primers designed with PrimerBlast (www.ncbi.nlm.nih.gov/tools/primer-blast/). The threshold cycle (Ct) values were then determined for each primer pair. Expression levels were normalized according to the housekeeping gene β-actin, by using software provided by ABI (ABI, Foster City, USA) and co-run with each sample. The primer pairs from Invitrogen were for the CLOCK

gene (F primer: TCA AGA AAA TCA TCT CAC ACG and R primer: CTA AAT GAT GAC CTT CTT TGC), for the ARNTL gene (F primer: CAG AAC ACC AAG GAA GGA TAA AA and R primer: GAC ATT GCG TTG CAT GTT G), for the STAT3 gene (F primer: GGG AAG AAT CAC GCC TTC TAC and R primer: ATC TGC TGC TTC TCC GTC AC), for the STAT5A gene (F primer: GGT GTT GAA GAA GCA CCA CA and R primer: CAC TAA AGC GCA ACA AGA AGG), and for the housekeeping gene β-actin (F primer: CTG GAA CGG TGA AGG TGA CA and R primer: AAG GGA CTT CCT GTA ACA ATG CA). The protocol for PCR amplification was initial denaturation at 95°C for 15 minutes, then 40 cycles of denaturation at 95°C for 15 seconds, annealing at 57°C for 30 seconds and elongation at 72°C for 30 seconds, and final extension of 40 cycles of denaturation at 95°C for 15 seconds, annealing at 57°C for 20 seconds and elongation at 95°C for 15 seconds. To minimize stochastic effects, all reactions were run in triplicates. We pipetted reactions in 96 well plates using the Precision Microplate Pipetting System. Reactions (11μl) were performed using the Applied Biosystems Power SYBR® Green™ master mix for quantitative PCR (qPCR) on an ABI PRISM®7000 sequence detection system (ABI, Foster City, USA). The final volume of 11μl reaction consisted of 5.5μl of Power SYBR® Green™ master mix, 0.1μl of each primer (reverse and forward) and 0.3μl of NFH_2O. The Power SYBR® Green™ dye specifically binds to double-stranded DNA and emits fluorescence only upon binding. Since dye binding is not sequence-specific, each primer test was first tested for specificity and reaction efficacy by dissociation and melt curve. Only specific primer pairs were selected for further analysis. After the analysis, we verified whether the baseline and threshold were called correctly for each well by viewing the resulting amplification plots, and adjusted the baseline and threshold values manually. All gene expressions studies were performed at the same location at the Laboratory of the Institute of Medical Genetics, University Medical Centre Ljubljana, Slovenia, and were interpreted independently by two investigators.

2.8. Statistical analyses:

The significance of association was analysed using the Chi-Square test (χ^2). Odds ratios (OR) and their respective 95% confidence intervals (CI) were also calculated to compare the allelic frequency and genotype distribution in patients and control subjects. Analyses were performed using the Genetic Analysis Package for R available from http://cran.rproject.org/web/packages/gap/. χ^2 test for deviation from Hardy-Weinberg was calculated. Associations were regarded as significant when they reached the P-value equal to or less than 0.05. Haplotype frequencies were estimated using the haplo.stats package [135]. Differences in gene expression between samples could then be calculated using the delta-delta Ct method, as previously described [136]. The significance of expression differences in the two groups investigated in the validation phase was calculated using the two-sample t-distribution test, and differences were deemed significant at alpha values under 0.05. For measuring the statistical strength of a monotonic relationship between SNPs in investigated genes compared with gene expression in blood, we used a Spearman's correlation coefficient. Associations were regarded as significant when they reached the P-value equal to or less than 0.05.

3. RESULTS AND DISCUSSION

The experimental group of patients with MS consisted of 749 females and 309 males, 47.5 ± 26.5 years of age at blood sampling. The female to male ratio was 2.49. At disease onset, they were 40 ± 25 years old and the mean disease duration was 23 ± 22 years. The control group consisted of 564 female and 476 male subjects of the same ethnic background (mean age 46.5 ± 24.5 years). Table 2 presents disease-related characteristics on the studied patients' population used for association studies. Table 3 presents disease-related characteristics on the studied patients' population used for gene expression studies in blood.

The genotype frequencies of investigated polymorphisms were in accordance with those predicted by the Hardy-Weinberg equilibrium in patients and in the control group ($p<0.05$). As multiple SNPs were investigated, appropriate corrections of significance values were also applied using the Bonferroni correction method.

Table 2: Patients' disease-related characteristics for the studied group in association studies

Course	°N (%)	Female (%)	Male (%)
RR MS	722 (68.24)	521 (49.24)	201 (19.00)
SP MS	229 (21.64)	159 (15.03)	70 (6.62)
PP MS	70 (6.62)	41 (3.88)	29 (2.74)
Benign MS	22 (2.08)	15 (1.42)	7 (0.66)
CIS	15 (1.42)	13 (1.23)	2 (0.19)

Table 3: Patients' disease-related characteristics for the studied group in gene expression in blood

MS course	°N (%)	Female (%)	Male (%)
RR MS	19 (38.00)	10 (20.00)	9 (18.00)
SP MS	10 (20.00)	6 (12.00)	4 (8.00)
PP MS	15 (30.00)	7 (14.00)	8 (16.00)
Benign MS	4 (8.00)	3 (6.00)	1 (2.00)
CIS	2 (4.00)	2 (4.00)	0 (0.00)

3.1. Results of the STAT3 and STAT5A genes

3.1.1 Association studies in the STAT3 and STAT5A genes

The results of genotyping and allelic distribution of the STAT3 and STAT5A polymorphisms in MS patients and healthy controls are presented in Table 4. A significant difference presented in Table 4 was found in the distribution of the STAT3 gene in rs963987 polymorphism genotypes with P-value 0.025, and in the distribution of the STAT5A gene in rs6503695 polymorphism genotypes with P-value $7.5 \cdot 10^{-4}$.

Table 4: Genotype and allelic distribution of the STAT3 and STAT5A polymorphisms in MS patients and healthy controls

Genotypes			Control (%)	MS (%)	χ^2	P	P corr.	OR [95% confidence interval]
STAT3	**rs7211777**	**AA**	401 (42.66)	408 (42.46)	2.26	0.20	1.00	AA+AG/GG 1.24 [0.93–1.65]
		AG	442 (47.02)	433 (45.06)				
		GG	97 (10.32)	120 (12.49)				
	rs963987	**AA**	89 (9.36)	135 (13.85)	10.51	**$5.00 \cdot 10^{-3}$**	**0.02**	AG+GG/AA 1.56 [1.17–2.07]
		AG	392 (41.22)	361 (37.03)				
		GG	470(49.42)	479 (49.13)				
	rs1053004	**AA**	350 (38.00)	388 (41.28)	2.73	0.25	1.00	AG+GG/AA 1.15 [0.95–1.38]
		AG	413 (44.84)	411 (43.72)				
		GG	158 (17.16)	141 (15.00)				
STAT5A	**rs6503695**	**CC**	112 (12.71)	193 (20.86)	22.18	**$1.50 \cdot 10^{-5}$**	**$7.50 \cdot 10^{-4}$**	CT+TT/CC 1.81 [1.41–2.34]
		CT	395 (44.84)	393 (42.49)				
		TT	374 (42.45)	339 (36.65)				
	rs12601982	**AA**	599 (65.39)	660 (66.94)	3.81	0.15	0.74	AA+AG/GG 1.48 [0.9–2.47]
		AG	291 (31.77)	285 (28.90)				
		GG	26 (2.84)	41 (4.16)				

The results of genotyping and allelic distribution of the STAT3 and STAT5A polymorphisms in MS patients and healthy controls separated by gender are presented in Table 5. A significant difference was observed in the distribution of the STAT5A gene

in rs6503695 polymorphism genotypes of the female (P=0.01) and the male (P=0.045) MS population. No significant differences were observed in the STAT3 gene in terms of gender.

Table 5: Genotype and allelic distribution of the STAT3 and STAT5A polymorphisms in MS patients and healthy controls in terms of gender

Genotypes, females			Control (%)	MS (%)	χ^2	P	P corr.	OR [95% confidence interval]
STAT3	**rs7211777**	**AA**	233 (42.67)	290 (42.71)	1.63	0.44	1.00	AA+AG/GG
		AG	259 (47.44)	307 (45.21)				1.25 [0.87–1.81]
		GG	54 (9.89)	82 (12.08)				
	rs963987	**AA**	53 (9.74)	98 (14.16)	5.87	**0.05**	0.26	AG+GG/AA
		AG	220 (40.44)	256 (36.99)				1.53 [1.07–2.19]
		GG	271 (49.82)	338 (48.84)				
	rs1053004	**AA**	193 (36.42)	279 (41.58)	6.61	**0.04**	0.18	AG+GG/AA
		AG	238 (44.91)	300 (44.71)				1.24 [0.98–1.57]
		GG	99 (18.68)	92 (13.71)				
STAT5A	**rs6503695**	**CC**	64 (13.09)	136 (20.86)	12.82	**$2.00\cdot10^{-3}$**	**0.01**	CT+TT/CC
		CT	215 (43.97)	279 (42.79)				1.75 [1.27–2.43]
		TT	210 (42.94)	237 (36.35)				
	rs12601982	**AA**	350 (66.29)	469 (66.90)	1.33	0.51	1.00	AA+AG/GG
		AG	164 (31.06)	206 (29.39)				1.41 [0.74–2.81]
		GG	14 (2.65)	26 (3.71)				
Genotypes, males			**Control (%)**	**MS (%)**	χ^2	**P**	**P corr.**	**OR [95% confidence interval]**
STAT3	**rs7211777**	**AA**	158 (43.41)	117 (41.94)	0.99	0.61	1.00	AA+AG/GG
		AG	167 (45.88)	125 (44.80)				1.27 [0.79–2.06]
		GG	39 (10.71)	37 (13.26)				
	rs963987	**AA**	34 (8.99)	37 (13.21)	3.58	0.17	0.83	AG+GG/AA
		AG	157 (41.53)	103 (36.79)				1.54 [0.94–2.53]
		GG	187 (49.47)	140 (50.00)				
	rs1053004	**AA**	145 (40.17)	108 (40.60)	1.34	0.51	1.00	AG+GG/AA
		AG	162 (44.88)	110 (41.35)				1.02 [0.74–1.41]
		GG	54 (14.96)	48 (18.05)				
STAT5A	**rs6503695**	**CC**	44 (12.05)	57 (21.11)	9.53	**$9.00\cdot10^{-3}$**	**0.04**	CT+TT/CC
		CT	167 (45.75)	112 (41.48)				1.95 [1.27–3.01]
		TT	154 (42.19)	101 (37.41)				
	rs12601982	**AA**	231 (64.53)	189 (67.02)	3.28	0.19	0.97	AA+AG/GG
		AG	116 (32.40)	78 (27.66)				1.76 [0.8–4.03]
		GG	11 (3.07)	15 (5.32)				

The results of genotyping and allelic distribution of the STAT3 and STAT5A polymorphisms in MS patients in terms of the course of MS (RRMS and SPMS) are presented in Table 6.

Table 6: Genotype and allelic distribution of the STAT3 and STAT5A polymorphisms in MS patients and healthy controls in terms of the course of RRMS and SPMS

Genotypes, RRMS			**Control (%)**	**MS (%)**	**χ^2**	**P**	**P corr.**	**OR [95% confidence interval]**
STAT3	**rs7211777**	**AA**	401 (42.66)	280 (43.01)	3.69	0.15	0.79	AA+AG/GG
		AG	442 (47.02)	285 (43.78)				0.76 [0.55–1.02]
		GG	97 (10.32)	86 (13.21)				
	rs963987	**AA**	89 (9.36)	98 (14.50)	11.05	**$4.00 \cdot 10^{-3}$**	**0.02**	AG+GG/AA
		AG	392 (41.22)	249 (36.83)				0.61 [0.44–0.82]
		GG	470 (49.42)	329 (48.67)				
	rs1053004	**AA**	350 (38.00)	265 (41.41)	1.86	0.39	1.00	AG+GG/AA
		AG	413 (44.84)	273 (42.66)				0.87 [0.71–1.06]
		GG	158 (17.16)	102 (15.94)				
STAT5A	**rs6503695**	**CC**	112 (12.71)	225 (35.77)	131.39	**0.00**	**$1.97 \cdot 10^{-6}$**	CT+TT/CC
		CT	395 (44.84)	263 (41.81)				3.44 [2.74–4.31]
		TT	374 (42.45)	141 (22.42)				
	rs12601982	**AA**	599 (65.39)	452 (66.67)	1.20	0.55	1.00	AA+AG/GG
		AG	291 (31.77)	202 (29.79)				0.80 [0.45–1.40]
		GG	26 (2.84)	24 (3.54)				
Genotypes, SPMS			**Control (%)**	**MS (%)**	**χ^2**	**P**	**P corr.**	**OR [95% confidence interval]**
STAT3	**rs7211777**	**AA**	401 (42.66)	91 (43.54)	0.29	0.86	1.00	AA+AG/GG
		AG	442 (47.02)	99 (47.37)				0.87 [0.51–1.44]
		GG	97 (10.32)	19 (9.09)				
	rs963987	**AA**	89 (9.36)	29 (14.72)	7.61	**0.02**	0.11	AG+GG/AA
		AG	392 (41.22)	65 (32.99)				1.68 [1.05–2.6]
		GG	470 (49.42)	103 (52.28)				
	rs1053004	**AA**	350 (38.00)	83 (41.50)	4.65	0.09	0.49	AG+GG/AA
		AG	413 (44.84)	95 (47.50)				1.16 [0.85–1.58]
		GG	158 (17.16)	22 (11.00)				
STAT5A	**rs6503695**	**CC**	112 (12.71)	27 (13.71)	0.56	0.76	1.00	CT+TT/CC
		CT	395 (44.84)	92 (46.70)				1.09 [0.68–1.7]
		TT	374 (42.45)	78 (39.59)				
	rs12601982	**AA**	599 (65.39)	141 (69.12)	2.24	0.32	1.00	AA+AG/GG
		AG	291 (31.77)	55 (26.96)				1.42 [0.59–3.06]
		GG	26 (2.84)	8 (3.92)				

A significant difference was found in the distribution of polymorphism genotypes in patients with RRMS in both genes (the STAT3 gene in rs963987 P-value 0.02, and STAT5A gene in rs6503695 P-value 1.97 10^{-6}). On the other hand, no significant differences were observed in the distribution of the STAT3 and STAT5A genes in patients with SPMS.

Table 7 presents the results of genotyping and allelic distribution of the STAT3 and STAT5A polymorphisms in PPMS patients and healthy controls. A significant difference was observed in the distribution of the STAT5A gene in rs12601982 polymorphism genotypes, with P-value 0.005.

Table 7: Genotype and allelic distribution of the STAT3 and STAT5A polymorphisms in PPMS patients and healthy controls

Genotypes, PPMS			**Control (%)**	**MS (%)**	χ^2	**P**	**P corr.**	**OR [95% confidence interval]**
STAT3	**rs7211777**	**AA**	401 (42.66)	28 (41.79)	0.65	0.72	1.00	AA+AG/GG
		AG	442 (47.02)	30 (44.78)				1.37 [0.61–2.72]
		GG	97 (10.32)	9 (13.43)				
	rs963987	**AA**	89 (9.36)	4 (5.97)	1.03	0.59	1.00	AG+GG/AA
		AG	392 (41.22)	27 (40.30)				0.64 [0.19–1.6]
		GG	470 (49.42)	36 (53.73)				
	rs1053004	**AA**	350 (38.00)	28 (43.08)	0.73	0.69	1.00	AG+GG/AA
		AG	413 (44.84)	26 (40.00)				1.24 [0.74–2.05]
		GG	158 (17.16)	11 (16.92)				
STAT5A	**rs6503695**	**CC**	112 (12.71)	16 (25.00)	8.18	**0.02**	0.08	CT+TT/CC
		CT	395 (44.84)	22 (34.38)				2.30 [1.22–4.12]
		TT	374 (42.45)	26 (40.63)				
	rs12601982	**AA**	599 (65.39)	48 (69.57)	13.13	**$1.00 \cdot 10^{-3}$**	**$5.00 \cdot 10^{-3}$**	AA+AG/GG
		AG	291 (31.77)	14 (20.29)				3.91 [1.5–8.98]
		GG	26 (2.84)	7 (10.14)				

In Table 8, the results of inferred haplotypes in the STAT3 and STAT5A genes are presented. Only the results of 5% and above of present haplotypes are presented. Statistically significant differences in haplotype distribution supported our results

obtained in association studies. Significant results were observed at the STAT3 gene locus in haplotype AGG with P-value $2.5 \cdot 10^{-4}$, and at the STAT5A gene locus in haplotypes CA with P-value $3 \cdot 10^{-5}$ and TGA with P-value $6.6 \cdot 10^{-4}$.

Table 8: Haplotypes in the STAT3 and STAT5A genes

Control=986 , MS=1049								
STAT3 gene								
	rs7211777	**rs963987**	**rs1053004**	**Hap freq. (%)**	**P-value**	**P corr.**	**% control**	**%MS**
1	A	A	A	6.07	0.01	0.07	4.97	7.14
2	A	G	A	52.62	0.89	1.00	52.77	52.54
3	**A**	**G**	**G**	5.71	**$5.00 \cdot 10^{-5}$**	**$2.50 \cdot 10^{-4}$**	7.43	3.95
4	G	A	G	21.86	0.83	1.00	21.99	21.80
5	G	G	G	9.42	0.38	1.00	9.42	8.73
STAT5A gene								
		rs6503695	**rs12601982**	**Hap freq. (%)**	**P-value**	**P corr.**	**% control**	**%MS**
1		**C**	**A**	23.39	**$1 \cdot 10^{-5}$**	**$3 \cdot 10^{-5}$**	19.93	26.58
2		C	G	15.20	0.49	1.00	14.94	15.55
3		**T**	**A**	57.94	**$2.20 \cdot 10^{-4}$**	**$6.60 \cdot 10^{-4}$**	61.25	54.85

To conclude, we found statistically significant differences in the STAT3 and STAT5A genes in association studies. The differences in rs6503695 in the STAT5A gene were significant in the entire group of patients and also when separated by gender, whereas the differences in rs963987 in the STAT3 gene were significant only in the entire group of participants. We supported our results by means of haplotype distributions in both genes. In the studied population, the AGG haplotype in the STAT3 gene might be a protective factor in MS. Moreover, in the STAT5A gene, the CA haplotype might be a risk factor and the TA haplotype might be a protective factor in MS.

3.1.2. Gene expression in the STAT3 and STAT5A genes

In order to support our results from genotyping, we conducted gene expression studies in blood. The results of 50 patients and 40 controls with regard to gene expression in blood in the STAT3 and STAT5A genes as obtained within the entire group of patients and by gender are presented in Chart 1. Significant differences were limited to the male population in the STAT3 gene with P-value 0.02.

Chart 1: Differential expression of genes in MS in comparison to controls in the STAT3 and STAT5A genes in all participants and separated by gender

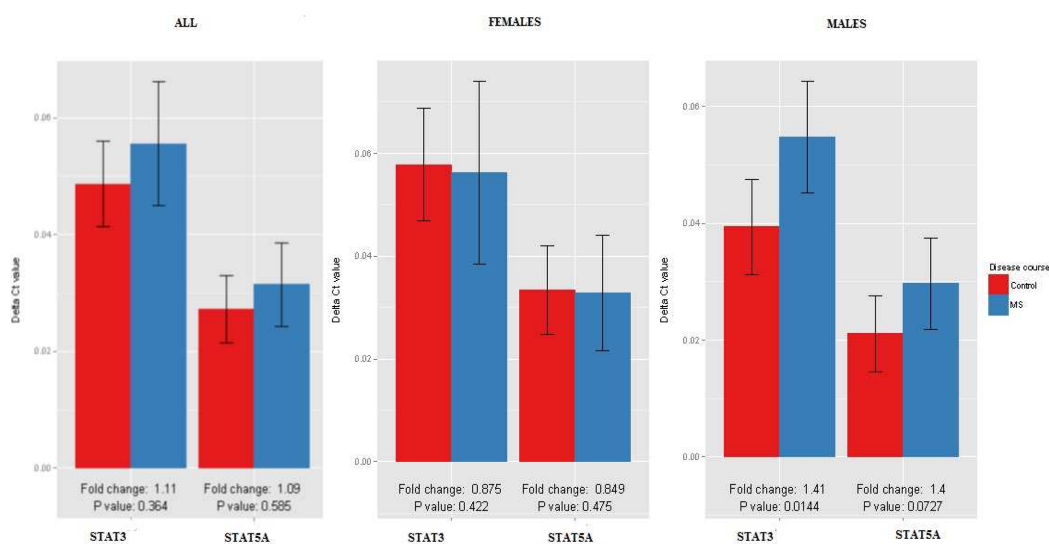
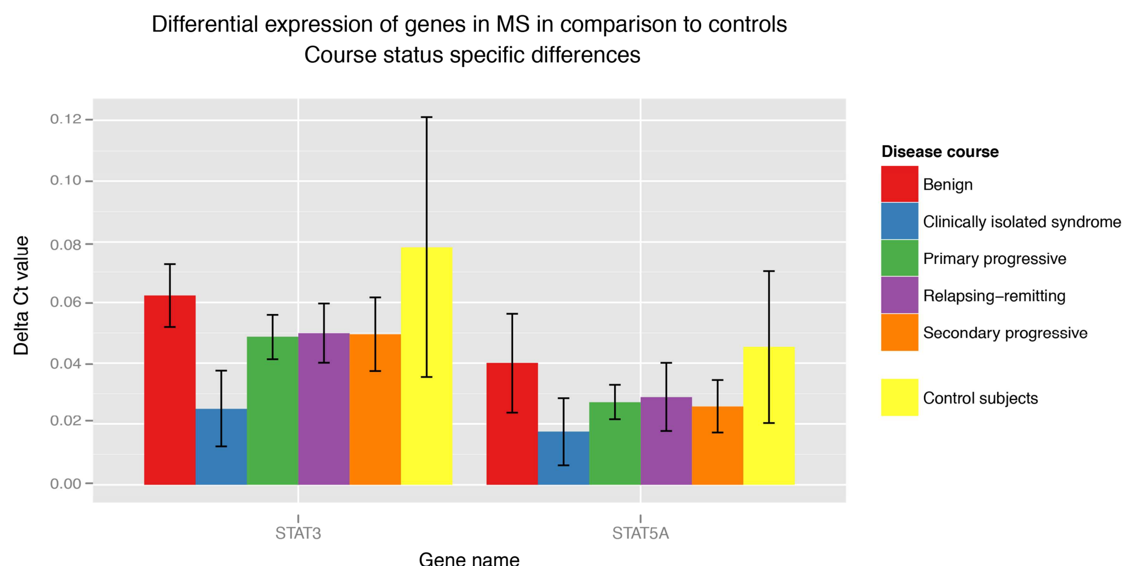


Chart 2 represents the results of gene expression profiling in blood for the STAT3 and STAT5A genes, stratified according to the course of the disease. The graphs represent expression profiles of STAT3 and STAT5A for subsets of all subjects affected by distinct MS course patterns, along with healthy controls. The height of bars correspond to the level of expression (measured quantitatively as delta Ct values). Colour of bars reflect expression in varying MS course patterns and yellow bars reflect expression in the group of healthy controls. Error bars in the plot correspond to the 95% confidence intervals for measured expression values.

Chart 2: Differential expression of genes in MS in comparison to controls in the STAT3 and STAT5A genes in all participants according to the disease course



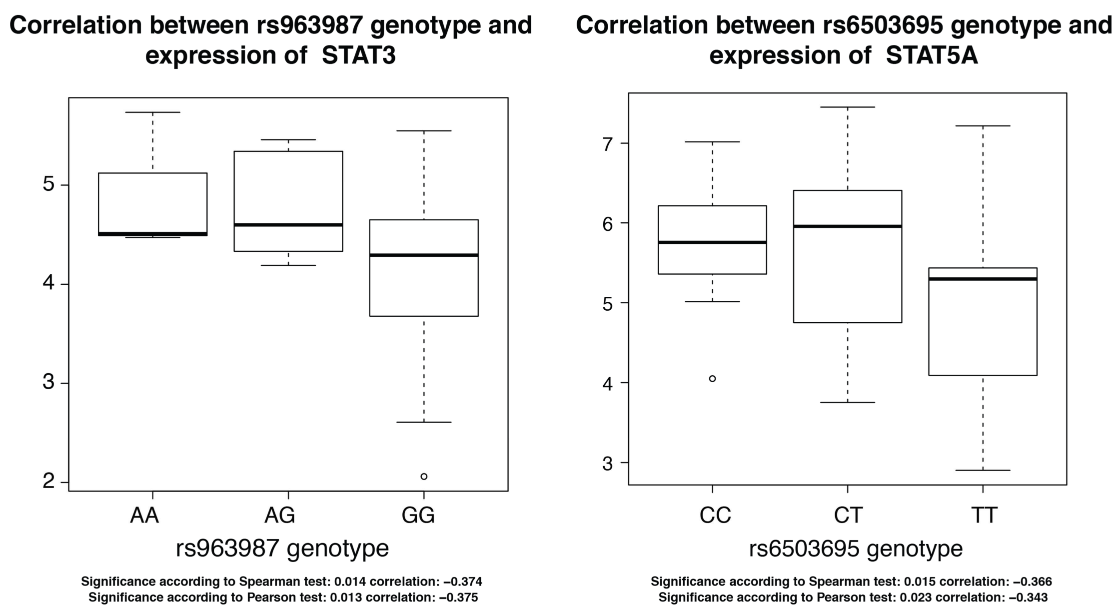
To conclude, we found statistically significant differences in gene expression of the STAT3 gene in blood of the male population.

3.1.3. Correlation between SNPs and blood expression in the STAT3 and STAT5A genes

Due to the observed results we were also interested in the correlation of SNPs in the examined STAT3 and STAT5A genes and their expression in blood. Correlating the genotype expression results from the same patients has revealed a significant relation between the genotype and expression levels for SNPs in STAT3 and STAT5A. We have detected a significant correlation of the rs963987 genotypes with the STAT3 gene expression (P=0.014) and of the rs6503695 (P=0.015) genotype with the STAT5A gene expression according to the non-parametric Spearman rank correlation test. The results are presented graphically in Chart 3. We have detected the same trend for genotype expression correlation in the GeneVar database (the rs963987 STAT3 gene with P=0.064 and the rs6503695 STAT5A gene with P=0.479). Correlation of MS-associated

SNP genotype with differences in gene expression in STAT3 and STAT5A may reveal the functional basis for associations of polymorphisms with MS.

Chart 3: Correlation between SNP rs963987A/G and expression of STAT3 and Correlation between SNP rs6503695 C/T and expression of STAT5A



3.2. Discussion on the STAT3 and STAT5A genes

The integratomic approach developed by our research group enables the synthesis of data originating from various types of genomic studies, and facilitates the discovery of disease candidate genes based on empiric evidence from a multitude of biological levels examined in omic studies. Using the information on the position and the significance of alterations detected in separate omic studies, it is possible to determine genomic regions where significant accumulation of signals from heterogeneous biological levels occurs. This approach might therefore serve in promoting the discovery of specific genomic variants playing a role in complex diseases where true biological signals are commonly lost in the noise of false-positive results, especially due to their low-effect sizes.

Detection of these biological alterations may, however, be promoted by the incorporation of several levels of biological information into integrated analyses of omic data. For this reason, we carried out a selection of candidate genes for multiple sclerosis using this approach.

Based on the results obtained through our integratomic approach, we hypothesized that genetic variants in STAT genes (the STAT3 and STAT5A genes) might be associated with multiple sclerosis. We found a statistically significant difference in the allelic distribution of the STAT5A and STAT3 genes. This association was supported by the examination of haplotype distribution at the STAT5A and STAT3 gene loci. Moreover, correlating the genotype-expression results from the same patients has revealed a significant correrelation between the genotypes at STAT3 and STAT5A SNPs and gene expression levels. Correlation of MS-associated SNPs with differences in gene expression in STAT3 and STAT5A might imply the functional basis for these associations. Although we have found correlation between genotype and expression, we have not found statistically significant differences in the expression of these two genes in blood samples, with the exception of the STAT3 gene expression in the male patient population.

The genomic region containing the STAT3 and STAT5A (17q11.2-22) genes could represent a plausible candidate region for MS for several reasons:

1. Genome-wide association studies and GWAS meta-analyses data suggested several SNPs in the STAT3-STAT5A genomic region to be associated with MS.
2. In a global expression profiling study of peripheral blood from MS patients, a differential expression of STAT3 and STAT5A was detected.
3. Proteins coded by the STAT3 and STAT5A genes were found to interact with proteins related to EAE in mouse knock-out models.
4. Activation of STATs results in the expression of genes that control critical cellular functions associated with immune responses.

Ad1. Meta-analysis of GWAS in MS implicated STAT3 as one of the candidate genes with a suggestive role in inflammatory bowel disease and multiple sclerosis [16]. Jakkula et al. [127] found an association between multiple sclerosis and rs744166 polymorphism in the STAT3 gene with the A allele conferring a protective effect against MS. The same group also found evidence for association between STAT3 and MS risk in other populations of European descent [127]. Patsopoulos et al. [137] found an association between rs2293152 and STAT3 locus with MS susceptibility. To validate findings by Jakkula et al. [127] and Patsopoulos et al. [137] obtained in GWAS studies, Lill et al. [138] performed a replication genetic association study of the STAT3 gene in a German case-control sample. There was a nominally significant association between the G allele of rs744166 and MS, and no association with rs2293152.

Ad2 In global expression profiling in peripheral blood samples, differential expression of the STAT3 and STAT5A genes between patients and healthy controls was detected [129,130,131]. It is known that STAT3 promotes differentiation of Th17 and follicular helper T cell subsets [139]. Although the exact pathogenesis of MS remains to be fully clarified, current evidence suggests an autoimmune etiology which includes infiltration of T cells, B cells and macrophages in active MS brain lesions. The immune attack in the CNS is mediated by autoreactive T cells that enter through a disturbed blood-brain barrier (BBB) and attack oligodendrocytes and myelin. Expression of genes tagging T cells was upregulated in blood samples from patients with all forms of MS [129].

Ad3 Studies on mice showed that the loss of STAT3 in $CD4^+$ T cells prevents development of experimental autoimmune diseases like experimental autoimmune encephalomyelitis (EAE), which is a multiple sclerosis mouse model [140]. This suggests that the expression of STAT3 in $CD4^+$ T cells is essentially related to the development of EAE [140]. STAT3 pathways and Th17 cells therefore represent attractive targets for inhibiting CNS autoimmune diseases.

Ad4 Activation of STATs results in the expression of genes that control critical cellular functions associated with immune responses [89–92]. In normal cells, these signal

transduction mechanisms are closely controlled to prevent unscheduled gene regulation and, consequently, inappropriate biological responses.

STAT3 is known to be activated in the brain and is part of a broader network that includes key regulators of cellular signalling [141]. Dysregulation of the STAT3 pathway may play a role in MS for several reasons, namely the importance of STAT3 in myeloid cell activation, T-cell differentiation and cytokine/chemokine induction [142]. Signalling through STAT3 is a critical component of Th17-dependent autoimmune processes [143]. Moreover, an increased phosphorylated (activated) STAT3 was reported in the T cells of patients evolving from a clinically isolated syndrome to the defined MS and in relapsing patients [143].

Through this work, we provided further evidence of the implication of STAT genes by identifying the STAT5A gene associated with MS susceptibility. Namely, in addition to STAT3, STAT5 has also been implicated in cellular functions of proliferation, differentiation and apoptosis with relevance to the processes of immunoregulation [106]. Unrestricted STAT5 activation can lead to pathological conditions such as inflammation or autoimmunity, cancer promotion and progression, and myeloproliferative diseases. Therefore, the activity of STAT5 needs to be controlled closely [144]. Precise mechanisms through which the STAT5 proteins exert their function are still largely unknown, as are the post-translational modifications of STAT5 necessary for its proper function. In general, proper functioning of the STAT5 molecules is essential for all immune cells and these processes include lymphocyte development, NK cell activity, cytotoxic T cell function, T helper or suppressor/regulatory T cell function, mast cell function, platelet/megakaryocyte function, and macrophage responses or stress erythropoiesis [145–151]. Prominent direct transcriptional targets of STAT5 include (i) proteins important for cell cycle progression and cellular growth, (ii) tissue invasion, (iii) survival, (iv) negative feedback inhibition in tyrosine kinase signalling pathways, (v) lymphocyte function, (vi) cofactor regulation, (vii) liver function, (viii) major urinary proteins, (ix) ribosomal proteins, (x) acute phase response genes, but also (xi) genes involved in DNA damage repair, (xii) reproduction or (xiii) mammary gland function and differentiation [145–152]. The way in which the STAT5 proteins exert their distinct

and specific functions in different cell types under various physiological and pathological conditions remains to be elucidated.

3.3. Results of the CLOCK and ARNTL genes

3.3.1. Association studies on the CLOCK and ARNTL genes

Genotype frequencies of investigated polymorphisms in ARNTL and CLOCK genes were in accordance with those predicted by the Hardy-Weinberg equilibrium in the case and control groups, with the exception of rs12363415 which was excluded from further analyses.

The results of genotype and allelic distribution of the ARNTL and CLOCK polymorphisms in 900 patients with MS and 1026 controls are presented in Table 9. A significant difference in the distribution of the ARNTL gene in rs3789327 polymorphism genotypes was observed in patients with MS, with P-value $7.49 \cdot 10^{-5}$. A significant difference was also detected in the distribution of genotypes in rs6811520 polymorphism in the CLOCK gene, with P-value 0.021.

Table 9: Genotype and allelic distribution of the ARNTL and CLOCK polymorphisms in MS patients and healthy controls

Genotypes			Total control (%)	Total MS (%)	χ^2	P	P corr.	OR [95% confidence interval]
ARNTL GENE	rs3789327	CC	198 (19.68)	257 (29.11)				TT+CT/CC
		CT	489 (48.61)	375 (42.47)	22.89	**$1.07 \cdot 10^{-5}$**	**$7.49 \cdot 10^{-5}$**	1.67 [1.35–2.07]
		TT	319 (31.71)	251 (28.43)				
	rs1481892	CC	93 (9.57)	73 (9.97)				CC+GC/GG
		GC	424(43.62)	313 (42.76)	0.16	0.92	1.00	1.02 [0.84–1.23]
		GG	455 (46.81)	346 (47.27)				
	rs4757144	AA	280 (31.67)	190 (27.54)				AA+AG/GG
		GA	422 (47.74)	354 (51.30)	3.28	0.19	1.00	1.04 [0.81–1.32]
		GG	182 (20.59)	146 (21.16)				
CLOCK GENE	rs6811520	CC	598 (64.09)	512 (72.01)				CT+TT/CC
		CT	274 (29.37)	166 (23.35)	11.75	**$3.00 \cdot 10^{-3}$**	**0.02**	1.44 [1.17–1.78]
		TT	61 (6.54)	33 (4.64)				
	rs6850524	CC	151 (15.07)	138 (15.94)				CC+GC/GG
		GC	502 (50.10)	403 (46.54)	2.38	0.30	1.00	1.12 [0.93–1.36]
		GG	349 (34.83)	325 (37.53)				
	rs11932595	AA	306 (33.89)	250 (30.38)				AA+AG/GG
		GA	450 (49.83)	397 (48.95)	4.96	0.08	0.59	1.30 [1.02–1.67]
		GG	147 (16.28)	164 (20.22)				
	rs13124436	AA	88 (11.47)	83 (10.81)				AA+AG/GG
		GA	336 (43.81)	331 (43.10)	0.36	0.84	1.00	1.06 [0.86–1.29]
		GG	343 (44.72)	354 (46.09)				

The results of genotype and allelic distribution of the ARNTL and CLOCK gene polymorphisms in the male population are presented in Table 10. No statistically significant distributions of polymorphisms were observed in the ARNTL and CLOCK genes in the male population.

Table 10: Genotype and allelic distribution of the ARNTL and CLOCK polymorphisms in male MS patients and healthy male controls

Genotypes, male			Control (%)	MS (%)	χ^2	P	P corr.	OR [95 % confidence interval]
ARNTL GENE	**rs3789327**	**CC**	95 (20.21)	82 (29.71)	9.57	**$8.00 \cdot 10^{-3}$**	0.06	TT+CT/CC
		CT	226 (48.09)	125 (45.29)				1.67 [1.18–2.35]
		TT	149 (31.70)	69 (25.00)				
	rs1481892	**CC**	47 (10.54)	21 (9.77)	0.23	0.89	1.00	CC+GC/GG
		GC	193 (43.27)	97 (45.12)				0.96 [0.69–1.33]
		GG	206 (46.19)	97 (45.12)				
	rs4757144	**AA**	115 (28.19)	60 (27.52)	0.40	0.82	1.00	AA+AG/GG
		GA	211 (51.72)	118 (54.13)				0.90 [0.58–1.36]
		GG	82 (20.10)	40 (18.35)				
CLOCK GENE	**rs6811520**	**CC**	292 (68.71)	148 (73.27)	1.63	0.44	1.00	CT+TT/CC
		CT	108 (25.41)	42 (20.79)				1.25 [0.86–1.82]
		TT	25 (5.88)	12 (5.94)				
	rs6850524	**CC**	71 (15.27)	50 (18.45)	1.93	0.38	1.00	CC+GC/GG
		GC	231 (49.68)	122 (45.02)				1.07 [0.78–1.46]
		GG	163 (35.05)	99 (36.53)				
	rs11932595	**AA**	150 (34.25)	80 (31.13)	4.07	0.13	0.92	AA+AG/GG
		GA	225 (51.37)	125 (48.64)				1.51 [1–2.26]
		GG	63 (14.38)	52 (20.23)				
	rs13124436	**AA**	42 (11.35)	23 (9.87)	1.20	0.55	1.00	AA+AG/GG
		GA	165 (44.59)	97 (41.63)				1.20 [0.86–1.66]
		GG	163 (44.05)	113 (48.50)				

The results of genotype and allelic distribution of the ARNTL and CLOCK genes in the female population are presented in Table 11. A significant difference was observed in the distribution of rs3789327 polymorphism genotypes of the ARNTL gene with P-value 0.007. A significant difference was also observed in the distribution of rs6811520 polymorphism genotypes of the CLOCK gene with P-value 0.007, as presented in Table 11.

Table 11: Genotype and allelic distribution of the ARNTL and CLOCK polymorphisms in the female population of MS patients and healthy female controls

Genotypes, female			Control (%)	MS (%)	χ^2	P	P corr.	OR [95% confidence interval]
ARNTL GENE	**rs3789327**	**CC**	103 (19.22)	175 (28.83)	15.03	**$1.00 \cdot 10^{-3}$**	**$7.00 \cdot 10^{-3}$**	TT+CT/CC
		CT	263 (49.07)	250 (41.19)				1.70 [1.29–2.25]
		TT	170 (31.72)	182 (29.98)				
	rs1481892	**CC**	46 (8.75)	52 (10.06)	0.79	0.67	1.00	CC+GC/GG
		GC	231 (43.92)	216 (41.78)				1.03 [0.81–1.32]
		GG	249 (47.34)	249 (48.16)				
	rs4757144	**AA**	165 (34.66)	130 (27.54)	5.71	0.06	0.41	AA+AG/GG
		GA	211 (44.33)	236 (50.00)				1.09 [0.8–1.48]
		GG	100 (21.01)	106 (22.46)				
CLOCK GENE	**rs6811520**	**CC**	306 (60.24)	364 (71.51)	15.05	**$1.00 \cdot 10^{-3}$**	**$7.00 \cdot 10^{-3}$**	CT+TT/CC
		CT	166 (32.68)	124 (24.36)				0.57 [0.32–0.98]
		TT	36 (7.09)	21 (4.13)				
	rs6850524	**CC**	80 (14.90)	88 (14.79)	1.48	0.48	1.00	CC+GC/GG
		GC	271 (50.47)	281 (47.23)				1.16 [0.91–1.47]
		GG	186 (34.64)	226 (37.98)				
	rs11932595	**AA**	156 (33.55)	170 (30.69)	1.28	0.53	1.00	AA+AG/GG
		GA	225 (48.39)	272 (49.10)				1.15 [0.84–1.58]
		GG	84 (18.06)	112 (20.22)				
	rs13124436	**AA**	46 (11.59)	60 (11.21)	0.06	0.97	1.00	AA+AG/GG
		GA	171 (43.07)	234 (43.74)				0.99 [0.76–1.28]
		GG	180 (45.34)	241 (45.05)				

The analyses were also stratified by clinically defined subtypes of MS. The results of genotype and allelic distribution of the ARNTL and CLOCK polymorphisms in RRMS patients and healthy controls are presented in Table 12. A significant difference in the distribution of rs3789327 polymorphism genotypes of the ARNTL gene in RR form of MS was observed when compared to controls, with P-value $2.5 \cdot 10^{-5}$.

Table 12: Genotype and allelic distribution of the ARNTL and CLOCK polymorphisms in RRMS patients and healthy controls

RR MS n=596								
Genotypes			**Control (%)**	**MS (%)**	χ^2	**P**	**P corr.**	**OR [95% confidence interval]**
ARNTL GENE	**rs3789327**	**CC**	198 (19.68)	180 (30.20)	25.10	$\mathbf{3.50 \cdot 10^{-6}}$	$\mathbf{2.50 \cdot 10^{-5}}$	TT+CT/CC
		CT	489 (48.61)	245 (41.11)				1.81 [1.43–2.29]
		TT	319 (31.71)	160 (26.85)				
	rs1481892	**CC**	93 (9.57)	47 (7.89)	0.67	0.71	1.00	CC+GC/GG
		GC	424(43.62)	187 (31.38)				1.05 [0.84–1.32]
		GG	455 (46.81)	217 (36.41)				
	rs4757144	**AA**	280 (31.67)	125 (20.97)	1.54	0.46	1.00	AA+AG/GG
		GA	422 (47.74)	221 (37.08)				1.06 [0.8–1.4]
		GG	182 (20.59)	95 (15.94)				
CLOCK GENE	**rs6811520**	**CC**	598 (64.09)	320 (53.69)	7.62	0.02	0.15	CT+TT/CC
		CT	274 (29.37)	104 (17.45)				1.41 [1.11–1.81]
		TT	61 (6.54)	23 (3.86)				
	rs6850524	**CC**	151 (15.07)	92 (15.44)	1.71	0.42	1.00	CC+GC/GG
		GC	502 (50.10)	267 (44.80)				1.11 [0.9–1.37]
		GG	349 (34.83)	213 (35.74)				
	rs11932595	**AA**	306 (33.89)	171 (28.69)	2.64	0.27	1.00	AA+AG/GG
		GA	450 (49.83)	268 (44.97)				1.24 [0.94–1.64]
		GG	147 (16.28)	106 (17.79)				
	rs13124436	**AA**	88 (11.47)	56 (9.40)	0.12	0.94	1.00	AA+AG/GG
		GA	336 (43.81)	228 (38.26)				0.99 [0.79–1.24]
		GG	343 (44.72)	227 (38.09)				

The results of genotype and allelic distribution of the ARNTL and CLOCK polymorphisms in SPMS patients and healthy controls are presented in Table 13. We observed that SNPs in the CLOCK and ARNTL genes in the SP form of MS did not display significant associations with MS susceptibility.

Table 13: Genotype and allelic distribution of the ARNTL and CLOCK polymorphisms in SPMS patients and healthy controls

SP MS n=229								
Genotypes			**Control (%)**	**MS (%)**	χ^2	**P**	**P corr.**	**OR [95% confidence interval]**
ARNTL GENE		**CC**	198 (19.68)	61 (26.64)				TT+CT/CC
	rs3789327	**CT**	489 (48.61)	96 (41.92)	6.16	0.05	0.32	1.50 [1.07–2.08]
		TT	319 (31.71)	70 (30.57)				
		CC	93 (9.57)	16 (6.99)				CC+GC/GG
	rs1481892	**GC**	424(43.62)	88 (38.43)	1.44	0.49	1.00	1.17 [0.87–1.58]
		GG	455 (46.81)	107 (46.72)				
		AA	280 (31.67)	61 (26.64)				AA+AG/GG
	rs4757144	**GA**	422 (47.74)	102 (44.54)	0.36	0.84	1.00	0.97 [0.66–1.41]
		GG	182 (20.59)	41 (17.90)				
CLOCK GENE		**CC**	598 (64.09)	147 (64.19)				CT+TT/CC
	rs6811520	**CT**	274 (29.37)	49 (21.40)	8.16	0.02	0.12	1.52 [1.09–2.15]
		TT	61 (6.54)	5 (2.18)				
		CC	151 (15.07)	33 (14.41)				CC+GC/GG
	rs6850524	**GC**	502 (50.10)	110 (48.03)	0.05	0.98	1.00	1.03 [0.76–1.4]
		GG	349 (34.83)	79 (34.50)				
		AA	306 (33.89)	59 (25.76)				AA+AG/GG
	rs11932595	**GA**	450 (49.83)	106 (46.29)	3.930	0.14	0.98	1.37 [0.93–1.99]
		GG	147 (16.28)	44 (19.21)				
		AA	88 (11.47)	17 (7.42)				AA+AG/GG
	rs13124436	**GA**	336 (43.81)	75 (32.75)	2.88	0.24	1.00	1.30 [0.95–1.79]
		GG	343 (44.72)	97 (42.36)				

The results of genotype and allelic distribution of the ARNTL and CLOCK polymorphisms in PPMS patients and healthy controls are presented in Table 14. We observed that SNPs in the CLOCK and ARNTL genes in SPMS did not display significant associations with MS susceptibility as presented in Table 14.

Table 14: Genotype and allelic distribution of the ARNTL and CLOCK polymorphisms in PP MS patients and healthy controls

PP MS n=48								
Genotypes			**Control (%)**	**MS (%)**	χ^2	**P**	**P corr.**	**OR [95% confidence interval]**
ARNTL GENE	**rs3789327**	**CC**	198 (19.68)	9 (18.75)	0.01	0.99	1.00	TT+CT/CC
		CT	489 (48.61)	22 (45.83)				1.03 [0.46–2.1]
		TT	319 (31.71)	14 (29.17)				
	rs1481892	**CC**	93 (9.57)	4 (8.33)	2.65	0.26	1.00	CC+GC/GG
		GC	424 (43.62)	24 (50.00)				0.61 [0.31–1.15]
		GG	455 (46.81)	15 (31.25)				
	rs4757144	**AA**	280 (31.67)	4 (8.33)	7.18	**0.03**	0.19	AA+AG/GG
		GA	422 (47.74)	23 (47.92)				0.87 [0.32–2.02]
		GG	182 (20.59)	6 (12.50)				
CLOCK GENE	**rs6811520**	**CC**	598 (64.09)	31 (64.58)	4.93	0.08	0.59	CT+TT/CC
		CT	274 (29.37)	6 (12.50)				2.44 [1.12–6.14]
		TT	61 (6.54)	1 (2.08)				
	rs6850524	**CC**	151 (15.07)	9 (18.75)	6.24	**0.04**	0.31	CC+GC/GG
		GC	502 (50.10)	14 (29.17)				1.79 [0.97–3.28]
		GG	349 (34.83)	22 (45.83)				
	rs11932595	**AA**	306 (33.89)	11 (22.92)	3.52	0.17	1.00	AA+AG/GG
		GA	450 (49.83)	18 (37.50)				1.97 [0.92–3.93]
		GG	147 (16.28)	11 (22.92)				
	rs13124436	**AA**	88 (11.47)	7 (14.58)	1.59	0.45	1.00	AA+AG/GG
		GA	336 (43.81)	15 (31.25)				1.12 [0.6–2.11]
		GG	343 (44.72)	20 (41.67)				

In Table 15, the results of inferred haplotypes in the ARNTL and CLOCK genes are presented. Only haplotypes with estimated frequencies above 5% are presented. A statistically significant difference in genotyping was supported by haplotype distributions at the ARNTL gene locus for haplotypes CGG (P=0.004) and TGA (P=0.030). A statistically significant difference was also observed in haplotype distributions at the CLOCK gene locus for TCAG haplotype (P= 0.001).

Table 15: Haplotypes in the ARNTL and CLOCK genes

control=1024, MS=900								
ARNTL gene								
	rs3789327	**rs1481892**	**rs4757144**	**Hap freq. (%)**	**P-value**	**P corr.**	**% control**	**% MS**
1	C	C	G	14.95	0.20	1.00	14.70	15.20
2	C	G	A	22.10	**0.03**	0.18	21.02	23.17
3	**C**	**G**	**G**	8.44	**$7.00 \cdot 10^{-4}$**	**$4.00 \cdot 10^{-3}$**	6.93	10.34
4	T	C	G	11.94	0.22	1.00	12.15	11.66
5	**T**	**G**	**A**	27.61	**$5.00 \cdot 10^{-3}$**	**0.03**	29.31	25.58
6	T	G	G	10.51	0.17	1.00	11.40	9.63

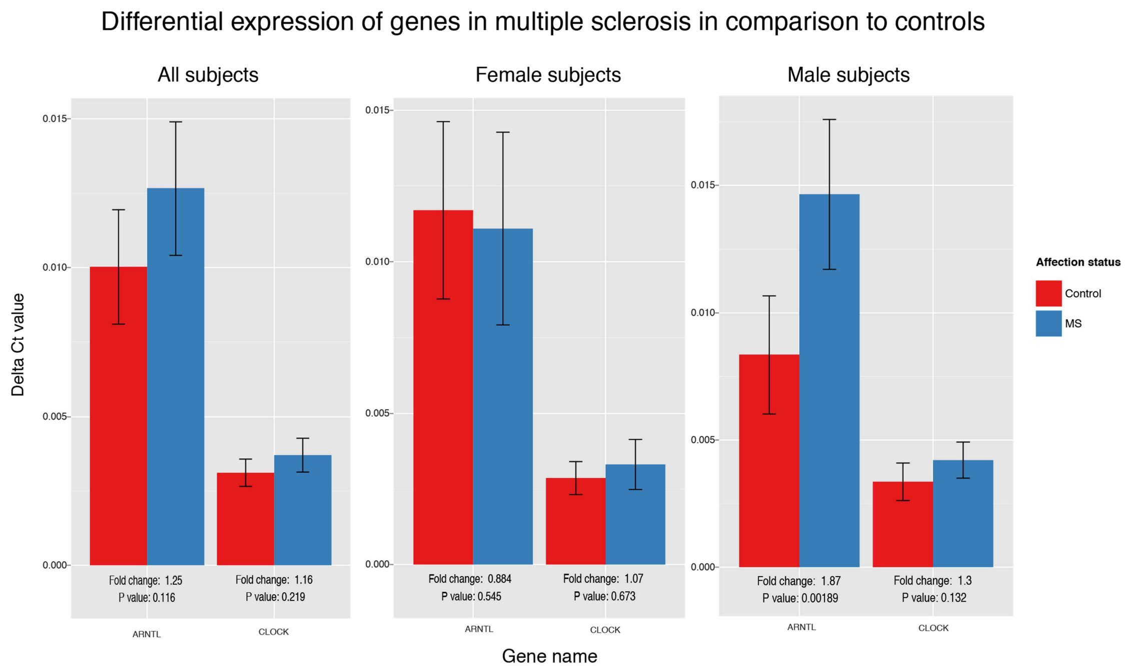
CLOCK gene									
	rs6811520	**rs6050524**	**rs11932595**	**rs13124436**	**Hap freq. (%)**	**P-value**	**P corr.**	**% control**	**% MS**
1	C	C	A	G	15.56	**0.04**	0.23	14.17	17.46
2	C	C	G	G	5.21	0.08	0.53	4.77	5.62
3	C	G	A	A	19.54	0.21	1.00	20.66	18.44
4	C	G	G	A	12.43	0.24	1.00	11.49	13.29
5	C	G	G	G	24.49	0.55	1.00	24.49	24.59
6	**T**	**C**	**A**	**G**	18.14	**$2.00 \cdot 10^{-3}$**	**$1.00 \cdot 10^{-3}$**	20.68	14.99

To conclude, our genotyping results indicate a statistically significant association of variants in the ARNTL and CLOCK genes with MS. The results of individual polymorphism analyses were supported by the analyses of interfered haplotype frequencies in case and control groups. The association for rs3789327 in the ARNTL gene was significant in all participants, as well in stratified sub-populations separated by gender and in the RR-MS subtype of MS. Additionally, we have shown that, in Caucasian population of Slavic origin, the ARNTL gene CGG haplotype might be a risk factor for MS and the ARNTL TGA haplotype might be a protective factor for MS. The association of rs6811520 in the CLOCK gene was also significant in all participants, but when gender-specific stratification analyses were performed, only the female population showed a statistically significant association in this gene. Furthermore, we have shown that, in Caucasian population of Slavic origin, haplotype TCAG in the CLOCK gene might be a protective factor in MS.

3.3.2. Gene expression in the CLOCK and ARNTL genes

Characterization of gene expression in whole blood might lead to the identification of useful clinical biomarkers for MS. For this reason, we also performed gene expression profiling of the ARNTL and CLOCK genes in blood samples in cases with MS and healthy controls. Chart 4 represents the relative gene expression level in blood for the ARNTL and CLOCK genes in all participants, as well as results stratified by gender. We haven't observed any significant differences in gene expression in blood for the ARNTL and CLOCK genes within the entire study population. In gender-separated results of gene expression, we observed significant differences in the ARNTL gene in male patients (P=0.002).

Chart 4: Differential gene expression of blood in the CLOCK and ARNTL genes in all participants and separated by gender



Furthermore, we performed a separate comparison of the ARNTL and CLOCK gene expression levels in blood according to the course of disease and separated by gender. We observed a significant differential expression of the ARNTL gene in female subjects (RR-MS P-value 0.047 and SP-MS P-value 0.027) and in male patients (SP-MS P-value 0.019).

To conclude, we found statistically significant differences in the ARNTL gene expression in MS.

3.3.3. Correlation between SNPs and blood expression in the CLOCK and ARNTL genes

Due to the observed results, we were also interested in the correlation of SNPs in the examined CLOCK and ARNTL genes, and their expression in blood. To determine this relationship, we utilized the Spearman's correlation. Correlating the genotype expression results from the same patients has not revealed a significant correlation between the genotype and expression levels in blood (ARNTL gene: rs3789327 P=0.60; rs1481892 P=0.16; rs4757144 P=0.24 and CLOCK gene: rs6811520P= 0.62; rs6850524 P=0.62; rs11932595 P=0.06 and rs13124436 P=0.17).

3.4. Discussion on the CLOCK and ARNTL genes

We hypothesized that the differences in epidemiology of MS related to the geographical latitude effect might be associated with chronobiology. Moreover, gene variability in key regulators of circadian rhythms genes (CLOCK and ARNTL) might be associated with multiple sclerosis. In answer to our hypothesis, we found a statistically significant association of the ARNTL and CLOCK genes with MS. The differences in the ARNTL gene were detected on a sample of all study participants, but also in the male and female sub-population and in RR-MS patients. We supported our results by determining the

ARNTL and CLOCK gene expression in blood. The CLOCK gene exhibited significant association in a sample of all participants, as well as in the female sub-population. Results of the association study were also supported when comparing differences in the frequencies of inferred haplotypes in a sample of cases and controls.

The association between MS and circadian rhythms is further supported by evidences that:

1. exposure to shift work and migration between high- and low-risk geographic regions increase the risk for MS later on,
2. the lack of exposure to darkness/light is a strong stimulus for melatonin production and is geographically associated with areas of high MS prevalence,
3. The month of one's birth and the risk of MS are related.

Ad1 Studies have shown that exposure to shift work and migration at a young age increase the risk for MS later on [121, 153, 154]. A key issue when discussing health problems in shift workers is the role of the disturbed circadian rhythm. Circadian disruption and sleep restriction are related to the disturbed melatonin secretion and enhanced proinflammatory responses [121]. Many biological variables follow the circadian rhythm, and shift work can interfere with disorders involving these variables. Shift work influences the mismatch of circadian rhythms (light intensity and day length), sleep/wake disturbances, increased susceptibility, internal desynhronization of circadian rhythms, behavioural changes and stress and jet lag-like symptoms and gastrointestinal malfunction [153,155–157]. Chronic circadian disruption may also increase susceptibility to disorders like obesity, diabetes mellitus, cardiovascular disease, thrombosis and even inflammation [153,155–157]. Circadian disruption and sleep deprivation may be contributory factors explaining the association between early exposure to shift work and an increased MS risk [121].

Studies of migrants indicate that the MS risk is strongly associated with the place of residence in early life [154]. The effects of migration between high- and low-risk geographic regions have been examined in several populations (UK immigrants to South

Africa, or Asian and Caribbean immigrants to the UK). Individuals born in low-risk areas appear to benefit from some long-lasting protection that is not, however, transmitted to their children. Immigrants who migrate before adolescence acquire the risk of their new country, while those who migrate later retain the risk of their home country [154]. MS is very rare among black people living in Africa. Although it is somewhat more frequent among black people in the USA, it is still less frequent than among white people of European descent living in the USA [158, 159]. Studies suggest that the risk at the same latitude is lower in individuals with darker skin pigmentation [158, 159].

Ad2 The daily molecular oscillator, known as the circadian clock, senses changes in the photoperiod and mediates a diverse number of photoperiodic responses such as hormone secretion in mammals [160]. Lack of exposure to light/darkness is a strong stimulus for melatonin production and is geographically associated with areas of high MS prevalence [161]. Alternatively, it is possible that producing more melanin represents an evolutionarily advantageous adaptation for populations exposed to higher levels of environmental factors (sun exposure, UV radiation), and that these factors may be involved in triggering autoimmune demyelination. There are evidences that also support a strong seasonal pattern in subclinical MS activity based on non-contrast brain MRI. The disease activity pattern (new T2 activity) is two to three times higher in the period from March to August than during the rest of the year, and correlates strongly with regional climate data, particularly solar radiation and light intensity [162].

An essential fact that supports the relation between melatonin and the immune system is the presence of melatonin receptors in immune organs and cells. Conversely, lack of exposure to the light-darkness cycle is a strong stimulus for melatonin production [161] and is geographically associated with areas of high MS prevalence. Melatonin secretion undergoes annual rhythms with a zenith in winter and declines to a nadir in the spring [163]. Thus, the fall in melatonin secretion in the spring may account for epidemiological findings revealing a high incidence of relapse of MS in the spring [163].

The main role of melatonin is the regulation of circadian rhythms [164, 165]. Circadian rhythms are controlled by circadian rhythm genes.

Ad3 The performed pooled analysis of datasets from Canada, Great Britain, Denmark and Sweden showed that significantly fewer people with MS were born in November and significantly more of them were born in May [166]. The results were derived from over 42000 patients with MS. This effect is greater in Scotland, where the population prevalence of this disease is highest [166]. The month of one's birth and the risk of MS are related, more so in familial cases, implying climate-related interactions between genes and the environment [166]. The risk factors responsible for the effect of timing of birth must vary seasonally and probably interact with the development of the central nervous system or immune systems, or both. Such interactions may take place during gestation or shortly after the birth in individuals born in the northern countries studied [166].

Circadian rhythms serve to align physiological functions with the environment. The CLOCK-ARNTL heterodimer regulates the transcription of many clock-controlled genes (CCGs), which in turn influence a wide array of physiological functions external to the oscillatory mechanism. Genome-wide array analyses have indicated that ~10% of all expressed genes in any tissue are under circadian regulation [167]. Moreover, CCGs directly control genes encoding regulators of the cell cycle, proliferation and cellular metabolism [167].

The results of our study can contribute to the possible explanation for the latitude effect on the disease, with association in correlation between MS and circadian rhythm genes presented for the first time. In light of our results with regard to the circadian rhythm genes, further investigations underlying the etiology and pathogenesis of MS will be needed. Replication in large samples is required to validate these associated loci.

4. CONCLUSIONS

We have confirmed both of our hypotheses. Our first hypothesis supposed that genetic variability in STAT genes (the STAT3 and STAT5A genes) might be associated with multiple sclerosis. We have found a statistically significant association of genetic variants in the STAT5A and STAT3 genes with MS. A statistically significant difference was supported by haplotype distributions at the STAT5A and STAT3 gene loci. Moreover, correlating the genotype expression results in STAT3 and STAT5A from the same patients has revealed a significant relation between the genotype and expression levels. Correlation of MS-associated SNPs with differences in gene expression in STAT3 and STAT5A may reveal the functional basis for detected associations. We thus provided further evidence for the implication of the STAT system in the pathogenesis of MS.

According to our second hypothesis, the differences in epidemiology of MS related to geographical latitude effect might be associated with chronobiology. The gene variability in key regulators of the circadian rhythm genes (the CLOCK and ARNTL genes) might thus be associated with multiple sclerosis. We found a statistically significant association of the ARNTL gene and the CLOCK gene with MS. These results were supported by inferred haplotypes in both genes. Genetic variability in the ARNTL gene and the CLOCK gene might therefore be associated with multiple sclerosis and could contribute to the possible explanation for the latitude effect on this disease.

Due to the inherent limitations of association and expression studies, further investigations underlying the etiology and pathogenesis of MS will be needed. Replication studies on independent and large cohorts are required to validate association.

5. SUMMARY

Multiple sclerosis (MS) is a serious neurological disorder. There is still a substantial gap (missing heritability) between the estimated heritability and the evidence of contribution of already identified genetic biomarkers. We therefore integrated different levels of genomic data to identify new gene candidates which we tested for association and blood expression. Among the 381 genomic regions and 409 genes, we selected STAT3 and STAT5A for association and expression studies. According to our hypothesis, the differences in gene variability in the selected top gene candidate (STAT3 and STAT5A) genes might be associated with multiple sclerosis. There are still some unexplained epidemiological data associated with the disease, including the notion of association between the gradient of increasing MS risk and the increasing latitude in both hemispheres. We hypothesized that the differences in epidemiology of MS related to the geographical latitude effect might be associated with chronobiology. Thus, the gene variability in circadian rhythms genes (CLOCK and ARNTL) might be associated with multiple sclerosis. Our study was performed on a Caucasian population of 1058 patients with MS from the Central South East Europe and a group of 1040 individuals matched healthy individuals with no family history of MS or any other inflammatory demyelinating disease in terms of ethnicity, age and gender. We also conducted a comparative gene expression study in blood samples obtained from 50 patients and 40 controls. Expression levels were normalized according to the housekeeping gene β-actin and were co-run with each sample. In total, we genotyped 13 SNPs in STAT genes (STAT 3 and STAT5A) and in both key circadian rhythm-regulating genes (Clock and ARNTL). For all analyses carried out in association studies and gene expression studies, we used the software provided by ABI (ABI, Foster City, USA).

We found a statistically significant difference in the allelic distribution of the STAT5A gene ($P=7.5\cdot10^{-4}$), which was particularly strong in patients with RR MS ($P=1.97\cdot10^{-6}$). These results were supported by determined haplotype distributions of the STAT5A gene (haplotype CA ($P=3\cdot10^{-5}$)). A significant association of genetic variants in the STAT3 gene and MS was also found (P=0.025) and was supported by differences in

haplotype distribution (haplotype AGG ($P=2.5 \cdot 10^{-4}$)). Furthermore, we observed significant differences in the STAT3 gene expression in male patients (P=0.01).

We found a statistically significant difference in the allelic distribution of ARNTL in all participants ($P=7.5 \cdot 10^{-5}$) and confirmed our results by means of haplotypes distribution (haplotypes CGG (P=0.004) and TGA (P=0.03)). The significant association in the ARNTL gene was also supported by gene expression in blood in male patients (P=0.002). In addition, we found a statistically significant association of the CLOCK gene with MS (P=0.021) and supported our results by means of haplotypes distribution in the CLOCK gene (haplotype TCAG (P= 0.001)).

In this dissertation study, we provided evidence that genetic variability in the STAT3 and STAT5A genes and the ARNTL and CLOCK genes might be associated with multiple sclerosis.

Keywords: Multiple sclerosis (MS), circadian rhythms genes, STAT genes

5. POVZETEK

Multipla skleroza (MS) je resna avtoimuna nevrološka bolezen. Kljub intenzivnemu raziskovanju vloge genetskih dejavnikov pri nastanku MS, njenega poteka in zdravljenja, večjega dela heritabilnosti pri MS še ne znamo pojasniti. Zaradi zapletenosti klinične slike bolezni in velike količine genskih podatkov je težko razložiti razpoložljive informacije. Povezali smo različne genomske podatke, ki so pokazali pomembno dovzetnost do MS na izraženih genomskih regijah za določitev top kandidatnih genov. Testirali smo jih v asociacijski študiji, kot tudi z ekspresijo izbranih genov v krvi. Med 381 genomskimi regijami in 409 geni smo izbrali STAT3 in STAT5A za asociacijsko študijo in ekspresijo genov v krvi. Domnevali smo, da bi razlika v genski variabilnosti v top izbranih genih (v STAT3 in STAT5A genu) lahko bila povezana z dovzetnostjo do MS. Vsemu trudu navkljub še vedno manjka neka povezava med dejavniki tveganja in MS, ki bi lahko pojasnila nekatere nepravilnosti, povezane z boleznijo, kot na primer, da je MS pogostejša pri ljudeh, ki živijo dlje od ekvatorja. Obstaja namreč geografski vzorec frekvence bolezni, ki narašča z oddaljenostjo od ekvatorja v smeri obeh polobel. Razlike v epidemiologiji MS v zvezi z zemljepisno širino bi lahko bile povezane s kronobiologijo, zato smo domnevali, da je genska variabilnost v dveh glavnih regulatorjih cirkadialnega ritma (bio ritma) v genih ARNTL in CLOCK povezana s tveganjem za nastanek multiple skleroze. Študija je bila izvedena s 1058 pacienti kavkazijske rase z MS iz osrednje jugo-vzhodne Evrope. Kontrolna skupina je bila sestavljena iz 1040 etnično, starostno in po spolu usklajene populacije, ki v družini ni imela multiple skleroze ali druge demielizacijske bolezni. Za gensko ekspresijo v krvi smo testirali 50 pacientov z MS in 40 zdravih kontrol. Nivoji genske ekspresije so bili usklajeni z notranjo endogeno kontrolo (housekeeping gene) β-actinom, ki je bila usklajevana z vsakim vzorcem posebej. Skupaj smo pri študiji STAT genov (STAT3 in STAT5A) in pri študiji bioritmičnih genov (CLOCK in ARNTL genu) testirali 13 SNPjev (polimorfizmov). Za vse analize asociacije, kot tudi genske ekspresije, smo uporabljali SDS program, dobavljen s strani podjetja ABI (ABI, Foster City, ZDA).

Statistično pomembno razliko pri STAT5A smo našli pri vseh pacientih ($P=7{,}5\cdot10^{-4}$) in pri pacientih z RR obliko MS ($P=1{,}97\cdot10^{-6}$), hkrati pa smo to pomembno razliko potrdili tudi z distribucijo haplotipov pri STAT5A genu (haplotip CA $P=3\cdot10^{-5}$). Statistično pomembno razliko smo našli tudi pri STAT3 genu (P=0,025) in jo potrdili s haplotipi (haplotip AGG $P=2{,}5\cdot10^{-4}$). Statistično pomembna razlika je bila potrjena z ekspresijo v krvi pri moških pacientih (P=0,01).

Statistično pomembno razliko smo našli pri ARNTL genu ($P=7{,}5\cdot10^{-5}$) in jo potrdili z distribucijo pri haplotipih (haplotip CGG (P=0,004) in haplotipu TGA (P=0,03)). Ta pomembna razlika je bila potrjena tudi z ekspresijo v krvi pri moških pacientih (P=0,002). Statistično pomembno razliko smo našli tudi pri CLOCK genu (P=0,021) in jo potrdili z distribucijo pri haplotipih (haplotip TCAG (P= 0,001)).

V tej disertaciji smo pokazali, da bi lahko genska variabilnost pri STAT3 in STAT5A genu ter ARNTL in CLOCK genu lahko bila povezana z nastankom multiple skleroze.

Ključne besede: multipla skleroza (MS), bioritmični geni, STAT geni

6. REFFERENCES

[1] Amedei A., Prisco D. and D'Elios M.M. Multiple sclerosis: the role of cytokines in pathogenesis and in therapies. International journal of molecular sciences, 13, 10 (2012) 13438–13460

[2] Ebers G.C. and Sadovnick A.D. The geographic distribution of multiple sclerosis: A review. Neuroepidemiology 12 (1993)1–5

[3] Goris A.; Epplen C.; Fiten P.; Andersson M.; Murru R.; Sciacca F.L.; Ronsse I.; Jäckel S.; Epplen J.T.; Marrosu M.G.; et al. Analysis of an IFN-γ gene (IFNG) polymorphism in multiple sclerosis in Europe: Effect of population structure on association with disease. J. Interferon Cytokine Res. 19 (1999) 1037–1046

[4] Oksenberg J.R. and Barcellos L.F. Multiple sclerosis genetics: leaving no stone unturned. Genes and immunity, 6, 5 (2005) 375–387

[5] Murray T.J. Multiple Sclerosis, the History of a Disease; Demos Medical Publishing: New York, NY, USA, 2005

[6] Compston A.; McDonald I.; Noseworthy J.; Lassmann H.; Miller D.; Smith K.; Wekerle H.; Confavreux C. McAlpine's Multiple Sclerosis, 4th ed.; Churchill Livingstone: London, UK, 2005

[7] Rejdak K., Jackson S. and Giovannoni G. Multiple sclerosis: a practical overview for clinicians. British medical bulletin 95 (2010) 79–104

[8] Achiron A., Gurevich M., Magalashvili D., Kishner I., Dolev M. and Mandel M. Understanding autoimmune mechanisms in multiple sclerosis using gene expression microarrays: treatment effect and cytokine-related pathways. Clin. Dev. Immunol. 11 (2004) 299-305

[9] Stüve O. and Oksenberg J. Multiple Sclerosis Overview, 2006; Bookshelf ID: NBK1316 PMID: 20301492

[10] Compston A and Coles A. Multiple sclerosis. Lancet 359, 9313 (2002) 1221–1231

[11] Compston A and Coles A. Multiple sclerosis. Lancet 372, 9648 (2008) 1502–1517

[12] Willer C.J., Dyment D.A., Risch N.J., Sadovnick A.D., Ebers G.C., the Canadian Collaborative Study Group: Twin concordance and sibling recurrence rates in multiple sclerosis. Proc. Natl. Acad. Sci. (USA) 100 (2003) 12877-12882

[13] Dyment D.A., Herrera B.M., Cader Z., et al: Complex interactions among MHC haplotypes in multiple sclerosis: susceptibility and resistance. Hum. Mol. Genet. 14 (2005) 2019-2026

[14] Hafler D.A., Compston A., Sawcer S., et al: Risk alleles for multiple sclerosis identified by a genome wide study. N. Engl. J. Med. 357 (2007) 851-862

[15] Ramagopalan S.V., Anderson C., Sadovnick A.D., Ebers G.C. Genomewide study of multiple sclerosis. N. Engl. J. Med. 357 (2007) 2199-2200

[16] De Jager P.L., Jia X., Wang J., de Bakker P.I., Ottoboni L., Aggarwal N.T., Piccio L., Raychaudhuri S., Tran D., Aubin C., et al. International MS Genetics Consortium. Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. Nat. Genet. 41(2009)776–782

[17] Katsavos S. and Anagnostouli M. Biomarkers in Multiple Sclerosis: An Up-to-Date Overview. Multiple sclerosis international, 2013, ii (2013)340508

[18] Schmidt H., Williamson D. and Ashley-Koch A.HLA-DR15 haplotype and multiple sclerosis: a HuGE review. American Journal of Epidemiology 165, 10 (2007)1097–1109

[19] Oksenberg J.R., Barcellos L.F. and Creeetal B.A.C. Mapping multiple sclerosis susceptibility to the HLA-DRlocus in African Americans. American Journal of Human Genetics 74, 1 (2004)160–167

[20] Bozikas V.P., Anagnostouli M.C., Petrikis P. et al. Familial bipolar disorder and multiple sclerosis: a three-generation HLA family study. Progress in Neuro-Psychopharmacology and Biological Psychiatry 27, 5 (2003)835–839

[21] Al-Din A.S.N., El-Khateeb M., Kurdi A. et al. Multiple sclerosis in Arabs in Jordan. Journal of the Neurological Sciences 131, 2 (1995)144–149

[22] Kwon O.J., Karni A., Israel S. et al., HLA class II susceptibility to multiple sclerosis among Ashkenazi and non-Ashkenazi Jews. Archives of Neurology 56, 5 (1999) 555–560

[23] Goodin D.S. The genetic basis of multiple sclerosis: a model for MS susceptibility. BMC Neurology 10 (2010) 101.

[24] Wellcome Trust Case Control Consortium (WTCCC) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature, 447 (2007) 661–678

[25] Tajouri L., Fernandez F. and Griffiths L.R. Gene Expression Studies in Multiple Sclerosis. Current. Genomics, 8 (2007) 181-189

[26] Alonso A. and Hernán M.A. Temporal trends in the incidence of multiple sclerosis: a systematic review. Neurology 71, 2 (2008) 129–135

[27] Ascherio A. and Munger K.L. Environmental risk factors for multiple sclerosis. Part II: Noninfectious factors. Ann. Neurol. 61, 6 (2007) 504–513

[28] Ascherio A., Munger K.L. and Simon K.C. Vitamin D and multiple sclerosis. Lancet Neurol 9,6 (2010) 599–612

[29] Simpson S. Jr., Jr, Blizzard L., Otahal P., Van Der Mei I. and Taylor B. Latitude is significantly associated with the prevalence of multiple sclerosis : a meta-analysis. J Neurol Neurosurg Psychiatry 82 (2011) 1132-1141

[30] Španinger K., Košir R., Fink M., Debeljak N., and Rozman D. Circadian rhythm in humans. Zdrav. Vestn. 78 (2009) 651–657

[31] Kutzelnigg A, Lucchinetti C.F., Stadelmann C., Brück W., Rauschka H., Bergmann M., Schmidbauer M., Parisi J.E. and Lassmann H. Cortical demyelination and diffuse white matter injury in multiple sclerosis. Brain 128, 11 (2005) 2705–2712

[32] Noseworthy J.H, Lucchinetti C., Rodriguez M. and Weinshenker B.G. Multiple sclerosis. N.Engl.J.Med 2000, 343,13 (2000) 938–952

[33] McDonald W. I., Compston A., Edan G., Goodkin D., Hartung H. P., Lublin F. D., et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. Ann. Neurol. 50 (2001)**,** 121–127Revieved by: - Polman, C.H. et al., 2005. Diagnostic Criteria for Multiple Sclerosis : 2005 Revisions to the " McDonald Criteria ". , 11840–846; - and Polman, C. H., Reingold, S. C., Banwell, B., Clanet, M., Cohen, J. a, Filippi, M., Fujihara, K., et al. (2011). Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Annals of neurology, 69, 2(2005) 292–302

[34] Zuvich R.L., McCauley J.L., Pericak-Vance M.A. and J.Haines J.L. Genetics and Pathogenesis of Multiple Sclerosis. Semin Immunol., 21, 6 (2010)328–333

[35]http://www.nationalmssociety.org/about-multiple-sclerosis/what-we-know-about ms/diagnosing-ms/cis/index.aspx 23.02.2013

[36] Miller D., Barkhof F., Montalban X., Thompson A. and Filippi M. Clinically isolated syndromes suggestive of multiple sclerosis, part I: natural history, pathogenesis, diagnosis, and prognosis. The Lancet Neurology, 4, 5(2005) 281–288

[37] Hawkins S.A. and McDonnell G.V. Benign multiple sclerosis ? Clinical course, long term follow up, and assessment of prognostic factors. J Neurol Neurosurg Psychiatry 67(1999)148–152

[38] Confavreux C., Vukusic S., Moreau T. and Adeleine P. Relapses and progression of disability in multiple sclerosis. N.Engl.J.Med 343, 20 (2000)1430–1438

[39] Trapp B.D. and Nave K.A. Multiple sclerosis: an immune or neurodegenerative disorder? Annu.Rev.Neurosci 31(2008) 247–269

[40] Kurtzke J.F. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology.33,11 (1983)1444-1452

[41] Lesko J.L. and Atkinson A.J. Use of biomarkers and surrogate endpoints in drug development and regulatory decision making: criteria, validation, strategies. Annual Review of Pharmacology and Toxicology 41 (2001) 347–366

[42] Kinter J., Zeis T. and Schaeren-Wiemers N. RNA Profiling of MS Brain Tissues. The International MS Journal 15 (2008) 51–58.

[43] Tang Y., Xu H., Du X., et al. Gene expression in blood changes rapidly in neutrophils and monocytes after ischemic stroke in humans: a microarray study. J Cereb. Blood Flow. Metab. 26 (2006) 1089–1102

[44] Tang Y., Schapiro M.B., Franz D.N., et al. Blood expression profiles for tuberous sclerosis complex 2, neurofibromatosis type 1 and Down’s syndrome. Ann. Neurol. 56 (2004) 808–814

[45] Tang Y., Gilbert D.L., Glauser T.A., Hershey A.D., Sharp F.R. Blood gene expression profiling of neurologic diseases: a pilot microarray study. Arch. Neurol. 62 (2005) 210–215

[46] Scherzer C.R., Eklund A.C., Morse L.J., et al. Molecular markers of early Parkinson's disease based on gene expression in blood. Proc. Natl. Acad. Sci. USA 104 (2007) 995–960

[47] Borovecki F., Lovrecic L., Zhou J., et al. Genome-wide expression profiling of human blood reveals biomarkers for Huntington's disease. Proc. Natl. Acad. Sci. USA 102 (2005) 11023– 11028

[48] Sullivan P.F., Fan C. and Perou C.M. Evaluating the comparability of gene expression in blood and brain. Am. J Med. Genet. B. Neuropsychiatr. Genet. 141, 5 (2006) 261-268

[49] Kermode A.G., Thompson A.J., Tofts P., et al. Breakdown of the blood-brain barrier precedes symptoms and other MRI signs of new lesions in multipke sclerosis: pathogentic and clinical implications. Brain 113 (1990) 1477–1489

[50] Chataway J., Feakes R., Coraddu F., Gray J., Deans J., Fraser M., Robertson N., Broadley S., Jones H., Clayton D.,Goodfellow P., Sawcer S., and Compston A. The genetics of multiple sclerosis: principles, background, and updated results of the United Kingdom systematic genome screen. Brain 121 (1998) 1869-1887

[51] Sadovnick A.D., Armstrong H., Rice G.P., Bulman D., Hashimoto L., Paty D.W., et al. A population based study of multiple sclerosis in twind: update. Ann. Neurol. 33 (1993) 281-285

[52] Robertson N.P., Fraser M., Deans J., Clayton D., Walker N. And Compston D.A. Age adjusted recurencce risk for relatives of patients with multiple sclerosis. Brain 119 (1996) 449-455

[53] Sadovnick A.D., Ebers G.C., Dyment D.A., Risch N.J. and the Canadian Collaborative Study Group. Evidence for genetic basis of multiple sclerosis: Lancet 347 (1996) 1728-1730

[54] Robertson N.P., O' Riordan J.I., Chataway J., Kingsley D.P., Miller D.H, Clayton D., et al. Offspring recurrence rates and clinical characteristics of conjugal multiple sclerosis. Lancet 349 (1997) 1587-1590

[55] Ebers G.C., Sadovnick A.D., Risch N.J. and The Canadian Collaborative Study Group. A genetic basis for familial aggregation in multiple sclerosis. Nature 377 (1995) 150-151

[56] Doolitle T.H., Myers R.H., Lenrich J.R., Birnbaum G., Sheremata W., Franklin G.M., et al. Multiple sclerosis sibling pairs : clustered onset and familial predisposition. Neurology 40 (1990) 1546-1552

[57] Bulman D.E., Sadovnick A.D. and Ebers G.C. Age of onset in siblings concordant for multiple sclerosis. Brain 114 (1991) 937-950

[58] http://www.omim.org/

[59] Strachan T. and Read A. Human Molecular Genetics (4 ed.). Garland Science. 467–495

[60] Altmüller J., Palmer L.J., Fischer G., Scherb H. and Wjst M. Genomewide Scans of Complex Human Diseases: True Linkage Is Hard to Find". Am. J. Hum. Genet. 69, 5 (2001) 936–950

[61] Risch N. and Merikangas K. (September 1996).The future of genetic studies of complex human diseases. Science 273, 5281 (1996) 1516–1517

[62] Zhernakova A., van Diemen C. and Wijmenga C. Detecting shared pathogenesis from the shared genetics of immune-related diseases. Nat. Rev. Genet. 10 (2009) 43–55

[63] Sánchez-Pla A., Reverter F., Ruíz de Villa C.M. and Comabella M. Transcriptomics: mRNA and alternative splicing. Journal of Neuroimmunology 248,1 (2012) 23-31

[64]http://www.hugenavigator.net/HuGENavigator/phenoPedia.do?firstQuery=Multiple Sclerosis

[65] Lill C.M., Roehr J.T., McQueen M.B., Bagade S., Schjeide B.M., Zipp F. and Bertram L. The MSGene Database. Alzheimer Research Forum. Available at http://www.msgene.org/. Accessed on [01.03.2013]

[66] The International Multiple Sclerosis Genetics Consortium,The Wellcome Trust Case Control Consortium2, “Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis,” Nature 476,7359 (2011) 214–219

[67] Frohman E.M., Filippi M., Stuve O., Waxman S.G., Corboy J., Phillips J.T., Lucchinetti C., Wilken J., Karandikar N., Hemmer B. et al. Characterizing the mechanisms of progression in multiple sclerosis: evidence and new hypotheses for future directions. Arch. Neurol. 62 (2005) 1345–1356

[68] International Multiple Sclerosis Genetics Consortium, “Risk alleles for multiple sclerosis identified by a genomewide study,” The New England Journal of Medicine 357, 9 (2007) 851– 862

[69] D'Netto M.J., Ward H., Morrison K.M., Ramagopalan S.V., Dyment D.A., DeLuca G.C., Handunnetthi L., Sadovnick A.D. and Ebers G.C. Risk alleles for multiple sclerosis in multiplex families. Neurology 72, 23 (2009)1984–1988

[70] Kemppinen A., Sawcer S. and Compston A. Genome-wide association studies in multiple sclerosis: lessons and future prospects. Briefings in functional genomics 10, 2 (2011) 61-70

[71] Morley M., Molony, C., Weber T., Devlin J., Ewens K., Spielman R. and Cheung V. Genetic analysis of genome-wide variation in human gene expression. Nature 430 (2004) 743–747

[72] Dixon A., Liang L., Moffatt M., Chen W., Heath S., Wong K., Taylor J., Burnett E., Gut I., Farrall, M. et al. A genome-wide association study of global gene expression. Nat. Genet. 39 (2007) 1202–1207

[73] Stranger B., Forrest M., Dunning M., Ingle C., Beazley C., Thorne N., Redon R., Bird C., de Grassi A., Lee, C. et al. Relative impact of nucleotide and copy number variation on gene expression phenotypes. Science 315 (2007) 848–853

[74] Goring H., Curran J., Johnson M., Dyer T., Charlesworth J., Cole S., Jowett J., Abraham L., Rainwater D., Comuzzie, A. et al. Discovery of expression QTLs using large-scale transcriptional profiling in human lymphocytes. Nat. Genet. 39 (2007) 1208–1216

[75] Heap, G.A., Trynka, G., Jansen, R.C., Bruinenberg, M., Swertz, M.A., Dinesen, L.C., Hunt, K.A., Wijmenga, C., Vanheel, D.A. and Franke, L. Complex nature of SNP genotype effects on gene expression in primary human leucocytes. BMC Med. Genom. 2 (2009) 1

[76] Wang Z., Gerstein M. and Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nature Reviews Genetics, 10,1 (2009) 57-63

[77] Omar de F. Jr., Moore C.S., Kennedy T.E., Antel J.P., Bar O.A. and Dhaunchak A.S.MicroRNA dysregulation in multiple sclerosis. Frontiers in genetics 3 (2012) 311

[78] Zhou L., Ivanov I.I., Spolski R., Min R., Shenderov K., Egawa T., Levy D.E., Leonard W.J. and Littman D.R. IL-6 programs T (H)-17 cell differentiations by promoting sequential engagement of the IL-21 and IL-23 pathways.Nat. Immunol. 8, 967 (2007)

[79] Boeuf H., Hauss C., Graeve F.D., Baran N. and Kedinger C. Leukemia inhibitory factor-dependent transcriptional activation in embryonic stem cells. J. Cell Biol. 138 (1997) 1207–1217

[80] Niwa H., Burdon T., Chambers I., and Smith A. Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. Genes Dev. 12 (1998) 2048–2060

[81] Raz R., Lee C.K., Cannizzaro L.A., d'Eustachio P., and Levy D.E. Essential role of STAT3 for embryonic stem cell pluripotency. Proc. Natl. Acad. Sci. USA. 96 (1999) 2846–2851

[82] Curtis R.K., Oresic M, Vidal-Puig A. Pathways to the analysis of microarray data. Trends Biotechnol. 23 (2005) 429–435

[83] Aerts S., Lambrechts D., Maity S., Van Loo P., Coessens B., De Smet F., Tranchevent L.C., De Moor B., Marynen P., Hassan B., Carmeliet P. and Moreau Y. Gene prioritization through genomic data fusion. Nat. Biotechnol. 24 (2006) 537–544

[84] Rasche A., Al-Hasani H. and Herwig R. Meta-analysis approach identifies candidate genes and associated molecular networks for type-2 diabetes mellitus. BMC Genomics, 9 (2008) 310

[85] Sun J., Jia P., Fanous A.H., Webb B.T., van den Oord E.J., Chen X., Bukszar .J, Kendler K.S., and Zhao Z. A multi-dimensional evidence-based candidate gene prioritization approach for complex diseases-schizophrenia as a case. Bioinformatics 25 (2009) 2595–6602

[86] Middleton F.A., Rosenow C., Vailaya A., Kuchinsky A., Pato M.T., Pato C.N. Integrating genetic, functional genomic, and bioinformatics data in a systems biology approach to complex diseases: application to schizophrenia. Methods Mol. Biol. 401 (2007) 337–364

[87] Maver A., Peterlin B. Positional integratomic approach in identification of genomic candidate regions for Parkinson's disease. Bioinformatics 27, 14 (2011) 1971-1978

[88] Buettner R., Mora L.B. and Jove R. Activated STAT Signaling in Human Tumors Provides Novel Molecular Targets for Therapeutic. Clin Cancer Res. 8,4 (2002) 945-954

[89] Horvath C.M.Review STAT proteins and transcriptional responses to extracellular signals. TIBS 25 (2000) 496–502

[90] Darnell Jr. J.E., Kerr I.M. and Stark G.R. Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins Science, 264(1994)1415 -1421

[91] Stark G.R., Kerr I.M., Williams B.R., Silverman R.H. and Schreiber R.D. How cells respond to interferons. Annu. Rev. Biochem., 67 (1998) 227- 264

[92] Darnell Jr J.E. STATs and gene regulation. Science, 277 (1997)1630 -1635

[93] Egwuagu C.E. STAT3 in CD4þ T helper cell differentiation and inflammatory diseases. Cytokine 47(2009)149–156

[94] Heinrich P.C., Behrmann I., Haan S.,Hermanns H. M., Muller-Newen G. And Schaper F. Principles of interleukin (IL)- 6-type cytokine signalling and its regulation.Biochemical J. 374(2003)1-20

[95] Rane S.G. and Reddy E.P. Janus kinases: components of multiple signaling pathways. Oncogene 19 (2000) 5662-5679

[96] Shuai K. Modulation of STAT signaling by STAT-interacting proteins. Oncogene 19 (2000) 2638-2644

[97] Veldhoen M., Hocking R.J., Atkins C.J., Locksley R.M. and Stockinger B., TGF beta in the context of an inflammatory cytokine milieu supports de novo differentiation of IL-17-producing T cells Immunity 24 (2006)179-189

[98] Yang X.O., Pappu B.P., Nurieva R., Akimzhanov A., Kang H.S., Chung Y., Ma L., Shah B., Panopoulos A.D., Schluns K.S., Watowich S.S., Tian Q., Jetten A.M. and Dong C. T helper 17 lineage differentiation is programmed by orphan nuclear receptors ROR alpha and ROR gamma Immunity 28 (2008) 29-39

[99] http://ifatwww.et.uni-magdeburg.de/syst/systembio/research/index.shtml

[100] Azam M., Erdjument-Bromage H., Kreider B.L., Xia M., Quelle F., Basu R., Saris, C., Tempst P., Ihle J.N. and Schindler C. Interleukin-3 signals through multiple isoforms of Stat5.EMBOJ 14 (1995)1402– 1411

[101] Mui A., Wakao H., O'Farrell A.M., Harada N. and Miyajima A. Interleukin-3, granulocyte-macrophage colony stimulating factor and interleukin-5 transduce signals through two Stat5 homologues. EMBO J 14 (1995) 1166–1175

[102] Schindler C. and Strehlow I. Cytokines and STAT signaling. Adv. Pharmacol. 47 (2000)113–174

[103] Dhennin-Duthille I., my Nyga R., Yahiaoui S., Gouilleux-Gruart V., Re´gnier A., Lassoued K., and Gouilleux F. The tumor suppressor hTid1 inhibits STAT5b activity via functional interaction. The Journal of biological chemistry. 286, 7 (2011) 5034–5042

[104] Benekli M., Baer M. R., Baumann H. and Wetzler M. Signal transducer and activator of transcription proteins in leukemia's. Blood 101(2003) 2940–2954

[105] Tan S. H. and Nevalainen M. T. Signal transducer and activator of transcription 5A/B in prostate and breast cancers. Endocr. Relat. Cancer 15 (2008) 367–390

[106] Grimley P.M., Dong F. and Rui H. Stat5a and Stat5b: fraternal twins of signal transduction and transcriptional activation. Cytokine Growth Factor Rev. 10, 2(1999), 131-157

[107] Turlej R.K., Fiévez L., Sandersen C.F., Dogné S., Kirschvink N., Lekeux P. and Bureau F. Enhanced survival of lung granulocytes in an animal model of asthma: evidence for a role of GM-CSF activated STAT5 signalling pathway. Thorax. 56, 9(2001) 696-702

[108] Yu H.H., Liu P.H., Lin Y.C., Chen W.J., Lee J.H., Wang L.C., Yang Y.H. and Chiang B.L. Interleukin 4 and STAT6 gene polymorphisms are associated with systemic lupus erythematosus in Chinese patients. Lupus. 19,10 (2010) 1219-1228

[109] Nagy Z.S., Rui H., Stepkowski S.M., Karras J. andKirken R.A. A preferential role for STAT5, not constitutively active STAT3, in promoting survival of a human lymphoid tumor. J Immunol. 177, 8 (2006) 5032-5040

[110] Hassel J.C., Winnemöller D., Schartl M. and Wellbrock C. STAT5 contributes to antiapoptosis in melanoma. Melanoma Res.18, 6 (2008) 378-385

[111] Sultan A.S., Xie J., LeBaron M.J., Ealley E.L., Nevalainen M.T. and Rui H. Stat5 promotes homotypic adhesion and inhibits invasive characteristics of human breast cancer cells. Oncogene. 27, 24/5 (2005) 746-760

[112] Alonso A. and Hernán M.A. Temporal trends in the incidence of multiple sclerosis: a systematic review. Neurology 71, 2 (2008) 129–135

[113] Ciarleglio C.M., Ryckman K., Servick S.V., Hida A., Wells N. et al. Genetic Differences in Human Circadian Clock Genes Among Worldwide Populations. J. Biol Rhythms. 23,4 (2008) 330-340

[114] Alberts B., Johnson A., Lewis J., Raff M., Roberts K. And Walter P. Molecular biology of the cell. Fifth edition. Garland Science, Taylor&Francis Group, LLC, New York, USA 2008. Chapter 7: Control of gene expression. page 460-462

[115] Vieira E., Nilsson E.C., Nerstedt A., Ormestad M., Long Y.C., Garcia-Roves P.M., Zierath J.R. and Mahlapuu M. Relationship between AMPK and the transcriptional balance of clock-related genes in skeletal muscle. Am J Physiol Endocrinol Metab 295 (2008) E1032–E1037

[116] Lawrey P.L. and Takahashi J.S. Mammalian circadian biology: elucidating genome-wide levels of temporal organization. Annu. Rev. Genomics Hum. Genet. 5 (2004) 407-441

[117] Reppert S.M. and Weaver D.R. Coordination of circadian timing in mammals. Nature 418 (2002) 935-941

[118] Guillaumond F., Dardente H., Giguere V. and Cermakian N. Differntial control of Bmal1 circadian transcription by REV-ERB and ROR nuclear receptors. J. Biol. Rhythms 20 (2005) 391-403

[119] Son G.H., Chung S. and Kim K. Biological Rhythms and Neuroendocrine Systems Endocrinol Metab. 25, 4 (2010)249-257

[120] Lowrey P.L. and Takahashi J.S. Mammalian Circadian Biology: Elucidating Genome-Wide Levels of Temporal Organization. Annual Review of Genomics and Human Genetics 5 (2004) 407-441

[121] Hedström A. K., Åkerstedt T., Hillert J., Olsson T. and Alfredsson, L. Shift work at young age is associated with increased risk for multiple sclerosis. Ann Neurol., 70 (2011) 733-741

[122] Meijer J.H., Watanabe K., Detari L. and Schaap J. Circadian rhythm in light response in suprachiasmatic nucleus neurons of freely moving rats. Brain Res 741 (1996) 352-355

[123] Meijer J.H. and Schwartz W.J. In search of the pathways for light induced pacemaker resetting in the suprachiasmatic nucleus. J Biol Rhythms 18 (2003) 235-249

[124] Pevet P., Jacob N., Lakhdar-Ghazal N. and Vuillez P. How do the suprachiasmatic nuclei of the hypothalamus integrate photoperiodic information? Biol Cell 89 (1997) 569-577

[125] Sahar S. and Sassone-Corsi P. Metabolism and cancer: the circadian clock connection. Nat. Rev. Cancer 9 (2009) 886-896

[126] Bechtold D.A., Gibbs J.E. and Loudon A.S. Circadian dysfunction in disease. Trends Pharmacol Sci 31 (2010) 191-198

[127] Jakkula E., Leppa V., Sulonen A.-M., Varilo T., Kallio S., Kemppinen A., Purcell S., Koivisto K., Tienari P., Sumelahti M.-L., Elovaara I., Pirttila T., et al. Genome-wide association study in a high-risk isolate for multiple sclerosis reveals associated variants in STAT3 gene. Am. J. Hum. Genet. 86 (2010)285-291

[128] Baranzini, S.E. et al. Pathway and network-based analysis of genome-wide association studies in multiple sclerosis. Human molecular genetics 18, 11 (2009) 2078–2090

[129] Gandhi K.S., McKay F.C., Cox M., Riveros C., Armstrong N., Heard R.N.,Vucic S.,Williams D.W., Stankovich J., Brown M., Danoy P., Stewart G.J., Broadley S., Moscato P., Lechner-Scott J., Scott R.J., and Booth D.R. The multiple sclerosis whole blood mRNA transcriptome and genetic associations indicate dys regulation of specific T cell pathways in pathogenesis. Human Molecular Genetics, 19, 11 (2010) 2134–2143

[130] Keller A. et al. Multiple sclerosis: microRNA expression profiles accurately differentiate patients with relapsing-remitting disease from healthy controls. PloS one 4,10 (2009) p.e7440

[131] Junker A. et al. MicroRNA profiling of multiple sclerosis lesions identifies modulators of the regulatory protein CD47. Brain : a journal of neurology, 132 (2009) 3342–3352

[132] Barry S.P. and Kavanagh P.Investigating the role of the JAK / STAT and MAPK Pathways in Ischaemia / Reperfusion Injury and Inflammation. (2009)

[133] Whitney A.R., Diehn M., Popper S.J., Alizadeh A.A., Boldrick J.C., Relman D.A. and Brown P.O. Individuality and variation in gene expression patterns in human blood. Proc. Natl Acad. Sci. USA, 100 (2003)1896–1901

[134] de Bakker P.I., Yelensky R., Pe'er I., Gabriel S.B., Daly M.J. and Altshuler D. Efficiency and power in genetic association studies. Nat. Genet. 37, 11 (2005) 1217-1223

[135] Schaid D.J., Rowland C.M., Tines D.E., Jacobson R.M. and Poland G.A. Score tests for association between traits and haplotypes when linkage phase is ambiguous, American journal of human genetics 70 (2002) 425-434

[136] Livak K.J. and Schmittgen T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C (T)) Method. Methods 25, 4 (2001) 402-408

[137] Patsopoulos N.A., BSP MS Genetics working group and the steering committees of studies evaluating IFNb-1b and a CCR1-antagonist, ANZgene Consortium, GeneMSA, International Multiple Sclerosis Genetics Consortium, and Paul I.W. de Bakker. Ann Neurol 70, 6(2012) 897–912

[138] Lill C. M., Schjeide B.M. M., Akkad D. A., Blaschke P., Winkelmann A., Gerdes L.-A., Hoffjan S., Luessi F., Dorner T., Li S. C., Steinhagen-Thiessen E., Lindenberger U., et al. Independent replication of STAT3 association with multiple sclerosis in a large German case-control sample**.** Neurogenetics 13 (2012) 83-86

[139] Stritesky G.L., Muthukrishnan R., Sehra S., Goswami R., Pham D., Travers J., Nguyen E.T., Levy D.E. and Kaplan M.H. The transcription factor STAT3 is required for T helper 2 cell development. 34,1 (2012) 39–49

[140] Liu X., Lee Y.S., Yu C.R and Egwuagu C.E. Loss of STAT3 in CD4+ T cells prevents development of experimental autoimmune diseases. J Immunol. 180, 9 (2008) 6070–6076

[141] Levy D.E. and Lee C. What does Stat3 do ? J. Clin. Invest 109, 9 (2002) 1143–1148

[142] Qin H., Yeh W.I., De Sarno P., Holdbrooksa A.T., Liua Y., Muldowneya M.T., Reynoldsa S.L., Yanagisawaa L.L., Thomas H. Fox T.H., Parka K., Harringtona L.E., Ramanc C. and Benvenistea E.N. Signal transducer and activator of transcription-3/suppressor of cytokine signaling-3 (STAT3/SOCS3) axis in myeloid cells regulates neuroinflammation PNAS 109 , 13 (2012) 5004–5009 |

[143] Cénit M.C., Alcina A., Márquez A., Mendoza J.L., Díaz-Rubio M., de las Heras V., Izquierdo G., Arroyo R., Fernández O., de la Concha E.G., Matesanz F. and Urcelay E. STAT3 locus in inflammatory bowel disease and multiple sclerosis susceptibility. Genes and Immunity 11 (2010) 264-268

[144] Kornfeld J. et al., . Europe PMC Funders Group. The different functions of Stat5 and chromatin alteration through Stat5 proteins. Front Biosci. 13 (2010) 6237–6254

[145] Noorbakhsh F., Ellestad K. K., Maingat F., Warren K. G., Han M. H., Steinman L., et al. Impaired neurosteroid synthesis in multiple sclerosis. Brain 134 (2011) 2703–2721

[146] Dolznig H., Grebien F., Deiner E.M., Stangl K., Kolbus A., Habermann B., Kerenyi M.A., Kieslinger M., Moriggl R., Beug H. and Mullner E.W. Erythroid progenitor renewal versus differentiation: genetic evidence for cell autonomous, essential functions of EpoR, Stat5 and the GR. Oncogene. 25 (2006) 2890–2900

[147] Hoelbl A., Kovacic B., Kerenyi M.A., Simma O., Warsch W., Cui Y., Beug H., Hennighausen L., Moriggl R. and Sexl V. Clarifying the role of Stat5 in lymphoid development and Abelson-induced transformation. Blood.107(2006) 4898–4906

[148] Kieslinger M., Woldman I., Moriggl R., Hofmann J., Marine J.C., Ihle J.N., Beug H. and Decker T. Antiapoptotic activity of Stat5 required during terminal stages of myeloid differentiation. Genes Dev. 14 (2000) 232–244

[149] Yao Z., Kanno Y., Kerenyi M.A., Stephens G., Durant L., Watford W.T., Laurence A., Robinson G.W., Shevach E.M., Moriggl R., Hennighausen L., Wu C. and O'Shea J.J. Nonredundant roles for Stat5a/b in directly regulating Foxp3. Blood.109 (2007) 4368–4375

[150] Burchill M.A., Goetz C.A., Prlic M., O'Neil J.J., Harmon I.R., Bensinger S.J., Turka L.A., Brennan P., Jameson S.C. and Farrar M.A. Distinct effects of STAT5 activation on CD4+ and CD8+ T cell homeostasis: development of CD4+CD25+ regulatory T cells versus CD8+ memory T cells. J Immunol.171 (2003) 5853–5864

[151] Burchill M.A., Yang J., Vogtenhuber C., Blazar B.R. and Farrar M.A. IL-2 receptor beta-dependent STAT5 activation is required for the development of Foxp3+ regulatory T cells. J Immunol. 178 (2007) 280–290

[152] Moriggl R., Sexl C., Piekorz R., Topham D. and Ihle J.N. Stat5 activation is uniquely associated with cytokine signaling in peripheral T cells. Immunity.11 (1999) 225–230

[153] Knutsson A. Shift work and coronary heart disease. Scand J Soc Med 44(1989):1–36

[154] Gale C.R. and Martyn C.N. Migrant studies in multiple sclerosis. Prog Neurobiol 47 (1995) 425-448[155] Folkard S. and Tucker P. Shift work, safety and productivity. Occup Med 53 (2003) 95–101

[156] Koller M., Kundi M., and Cervinka R. Field studies of shift work at an Austrian oil refinery. I. Health and psychosocial wellbeing of workers who drop out of shift work. Ergonomics 21 (1978)835–847

[157] Knutsson A. In-depth review : shift work. Occupational Medicine 53 (2003)103–108

[158] Hallpike J.F., Adams C.W.M. and Tourtellote W.W. Multiple Sclerosis: Pathology, Diagnosis and Management. London: Chapman and Hall (1983)

[159] Constantinescu C.S. Melanin, melatonin, melanocyte-stimulating hormone, and the susceptibility to autoimmune demyelination: a rationale for light therapy in multiple sclerosis. Medical hypotheses 45 (1995) 455-458

[160] Goldman B.D. Mammalian photoperiodic system: formal properties and neuroendocrine mechanisms of photoperiodic time measurement. J. Biol. Rhythms 16 (2001) 283–301

[161] Utiger R.D. Melatonin: the hormone of darkness. N Engl J Med 327 (1992) 1377-1379

[162] Meier D.S., Balashov K.E., Healy B., Weiner H.L. and Guttmann C.R.G. Seasonal prevalence of MS disease activity. Neurology 75, 9 (2010) 799-806

[163] Sandyk R. and Awerbuch G.I., Multiple Sclerosis: Relationship between seasonal variations of Relapse and age of onset. Int J Neurosci 71, 1-4 (1993) 147-157

[164] Constantinescu C.S. Melanin, melatonin, melanocyte-stimulating hormone, and the susceptibility to autoimmune demyelination: a rationale for light therapy in multiple sclerosis. Medical hypotheses 45 (1995) 455-458

[165] Reiter R.J. Pineal melatonin: cell biology of its synthesis and of its physiological interactions. Endocr. Rev 12 (1991) 151-180

[166] Willer C.J., Dyment D.A., Sadovnick A.D., Rothwell P.M., Murray T.J. and Ebers G.C. Timing of birth and risk of multiple sclerosis: population based study. BMJ 330, 7483 (2005)

[167] Sassone-Corsi P. Commentary: the year in circadian rhythms. Mol Endocrinol. 24, 11 (2010) 2081–2087

ACKNOWLEDGEMENTS:

First and foremost, I would like to express my sincere gratitude to my mentor Prof. Dr. Borut Peterlin for providing me with knowledge, support and suggestions. He was always available to advise me, and has shown patience and in-depth knowledge of the field. I would also like to thank Andrej Stegnar and Lili Simeunović for their invaluable assistance and technical support, as well as Aleš Maver for advice in statistical analyses and his invaluable assistance provided on several occasions. I am also very grateful to all my colleagues from the Division of Medical Genetics at the University Medical Centre (UMC), Ljubljana, Slovenia. Special thanks go to Mojca Gorjanc for all support needed.

I would like to thank the Ministry of Higher Education, Science and Technology, which, in cooperation with the University of Nova Gorica, enabled me to receive European funding for my PhD studies under Contract No. 24-221/2011. They gave me additional support in the last two years of my PhD studies.

Furthermore, a sincere thank you to all participating MS patients and healthy controls for their blood sample donations.

Special thanks go to my parents, Vida and Janez Lavtar, and to my dear friends, Jernej and Franja Jung, for all their support. They gave me the love, extra strength and motivation I needed to get things done, and stood by me in good times and bad.

Thank you from the heart.

I would like to dedicate my PhD thesis in memoriam of Prof. Dr. Mirko Jung, who always said to me: "Polona, all the answers are in genetics!" Thank you, Mirko!